

Universidade Federal do Rio Grande do Norte - UFRN
Centro de Ciências Exatas e da Terra - CCET
Programa de Pós-Graduação em Matemática Aplicada e
Estatística - PPGMAE

Estudo Comparativo de Gráficos de Probabilidade
Normal para Análise de Experimentos Fatoriais
Não Replicados

por

Manasses Pereira Nóbrega

Natal - RN
Maio de 2010

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Estudo Comparativo de Gráficos de Probabilidade Normal para Análise de Experimentos Fatoriais Não Replicados

Manasses Pereira Nóbrega

Dissertação apresentada ao Programa de Pós-Graduação em Matemática Aplicada e Estatística da Universidade Federal do Rio Grande do Norte como requisito parcial para a obtenção do título de Mestre em Matemática Aplicada e Estatística.

Área de Concentração: Métodos Estatísticos

Linha de Pesquisa: Planejamento de Experimentos

Orientadora: Prof^ª PhD Carla Almeida Vivacqua

Natal-RN

Mai de 2010

*Dedico esta dissertação a meus pais, Manoel e Naizia,
por me tratarem como seu presente, por torcerem por
mim e por acreditarem que este trabalho seria
concluído. Dedico também àqueles e àquelas que
poderão porventura, algum dia, fazer uso do aqui
exposto.*

*Inventing a graphical method is easy. Inventing one
that works is difficult. Figuring out if a particular
invention works is more difficult.*

William S. Cleveland

All plots are useful; some are less useful than others.

J. Stuart Hunter

Agradecimentos

Ao Soberano Senhor do Universo, Jeová Deus, o meu Deus, minha fonte de vida, luz e sabedoria. A Ti, Jeová, agradeço pela força dada nos momentos mais difíceis (e nos mais fáceis também), por Tua presença quando a solidão batia devido à saudade de casa, pela disciplina aplicada nos momentos em que dela precisei, e pelo imerecido privilégio de Te conhecer. *“Assim disse Jeová: “Não se jacte o sábio da sua sabedoria, nem se jacte o poderoso da sua potência. Não se jacte o rico das suas riquezas. Mas quem se jacta, jacte-se da seguinte coisa: de ter perspicácia e de ter conhecimento de mim, que eu sou Jeová, Aquele que usa de benevolência, de juízo e de justiça na terra; porque é destas coisas que me agrado”, é a pronúnciação de Jeová.” (Jeremias 9: 23, 24).*

À minha querida família, meu bem visível mais precioso. Aos meus pais, Manoel e Naizia, sem os quais eu não existiria, e cuja força, incentivo e valores me tornaram o que sou hoje. Muito obrigado. Eu os amo demais. À minha irmã Monnizia, pelo exemplo de coragem e determinação que sempre me passou, e pelas boas e duras broncas que me deu durante esse curso. À minha irmã Mellka, pelo incentivo, e principalmente pelo telefonema naquele que seria o dia do meu triste retorno. À minha irmã Mayla, pelas palavras bondosas e pelas orações a meu favor.

À grande amiga Patrícia, minha querida “vizinha”. Os momentos de alegria, tristezas, raivas e de reflexão que passamos juntos foram enriquecedores para mim. Em muito poucos pude observar tanto zelo, fé e amor a Deus como vi em você. Seu exemplo de moral, integridade, honestidade e boa conduta não serão por mim jamais esquecidos. Desculpe pelas raivas e pelas “brigas”; fazem parte do convívio. Que Deus realize os desejos de seu coração. Muito obrigado por tudo.

À querida amiga Renatinha, pelo exemplo de brandura, doçura, paciência e bravura que demonstrou. Obrigado pelas palavras de incentivo e encorajamento, pelas aulas de Inferência que me deu, e pelas ajudas que me prestou, nos bastidores, mesmo sem eu saber. Você mostrou-se uma amiga leal. Foi muito bom e divertido poder ter dividido o mesmo teto que você por algum tempo. Aprendi muito. Te devo muito. Obrigado.

À nossa grande “Renatona”, pelo exemplo de força, fibra e, principalmente, de inteligência. Um dia eu gostaria de saber Estatística tanto quanto você. Obrigado pela ajuda que me prestou e pelas “aulas” ministradas durante o curso. Foi uma honra morar com você.

Aos colegas de mestrado Cecílio, João Saturnino, Jackelia, Lenilson, Neto, Hermes, Allan, Wandell, Moisés, Aparecida e Narinho. A saudade de casa teria sido muito mais difícil de lidar sem a presença de vocês. Obrigado.

A Moisés, pela torcida e pela presença tanto na minha qualificação quanto na defesa.

À nossa grande e querida “xeroqueira” Liandra. Muito obrigado pelos momentos de descontração, pelo incentivo, pelos excelentes trabalhos que você fazia na xerox e pelos grandes favores que me prestou durante o curso. Você é uma excelente profissional, e pessoa também. Muito obrigado.

A todos os professores do PPGMAE. Obrigado pela qualidade no ensino e pelos valores humanos que transmitiram. Aqui eu tive o meu primeiro contato com a Estatística.

Ao professor Damião, pelo exemplo de pessoa que é, pela ajuda e pela paciência demonstradas durante todo o meu curso, bem como pela presença tanto em minha qualificação quanto na defesa. Obrigado, em especial, pelas excelentes contribuições dadas para o Capítulo 3 desta dissertação. Com certeza enriqueceram meu trabalho.

Ao professor André Pinho, pela ajuda com Probabilidade e Inferência, bem como com a programação no R. Suas contribuições permeiam toda esta dissertação, pelas quais sou extremamente grato. Muito do êxito dessa pesquisa devo ao senhor, às suas orientações e incentivos dados durante todo o processo de elaboração. Muito obrigado.

À minha querida orientadora, professora Carla, a grande idealizadora deste trabalho. Muito obrigado por ter me mostrado, e ensinado, esse interessante campo do Planejamento de Experimentos. Obrigado também pela excelente orientação que fez; pela paciência que teve comigo; por ter entendido e respeitado o meu tempo e minhas dificuldades; pela preocupação, interesse e cuidados nos momentos difíceis que enfrentei na parte final deste trabalho; por sempre me incentivar a continuar e a melhorar; e por nunca ter desistido de mim, acreditando que eu poderia chegar onde cheguei. O exemplo que a senhora dá como pessoa, como profissional, sua grande inteligência e humildade são de grande valia para mim. Muito obrigado mesmo, por tudo. Foi um prazer ser seu orientando.

À nossa amada Nadir (em memória), quem tanto falou em ver esse dia chegar, a fim de comemorarmos juntos essa conquista, mas a quem a morte tirou do nós tão cedo e de uma forma tão trágica e violenta. A dor ainda é imensa e a saudade é enorme. Sinto muita falta do seu cheiro carinhoso, do seu abraço aconchegante e de sua voz tão suave. Te amo muito. Nunca esqueceremos de você. Muito obrigado por tudo. Daquela a quem a senhora chamava de *“meu querido . . . o filho que nunca tive.”*

Resumo

Os experimentos fatoriais 2^k são muito utilizados na experimentação industrial. Contudo, quanto maior o número de fatores considerados maior será a quantidade de provas necessárias para a execução de um experimento, e realizar replicações dos tratamentos pode ser inviável, considerando as limitações de recursos e de tempo, tornando tal experimento dispendioso. Nestes casos, são utilizados os fatoriais 2^k não replicados. Mas, sem replicação, não é possível obter uma estimativa direta da variabilidade do erro experimental para se avaliar a significância dos efeitos. Uma das possíveis soluções para este problema é utilizar os gráficos normal ou semi-normal dos efeitos. Muitos pesquisadores usam o gráfico normal, ao passo que outros preferem o semi-normal e, em muitas vezes, para ambos os casos, sem alguma justificativa. A controvérsia sobre o uso destas duas técnicas gráficas é o que motiva a realização do presente trabalho, uma vez que não há registro de procedimento formal ou teste estatístico que indique “qual delas é melhor”. A escolha entre os dois gráficos parece ser uma questão subjetiva. O objetivo central desta dissertação é, então, realizar um estudo comparativo experimental dos gráficos normal e semi-normal no contexto da análise dos experimentos fatoriais 2^k não replicados. Tal estudo consiste na construção de cenários simulados, nos quais o desempenho dos gráficos em detectar os efeitos significativos e identificar valores discrepantes é avaliado com o intuito de verificar as seguintes questões: Um gráfico pode ser melhor que o outro? Em que situações? Que informações um gráfico acrescenta à análise do experimento que possam complementar aquelas fornecidas pelo outro gráfico? Quais as restrições no uso de cada gráfico? Com isso, propõe-se confrontar estas duas técnicas; examiná-las simultaneamente a fim de conhecer semelhanças, diferenças ou relações que possam contribuir para a construção de um referencial teórico que sirva como um subsídio para justificar ou auxiliar na decisão do pesquisador sobre qual das duas técnicas gráficas utilizar e o porquê deste uso. Os resultados das simulações mostram que o gráfico semi-normal é melhor para auxiliar no julgamento dos efeitos, ao passo que o gráfico normal é recomendado para detectar a presença de valores discrepantes nos dados.

Palavras chave: fatoriais com dois níveis; gráfico normal; gráfico semi-normal; efeitos significativos; parcelas subdivididas; valores discrepantes.

Abstract

Two-level factorial designs are widely used in industrial experimentation. However, many factors in such a design require a large number of runs to perform the experiment, and too many replications of the treatments may not be feasible, considering limitations of resources and of time, making it expensive. In these cases, unreplicated designs are used. But, with only one replicate, there is no internal estimate of experimental error to make judgments about the significance of the observed effects. One of the possible solutions for this problem is to use normal plots or half-normal plots of the effects. Many experimenters use the normal plot, while others prefer the half-normal plot and, often, for both cases, without justification. The controversy about the use of these two graphical techniques motivates this work, once there is no register of formal procedure or statistical test that indicates “which one is best”. The choice between the two plots seems to be a subjective issue. The central objective of this master’s thesis is, then, to perform an experimental comparative study of the normal plot and half-normal plot in the context of the analysis of the 2^k unreplicated factorial experiments. This study involves the construction of simulated scenarios, in which the graphics performance to detect significant effects and to identify outliers is evaluated in order to verify the following questions: Can be a plot better than other? In which situations? What kind of information does a plot increase to the analysis of the experiment that might complement those provided by the other plot? What are the restrictions on the use of graphics? Herewith, this work intends to confront these two techniques; to examine them simultaneously in order to identify similarities, differences or relationships that contribute to the construction of a theoretical reference to justify or to aid in the experimenter’s decision about which of the two graphical techniques to use and the reason for this use. The simulation results show that the half-normal plot is better to assist in the judgement of the effects, while the normal plot is recommended to detect outliers in the data.

Keywords: two-level factorials; normal plot; half-normal plot; significant effects; split-plot; outliers.

Sumário

1	Introdução	1
1.1	Motivação	3
1.2	Objetivos	5
1.3	Estrutura da Dissertação	7
2	Conceitos Fundamentais em Experimentação	8
2.1	Princípios Básicos	9
2.1.1	Replicação	11
2.1.2	Aleatorização	11
2.1.3	Blocagem	12
2.2	Experimentos Fatoriais	13
2.2.1	Experimentos Fatoriais 2^k	14
2.2.2	Experimentos Fatoriais 2^{k-p}	23
2.3	Experimentos Fatoriais 2^k e 2^{k-p} Não Replicados	25
2.4	Métodos Propostos para Análise de Experimentos Fatoriais Não Replicados	26
2.5	Sugestões de Nomenclatura	37
3	Gráficos de Probabilidade Normal	40
3.1	Os Gráficos e a Estatística	40
3.2	Gráficos Probabilísticos em uma Família de Localização-Escala	42
3.3	Gráfico de Probabilidade Normal	48
3.4	Gráfico de Probabilidade Semi-Normal	54
3.5	A Controvérsia	61

4	Levantamento Bibliográfico	64
4.1	Levantamento Bibliográfico na <i>Technometrics</i>	64
4.1.1	Artigos Seleccionados da <i>Technometrics</i>	70
4.2	Levantamento Bibliográfico em Outros Periódicos	75
4.2.1	Artigos Seleccionados dos Demais Periódicos	80
4.3	Síntese	83
5	Estudo Comparativo dos Gráficos de Probabilidade Normal	87
5.1	Metodologia	87
5.2	Cenários Construídos com os Dados dos Artigos	89
5.2.1	Julgamento da Significância dos Efeitos	89
5.2.2	Detecção de Valores Discrepantes nos Dados	107
5.2.3	Identificação de Restrições na Aleatorização (<i>Inadvertent Plot-Splitting</i>)	117
5.3	Cenários para as Simulações	126
5.3.1	Julgamento da Significância dos Efeitos	138
5.3.2	Detecção de Valores Discrepantes nos Dados	145
6	Considerações Finais	163
6.1	Conclusão	163
6.2	Pesquisas Futuras	165
	Referências Bibliográficas	166
A	Programação	176

Capítulo 1

Introdução

Experimentos fatoriais são muito utilizados, especialmente na indústria (OLGUIN; FEARN, 1997). Estes tipos de planos experimentais são mais vantajosos que experimentos com um fator de cada vez, visto que podem detectar e estimar interações entre os fatores. Por se tratarem de uma classe extremamente importante, os experimentos fatoriais 2^k , caso em que todos os fatores possuem dois níveis, destacam-se dentre os planos fatoriais (MONTGOMERY, 2001). Os planos fatoriais 2^k permitem uma análise estatística bastante simplificada e formam a base para outros planos úteis, tais como os experimentos de triagem fatorial ou experimentos exploratórios (*factor screening experiments*), que são experimentos em que muitos fatores estão sendo analisados e deseja-se saber quais deles podem ser desconsiderados por não terem um efeito estatisticamente significativo sobre a variável resposta em estudo. Os planos fatoriais 2^k são usados nos estágios iniciais destes tipos de estudos para que, com um número menor de provas, possam indicar os fatores que afetam significativamente a resposta. Esta indicação serve de direcionamento para outras etapas do experimento, tais como um estudo mais profundo dos fatores julgados ativos ou importantes.

Como se pode perceber pela notação 2^k , à medida que o valor de k aumenta, o número de provas requeridas para a realização do experimento também aumenta, fazendo com que a replicação dos tratamentos se torne inviável, levando em conta as limitações de recursos e de tempo. Os experimentos fatoriais 2^k não replicados são os planos experimentais comumente indicados para esta situação. A expressão “não replicados” indica que, em tais experimentos, há uma única unidade experimental para cada tratamento. A ausência de replicação torna impossível obtermos uma estimativa direta da varia-

bilidade do erro experimental para julgar a significância dos efeitos. Neste caso, dois procedimentos que podem ajudar na avaliação da magnitude dos efeitos, em experimentos não replicados, são as análises feitas a partir do gráfico de probabilidade normal (*full-normal plot*) ou do gráfico de probabilidade semi-normal (*half-normal plot*)¹.

A idéia de um gráfico de probabilidade é comparar duas distribuições de probabilidade: a função de distribuição da amostra (também chamada de distribuição empírica) com uma função de distribuição teórica (também chamada de distribuição hipotetizada). Quando a função teórica escolhida é a distribuição normal, o gráfico é chamado de gráfico de probabilidade normal. Neste gráfico, os pares ordenados são da forma $(\Phi^{-1}(\hat{p}_j), y_{(j)})$, com $j = 1, 2, \dots, n$; em que $y_{(j)}$ são os valores observados na amostra e $\Phi^{-1}(\hat{p}_j)$ seus respectivos quantis na distribuição normal². Se o conjunto de pontos marcados segue um padrão linear, então, sob H_0 : Os dados são normalmente distribuídos, há fortes evidências para se concluir que a distribuição normal serve como um modelo plausível para explicar a população que produziu a amostra. Nesse caso, o intercepto da reta e sua inclinação indicam, respectivamente, a média e o desvio padrão da população da qual a amostra foi obtida (JOHNSON; BHATTACHARYYA, 1992; UPTON; COOK, 2006). Embora sejam procedimentos estatísticos informais, uma vez que sua interpretação é visual e subjetiva, os gráficos de probabilidade normal têm o seu valor (WILK; GNANADESIKAN, 1964, 1968), pois são úteis para verificar a normalidade das amostras, detectar a existência de valores discrepantes nos dados, julgar a significância dos efeitos em planos fatoriais, como também identificar restrições na aleatorização.

Em 1959, Cuthbert Daniel desenvolveu o que ele chamou de *half-normal plot* ou gráfico semi-normal, como uma nova ferramenta estatística para analisar os experimentos fatoriais 2^k (completos e fracionados) não replicados. A idéia central deste gráfico é semelhante àquela apresentada para o gráfico normal: comparar a distribuição da amostra com uma distribuição teórica que, neste caso, é a distribuição semi-normal. Neste gráfico, os pares ordenados são representados pelos módulos dos valores observados na amostra e seus respectivos quantis na distribuição semi-normal. Sua construção,

¹De acordo com o glossário do Instituto Internacional de Estatística (ISI), a tradução para o português de *half-normal plot* é gráfico semi-normal. Ver <http://isi.cbs.nl/glossary>.

²Em inglês, um termo utilizado para se referir a este procedimento gráfico é *full-normal plot*. (ANSCOMBE; TUKEY, 1963; PASTERNAK; LIUZZI, 1965; BOX; MEYER, 1986)

definição e propriedades são apresentadas em seu artigo de origem, publicado em novembro de 1959 na revista *Technometrics* (DANIEL, 1959). Segundo Daniel (1959), o gráfico semi-normal pode ser utilizado para identificar efeitos ativos, estimar a medida da variância do erro experimental, detectar a presença de valores discrepantes (*outliers*), e ainda identificar restrições na aleatorização, como nos experimentos em parcelas subdivididas (*split-plot*). Maiores detalhes sobre os gráficos de probabilidade normal e semi-normal serão destacados no Capítulo 3 desta dissertação.

1.1 Motivação

Embora Daniel (1959) tenha desenvolvido e usado o gráfico semi-normal em suas análises, afirmando que ele é uma ferramenta útil para a estimação do erro experimental bem como para fazer julgamento sobre a significância dos efeitos, sua posição estava para mudar. Em 1976, ele publica seu livro *Applications of Statistics to Industrial Experimentation*, acerca do qual Box e Meyer (1986) comentam que Daniel, neste seu livro, muda de opinião, e prefere usar o gráfico normal, dando como justificativa, a habilidade de tal gráfico em revelar discrepâncias nas suposições acerca dos dados. Neste mesmo ano, Barry H. Margolin (1976) publica a primeira referência na *Technometrics* a usar o gráfico normal com o mesmo objetivo principal que Daniel propôs para o gráfico semi-normal: verificar a significância dos efeitos em experimentos fatoriais não replicados e estimar a variância do erro. Contudo, nenhuma citação é feita sobre o artigo de Daniel (1959), apesar de, anos antes, Margolin (1967) ter indicado o gráfico semi-normal como uma técnica para julgar a significância dos efeitos dos fatores. Schoen e Kaul (2000, p. 276) também comentam que Daniel (1976) recomenda o gráfico normal por este ser mais capaz, em relação ao gráfico semi-normal, de detectar “*anomalies in the data*”. Como se já não bastasse essa controvérsia, contribuindo ainda mais para essa questão, há uma afirmação (não confirmada) de que Daniel tenha mudado de opinião a favor do gráfico semi-normal novamente³.

Esta situação levanta algumas questões: Qual dos dois procedimentos é o “melhor”? Qual deles usar? Em que situações? Com que justificativa? Com isso, a controvérsia havia começado. Muitos pesquisadores, em seus trabalhos publicados, usam

³Ver www.statease.com/news/faqalert6-09.html

o gráfico normal, muitas vezes sem alguma justificativa; e outras vezes argumentando que, pelo fato do gráfico semi-normal utilizar apenas os valores absolutos dos efeitos, ele acaba perdendo a informação que os sinais podem dar na análise (OLGUIN; FEARN, 1997). Por outro lado, os usuários do gráfico semi-normal, quando se justificam, o fazem dizendo que, como em muitas análises os diferentes sinais (“+” e “-”) para os níveis dos fatores são arbitrários, estes podem produzir diferentes gráficos normais para um mesmo experimento, bastando para isso apenas mudar estas codificações, uma vez que estes gráficos usam os dados com seus sinais originais; ao passo que um único gráfico semi-normal é produzido, e será o mesmo, independente das configurações dos sinais escolhidas para os níveis dos fatores, haja visto que este gráfico usa os dados em seus valores absolutos (OLGUIN; FEARN, 1997). Outros autores afirmam que o gráfico semi-normal é mais sensível para identificar efeitos significativos (BRADU; HAWKINS, 1982).

Desta forma, a controvérsia sobre as duas técnicas gráficas é que motiva a realização da presente proposta de pesquisa, uma vez que não há registro de procedimento formal ou teste estatístico que indique “qual delas é melhor”. A escolha entre o gráfico normal e o semi-normal para a análise de experimentos fatoriais não replicados parece ser uma questão subjetiva.

Enfim, aos cinquenta anos de publicação do artigo de Daniel sobre o gráfico semi-normal, esta dissertação busca verificar as seguintes questões: Um gráfico pode ser melhor que o outro? Em que situações? Que informações um gráfico acrescenta à análise do experimento que possam complementar aquelas fornecidas pelo outro gráfico? Quais as restrições no uso de cada gráfico? O conhecimento construído pelas respostas dadas a estas e a outras questões com certeza contribuirá para o avanço da Estatística como um todo, e em particular, para a área de Planejamento de Experimentos, que é a base desta pesquisa.

1.2 Objetivos

Os objetivos desta dissertação são:

Objetivo central: realizar um estudo comparativo experimental dos gráficos normal e semi-normal no contexto da análise de experimentos fatoriais 2^k não replicados, no intuito de avaliar o desempenho destes gráficos sob dois importantes aspectos: sua utilidade em auxiliar no julgamento da significância dos efeitos, bem como sua capacidade de identificar valores discrepantes nos dados.

Objetivos secundários:

- apresentar um levantamento bibliográfico sobre o uso dos gráficos de probabilidade normal em artigos científicos;
- apresentar uma síntese dos métodos propostos para análise de experimentos fatoriais não replicados;
- apresentar sugestões de nomenclatura em português para termos utilizados em experimentação.

Inicialmente é realizada uma busca de artigos publicados na revista *Technometrics*, na qual Daniel lançou o gráfico semi-normal. Esta busca compreende o período de fevereiro de 1959 até agosto de 2008. O objetivo é encontrar artigos nos quais o gráfico normal ou o semi-normal tenham sido utilizados para a análise de experimentos, a fim de delinear ou quantificar o uso desses dois gráficos nos trabalhos científicos, como também verificar quais as justificativas apresentadas para o uso destes. Pesquisa semelhante também é feita em outros periódicos. Três motivos podem ser citados para a escolha da *Technometrics*: (1) trata-se de um importante veículo de divulgação de métodos e estudos de caso pertinentes ao uso da estatística na indústria e tecnologia (STEINBERG; BISGAARD, 2008); (2) Cuthbert Daniel publicou seu artigo sobre o gráfico semi-normal neste periódico; (3) o fato de tanto este periódico quanto a técnica desenvolvida por Daniel terem completado cinquenta anos em 2009.

Uma vez terminado esse trabalho de pesquisa bibliográfica, situações experimentais (reais e simuladas) de experimentos fatoriais não replicados são construídas, e

nestas, os dois métodos gráficos são utilizados e comparados, com o objetivo de analisar algumas características de seu desempenho e comportamento. Inicialmente, os cenários realizados são escolhidos dentre os artigos pesquisados e divididos em três grupos. Cada grupo representa o aspecto sob o qual o desempenho dos gráficos é avaliado: (1) o julgamento da significância dos efeitos dos fatores; (2) a capacidade de detectar a presença de valores discrepantes nos dados (*outliers*); (3) a identificação de restrições na aleatorização (*inadvertent plot-splitting*). Estes aspectos são escolhidos, pois, segundo a proposta de Daniel (1959), o gráfico semi-normal é uma ferramenta útil para estas três situações. Nesta fase inicial, os critérios para a comparação dos gráficos são as percepções do analista, sua sensibilidade de captar características distintas nos gráficos por meio da análise visual destes, tendo por base os dados e os resultados descritos nos artigos de origem destes experimentos. Os resultados obtidos e as recomendações feitas nessa fase inicial fornecem as indicações para a metodologia a ser adotada no próximo passo do trabalho: a construção de simulações de cenários experimentais de planos fatoriais não replicados para realizar um estudo comparativo dos gráficos normal e semi-normal, para cada um dos três aspectos mencionados anteriormente. Nesta segunda fase, os critérios usados para a comparação dos gráficos são métodos objetivos, isto é, testes e procedimentos formais, com resultados analíticos, a fim de que a confrontação gráfica não se dê de forma apenas subjetiva, como a que ocorre na fase inicial do estudo.

Trata-se, portanto, de um experimento comparativo em que se tenta controlar a situação através da qual as observações são feitas, ou seja, é um estudo comparativo experimental de simulação. Com isso, propõe-se confrontar estas duas técnicas; examiná-las simultaneamente a fim de conhecer semelhanças, diferenças ou relações que possam contribuir para a construção de um referencial teórico que sirva como um subsídio para justificar ou auxiliar na decisão do pesquisador sobre qual das duas técnicas gráficas utilizar e o porquê deste uso. As informações assim produzidas por esta dissertação evidenciam a importância do tema em questão, como também justificam a realização deste trabalho.

1.3 Estrutura da Dissertação

A presente dissertação encontra-se dividida em seis capítulos e um apêndice. O Capítulo 2 expõe uma revisão de conceitos importantes no planejamento de experimentos, dando ênfase às características e a uma teoria geral para os planos fatoriais 2^k . A seguir, é feita uma síntese dos principais métodos propostos para análise de experimentos fatoriais não replicados. Este capítulo é finalizado com sugestões de nomenclatura em português para termos utilizados em experimentação. O Capítulo 3 aborda os gráficos de probabilidade normal e semi-normal. O Capítulo 4 trata da revisão de literatura na análise dos artigos selecionados no levantamento bibliográfico, no intuito de esboçar o uso dos gráficos de probabilidade normal em artigos científicos. O Capítulo 5 discute o experimento aqui proposto e seus resultados. O Capítulo 6 apresenta as conclusões e algumas sugestões para trabalhos futuros. Finalmente, no Apêndice, constam os algoritmos dos programas e comandos utilizados na aplicação dos dados, como também nas simulações.

Capítulo 2

Conceitos Fundamentais em Experimentação

Um **experimento** é uma ação investigativa que visa gerar conhecimento sobre um determinado produto ou processo. É uma intervenção num sistema (DANIEL, 1976); um teste ou uma série de testes no funcionamento de um sistema operante com o intuito de aprender sobre seu desempenho (MONTGOMERY, 2001), como também analisar e avaliar as implicações das observações resultantes (KUEHL, 2000). O **planejamento** é a elaboração dessa intervenção. Trata-se de uma mediação para que uma proposta de ação consciente e intencional se torne o mais clara, precisa, eficiente e direcionada possível. Desta forma, pode-se entender o **Planejamento de Experimentos** “como uma parte da Estatística que busca permitir ao experimentador a obtenção de dados que lhe sejam úteis, no sentido de fornecer informações de acordo com o objetivo, de uma forma tão econômica quanto possível.” (VIVACQUA; PINHO, 2008); um conjunto de conhecimentos e técnicas que habilitam um investigador a conduzir melhor seus experimentos, analisar dados eficientemente e fazer conexões entre os resultados dessas análises e os objetivos originais da investigação (WU; HAMADA, 2000); um método estatístico, aplicado há muitos anos na indústria visando a melhoria da qualidade, que busca estabelecer que variáveis são importantes num processo e as condições sob as quais estas variáveis otimizam tal processo (ILZARBE et al, 2008). Trata-se, de fato, das estratégias de experimentação: elaboração (planejamento), execução e avaliação (análise e interpretação). Um experimento bem planejado permite a obtenção de dados úteis, a um mínimo de custos, tempo e riscos; como também garante a validade das conclusões (VIVACQUA; PINHO, 2008). O planejamento de experimentos, conforme

Box, Hunter e Hunter (2005) afirmam, dinamiza o método científico, aumentando sua eficiência e, com isso, proporcionando a geração de conhecimento.

Este capítulo trata de algumas considerações teóricas sobre Planejamento de Experimentos: seus princípios básicos, sua terminologia e, especialmente, sobre o tipo mais conhecido e utilizado de plano experimental que são os experimentos fatoriais; visto ser o cenário principal desta pesquisa. Não se pretende exaurir nem detalhar todos os conceitos e teorias que abrangem esse assunto, mas apenas dar um enfoque conciso e objetivo sobre os aspectos principais; essenciais para o entendimento do trabalho. Para maiores detalhes, consultar bibliografia especializada, tais como Kempthorne e Hinkelmann (2008); Box, Hunter e Hunter (2005); Montgomery (2001) e Daniel (1976). Para uma análise do desenvolvimento histórico da técnica de planejamento de experimentos e análise de dados, consultar Box (1989).

2.1 Princípios Básicos

Assim como toda ciência, a Estatística possui linguagem própria, específica para seu campo de atuação, que é bem vasto. E com Planejamento de Experimentos não é diferente. É preciso ter conhecimento da terminologia básica dessa área de estudo a fim de que o que será exposto no decorrer desta dissertação possa fazer sentido e ser claramente entendido. Por exemplo, chama-se de **resposta** a característica de interesse que será medida no estudo (VIVACQUA; PINHO, 2008). **Fator** é a característica que será controlada no experimento (VIVACQUA; PINHO, 2008); é a variável cuja influência sobre a resposta está sendo estudada (WU; HAMADA, 2000). Os **níveis** de um fator são os valores que um determinado fator pode assumir (WU; HAMADA, 2000); são os diferentes modos de presença de um fator no estudo. **Efeito** é a mudança na resposta produzida por uma mudança no nível do fator. Um **tratamento** é uma combinação específica dos níveis dos fatores (WU; HAMADA, 2000); é o elemento cujo efeito se deseja medir ou comparar num experimento; é o conjunto de circunstâncias criadas para o experimento em resposta à hipótese de investigação, sendo o foco da pesquisa (KUEHL, 2000). **Unidade observacional** é a unidade em que as observações ou medições são feitas (KEMPTHORNE; HINKELMANN, 2008); é ela que fornece a resposta. **Unidade experimental** é a unidade a qual um tratamento é atribuído e

aplicado (KEMPTHORNE; HINKELMANN, 2008). É importante distinguir estes dois tipos de unidades, pois em muitas situações elas podem ser idênticas, ao passo que em outras não o serão.

Associado à unidade observacional está o **erro observacional**, que reflete erros de medida ou de amostragem (KEMPTHORNE; HINKELMANN, 2008). Igualmente, associado às unidades experimentais está o **erro experimental**, que descreve a variação entre as unidades experimentais tratadas de maneira idêntica e independente (KUEHL, 2000); são as fontes de variação não conhecidas presentes em todo experimento. Estas fontes de variação contribuem para um comportamento aleatório, uma vez que não foram levadas em consideração por ocasião da realização do experimento. São também inevitáveis, e podem surgir em decorrência de vários motivos: a variação natural entre as unidades experimentais; a variabilidade na medição da resposta; a incapacidade de reproduzir de modo exato as condições do tratamento de uma unidade para outra; a interação dos tratamentos com as unidades experimentais; bem como qualquer outro fator externo que influencie nas características medidas (KUEHL, 2000). A falta de entendimento sobre o erro experimental pode levar o investigador a tirar conclusões equivocadas de seu experimento.

Bloco é um grupo de unidades experimentais relativamente homogêneas (MONTGOMERY, 2001), uma porção de material experimental mais homogênea que o restante; é a variável que causa uma maior heterogeneidade nas unidades. A **interação** mede a influência conjunta de dois ou mais fatores sobre a resposta; ela existe quando o efeito de um fator sobre a variável resposta depende do nível de outro fator (ANTONY, 1999). Não haverá interação se o efeito de um fator sobre a resposta é o mesmo a diferentes níveis de outro fator. Temos ainda a **prova** (ou ensaio, ou corrida), que é cada realização de uma determinada condição de interesse, isto é, cada prova corresponde à aplicação de um tratamento a uma unidade experimental (VIVACQUA; PINHO, 2008). O propósito básico da experimentação é chegar a uma combinação de níveis de fatores que otimizem a resposta; identificar fatores importantes que controlem a característica de interesse; como também encontrar níveis dos fatores que propiciem uma resposta robusta, isto é, que seja afetada o mínimo possível por fontes externas de variabilidade.

Agora, tendo em mente estas definições principais, serão considerados os três prin-

cípios básicos do planejamento de experimentos: replicação, aleatorização e blocagem.

2.1.1 Replicação

A replicação consiste em aplicar, independentemente, um tratamento sob estudo em mais de uma unidade experimental. Pode-se destacar dois objetivos práticos da replicação (MONTGOMERY, 2001). Primeiro, ela permite ao investigador obter uma estimativa do erro experimental. Esta estimativa do erro é importante pois, no processo de análise dos dados, ela se torna uma medida essencial para determinar se as diferenças observadas nos dados são estatisticamente significativas; bem como pode explicar a fonte principal dessa variação. Segundo, a replicação permite ao experimentador obter uma estimativa mais precisa dos efeitos dos fatores. Quanto maior o número de réplicas, menor será a variância da estimativa dos efeitos dos fatores e mais fácil será detectar diferenças entre eles (WU; HAMADA, 2000). De acordo com Kuehl (2000), há muitas razões para se replicar um experimento: a replicação oferece um certo grau de segurança contra resultados aberrantes; provê o meio para estimar a variância do erro experimental; aumenta a precisão na estimação dos efeitos.

Uma distinção precisa ser feita entre replicação e repetição. Na replicação, um mesmo tratamento é aplicado a diferentes unidades experimentais, enquanto na repetição, muitas observações são feitas de uma mesma unidade experimental (KUEHL, 2000). Por exemplo, três medidas de uma mesma unidade experimental são repetições, enquanto as medidas de três unidades experimentais diferentes, que receberam o mesmo tratamento, são replicações. A variância entre as medidas repetidas apenas medem a variação devida aos erros do processo de medida; ao passo que a variância entre as medidas das unidades experimentais é a variância do erro experimental, com o qual se compara os efeitos dos fatores, verificando qual deles é significativo.

2.1.2 Aleatorização

Aleatorização, no contexto de planejamento de experimentos, significa que a alocação dos tratamentos às unidades experimentais, bem como a ordem de execução dos ensaios (ou provas) é determinada de forma aleatória (isto é, cada unidade experimental tem probabilidade conhecida de receber qualquer tratamento); pelo uso

de algum procedimento físico (sorteio, uso de uma tabela de números aleatórios etc.). A aleatorização valida a aplicação dos métodos estatísticos para a análise dos dados pois gera uma distribuição de referência (KUEHL, 2000; BOX; HUNTER; HUNTER, 2005; KEMPTHORNE; HINKELMANN, 2008). Desta forma, quando o experimento é aleatorizado é possível fazer testes estatísticos sem a necessidade das suposições de normalidade, homocedasticidade e independência acerca do erro experimental, e por conseguinte das observações. É importante atentar para as suposições, pois, cada análise estatística que se pretende realizar é construída sob certas condições que precisam ser satisfeitas pelo experimento, a fim de que as inferências realizadas sejam válidas. Estas suposições basicamente referem-se ao modelo que será construído, bem como à distribuição das flutuações aleatórias.

A aleatorização é uma forma de tornar a designação tratamento-unidade experimental imparcial, como também de tentar balancear a ação do erro aleatório entre as unidades experimentais (VIVACQUA; PINHO, 2008), como que atenuando ou equilibrando o efeito das fontes de variação que possam afetar a resposta. Um experimento não aleatorizado pode fornecer estimativas tendenciosas dos efeitos dos fatores bem como da variância do erro experimental, o que conduziria a conclusões erradas ou equivocadas sobre o experimento. Em algumas situações, a aleatorização pode ser impossível de ser feita, e há casos em que seu mal uso pode ser também prejudicial. Para mais detalhes sobre estas situações e suas implicações práticas ver Daniel (1976, capítulo 3).

2.1.3 Blocagem

Um aspecto importante na experimentação diz respeito à variabilidade do material experimental. É essencial, para a validade das conclusões, que as unidades experimentais sejam similares. Quanto mais homogêneas forem as unidades experimentais, mais claro será avaliar o efeito dos tratamentos sobre elas. Mas em muitas situações, as unidades experimentais possuem características diferentes. Isto se deve a fatores característicos próprios destas unidades, que influenciam nas observações e medições feitas. Tais fatores, conhecidos pelo investigador, embora não sejam o objeto principal no estudo, acrescentam uma variabilidade adicional aos dados que pode mascarar os efeitos devidos aos fatores de interesse; pois o erro experimental, tendo sido “inflacio-

nado”, refletirá tanto o erro aleatório inerente ao experimento, quanto a variabilidade existente no material experimental. Uma vez identificada essa situação, estes fatores devem ser levados em conta. Na prática, isto é feito considerando-os no experimento na forma de blocos. A este procedimento dá-se o nome de blocagem.

A blocagem permite fazer comparações entre as condições de interesse no experimento dentro de cada bloco, uma vez que os tratamentos são sorteados dentro do bloco. Tem por objetivo, então, amenizar o efeito das fontes de variação (VIVACQUA; PINHO, 2008). As unidades experimentais são agrupadas de tal modo que a variabilidade dentro dos grupos seja a menor possível. As respostas dentro dos blocos podem ser comparadas com maior precisão, porque as diferenças dentro dos blocos, e conseqüentemente entre as unidades experimentais, tendem a ser menores. Com isso, as diferenças entre os tratamentos que estão sendo comparados não são confundidas com as diferenças entre as unidades experimentais. Desta forma, a blocagem serve para reduzir a variação do erro experimental (KUEHL, 2000; BOX; HUNTER; HUNTER, 2005; KEMPTHORNE; HINKELMANN, 2008).

Uma vez considerada esta parte introdutória dos conceitos e princípios básicos do planejamento de experimentos, a próxima seção discute uma forma bastante usual de plano experimental: os experimentos fatoriais 2^k . Além de sua importância e grande utilização na experimentação industrial, estes experimentos compõem os cenários que servem de base para esta pesquisa.

2.2 Experimentos Fatoriais

Em muitas situações de investigação científica, o pesquisador deseja estudar o efeito de determinados fatores sobre a variável resposta de interesse. Isso pode ser feito de vários modos. Um deles poderia ser o uso do método conhecido como Um Fator de Cada Vez. Neste método, os níveis dos fatores são modificados um de cada vez, enquanto os demais fatores são mantidos constantes. Depois desse processo, verifica-se como a variável resposta foi afetada. Este método possui a desvantagem de ter como suposição a idéia de que os fatores agem na resposta aditivamente, isto é, não considera alguma possível interação entre os fatores. Isto é um problema, pois, como a interação entre fatores é algo muito comum em experimentação, a sua não detecção (quando esta

realmente está presente) pode levar a sérios erros nas conclusões do experimento.

Visando superar este problema, uma abordagem correta para lidar com mais de um fator seria conduzir um experimento fatorial. Nessa estratégia experimental, os fatores são variados no mesmo experimento, ao invés de um de cada vez. Os experimentos fatoriais, portanto, são mais vantajosos que os planos Um Fator de Cada Vez, pois, como os fatores são estudados simultaneamente, os planos fatoriais podem detectar e estimar possíveis interações. Além disso, mesmo que os fatores atuem aditivamente, o plano fatorial requereria menos provas. E mesmo que os dois planos usassem o mesmo número de provas, o plano fatorial teria uma maior precisão (VIVACQUA; PINHO, 2008), devido a *hidden replications* (réplicas escondidas). Portanto, os experimentos do tipo Um Fator de Cada Vez são sempre menos eficientes que os experimentos fatoriais (BOX; HUNTER; HUNTER, 2005).

Um planejamento fatorial, então, é aquele no qual dois ou mais fatores são estudados simultaneamente e, para cada réplica completa do experimento, todas as possíveis combinações dos níveis dos fatores são pesquisadas, isto é, cada combinação dos níveis dos fatores aparecerá o mesmo número de vezes. E é exatamente aí, segundo Webster (2006), que se encontra a maior desvantagem dos experimentos fatoriais: o número de unidades experimentais requeridas para a execução do experimento aumenta rapidamente com o aumento do número de fatores. Mas, segundo Montgomery (2001), esta situação também apresenta uma vantagem: geralmente, a eficiência relativa dos experimentos fatoriais aos experimentos Um Fator de Cada Vez aumentará com o aumento do número de fatores.

Percebe-se, portanto, que os experimentos fatoriais são extremamente úteis. São também bastante utilizados na experimentação industrial (ANTONY, 1999; VOSS; WANG, 1999; LOEPPKY; SITTER, 2002; SCHOEN, 2004; VARIYATH; ABRAHAN; CHEN, 2005; LAWSON, 2008). Em vista disso, são destacados agora alguns importantes aspectos desse planejamento experimental.

2.2.1 Experimentos Fatoriais 2^k

Dentre os casos especiais dos planejamentos fatoriais, os experimentos 2^k acham-se entre os mais importantes, por serem altamente utilizados em trabalhos de pesquisa; em especial na área industrial (DANIEL, 1976), como também pelo seu considerável

valor prático. Box, Hunter e Hunter (2005) também citam alguns motivos para a importância especial desse plano experimental: requerem relativamente poucas provas por fator estudado; a interpretação dos resultados obtidos pode ser feita pelo uso de aritmética elementar e recursos gráficos; são particularmente úteis nos estágios iniciais de um trabalho experimental, pois frequentemente apontam a direção para uma investigação mais avançada; têm uma análise estatística bastante simplificada; e formam a base para muitos outros planejamentos úteis, tais como os *factor screening experiments* (MONTGOMERY, 2001), discutidos no Capítulo 1.

A notação 2^k indica que nestes experimentos existem k fatores, todos a dois níveis. Estes fatores podem ser quantitativos ou qualitativos. Os dois níveis dos fatores são geralmente denominados baixo e alto, e representados pelos sinais “-” e “+”, respectivamente. O número total de tratamentos utilizados será o resultado da potência 2^k ; daí porque são chamados de **fatoriais completos**. Nos experimentos **balanceados**¹ o número de provas requeridas será a quantidade de tratamentos utilizados, incluindo possíveis replicações, ou seja, se um experimento tem n réplicas, o número de provas neste experimento será $n2^k$.

Nos experimentos fatoriais 2^k existem k efeitos principais, $\binom{k}{2}$ interações de dois fatores ou de segunda ordem, $\binom{k}{3}$ interações de terceira ordem, \dots , e uma única interação de k fatores. Com isso, todas estas quantidades podem ser expressas como $\binom{k}{i}$, com $1 \leq i \leq k$. Com o uso desses planejamentos deseja-se estimar os **efeitos principais**² dos fatores e as interações, bem como avaliar suas significâncias estatísticas.

Três princípios fundamentais são considerados no estudo dos efeitos nos planos fatoriais (WU; HAMADA, 2000):

- **Princípio da Ordenação Hierárquica:** Efeitos principais e interações de baixa ordem parecem ser mais importantes que interações de ordem mais alta. Este princípio significa que, na prática, em experimentos com muitos efeitos a serem estimados, a evidência empírica tem mostrado que as interações de alta ordem (geralmente as interações a partir de terceira ordem) não serão estatisticamente

¹Um experimento é dito balanceado quando há o mesmo número de observações para cada tratamento (MONTGOMERY, 2001).

²Estes efeitos referem-se às comparações de interesse primário no estudo. No caso dos fatoriais 2^k o efeito principal de um fator é a mudança na média da resposta quando mudamos do nível “-” para o nível “+”.

significativas. Outra razão para a grande aceitação deste princípio é que interações de alta ordem são mais difíceis de interpretar ou justificar fisicamente.

- **Princípio da Esparsidade dos Efeitos (BOX; MEYER, 1986):** O número de efeitos significativos num experimento fatorial é relativamente pequeno.
- **Princípio da Hereditariedade do Efeito:** Para que uma interação seja significativa, pelo menos um dos fatores que a compõe deveria ser significativo.

Exemplo 2.1. A estrutura básica dos fatoriais 2^k completos será dada por meio de um exemplo prático e bastante simples. Este experimento foi realizado por um grupo de alunos da disciplina Planejamento de Experimentos do Curso de Estatística da UFRN, no decorrer do semestre 2008.2, sob orientação da Professora Carla Almeida Vivacqua. O experimento tem por objetivo verificar que fatores são relevantes para se conseguir um menor tempo de dissolução de comprimidos efervescentes de vitamina C. Os fatores considerados são os seguintes:

Tabela 2.1: Fatores considerados no experimento com os comprimidos de vitamina C.

Fator	Nível (-)	Nível (+)
A: Quantidade de água	50 ml	100 ml
B: Temperatura da água	23,8 °C(ambiente/natural)	14,5 °C(gelada)
C: Marca do comprimido	Bio - C	Cenevit

Trata-se de um fatorial 2^3 completamente aleatorizado, realizado com duas réplicas. A resposta de interesse é o tempo de dissolução do comprimido, medido em segundos. Os resultados encontram-se na Tabela 2.2, que geralmente é o modo como os arranjos fatoriais 2^k são apresentados³. Os efeitos principais A , B e C são escritos em letras maiúsculas, assim como suas interações AB , AC , BC e ABC . Como já dito anteriormente, os sinais “-” e “+” referem-se aos níveis baixo e alto de cada fator, respectivamente. A disposição dos sinais para os efeitos principais A , B e C é padrão, ao passo que os sinais das colunas das interações é o resultado do produto dos sinais dos efeitos envolvidos nesta interação. Já os $2^3 = 8$ tratamentos são representados por letras minúsculas, com o seguinte significado: quando o fator estiver em seu nível alto, este será representado pela presença da letra minúscula correspondente; a ausência de

³A tabela dos sinais para os efeitos principais e os efeitos das interações usada neste exemplo pode ser obtida similarmente para qualquer experimento fatorial 2^k .

uma letra indica que aquele fator encontra-se em seu nível baixo. Assim, por exemplo, *a* representa o fator **A** em seu nível alto e os fatores **B** e **C** em seu nível baixo; e *bc* representa **A** no nível baixo e **B** e **C** em seus níveis altos. A notação (1) é usada para representar o tratamento em que todos os fatores estão em seus níveis baixos. Convém lembrar ainda que a sequência (1), *a*, *b*, *ab*, *c*, *ac*, *bc*, *abc* é a **ordem padrão** na qual os tratamentos são apresentados; não significando que esta tenha sido a ordem de execução do experimento. Esta última é representada pelos números entre parênteses que acompanham os resultados escritos nas colunas das réplicas. A Figura 2.1 é uma representação geométrica desse experimento.

Tabela 2.2: Resultados do experimento com os comprimidos de vitamina C.

Tratamento	A	B	C	AB	AC	BC	ABC	Réplica I	Réplica II	Média
(1)	-	-	-	+	+	+	-	158,4 ⁽¹²⁾	151,3 ⁽⁶⁾	154,9
<i>a</i>	+	-	-	-	-	+	+	126,9 ⁽⁵⁾	125,4 ⁽¹⁾	126,2
<i>b</i>	-	+	-	-	+	-	+	102,6 ⁽²⁾	137,3 ⁽¹⁰⁾	119,9
<i>ab</i>	+	+	-	+	-	-	-	106,9 ⁽¹³⁾	102,1 ⁽¹⁴⁾	104,5
<i>c</i>	-	-	+	+	-	-	+	232,1 ⁽⁸⁾	231,5 ⁽¹⁶⁾	231,8
<i>ac</i>	+	-	+	-	+	-	-	219,1 ⁽¹⁵⁾	205,7 ⁽¹¹⁾	212,4
<i>bc</i>	-	+	+	-	-	+	-	137,7 ⁽⁹⁾	148,8 ⁽⁷⁾	143,3
<i>abc</i>	+	+	+	+	+	+	+	128,1 ⁽³⁾	127,5 ⁽⁴⁾	127,8

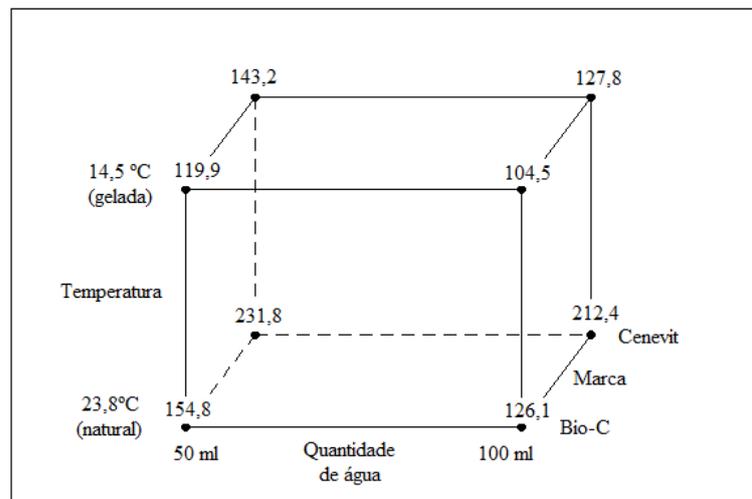


Figura 2.1: **Cubo gráfico** (MINITAB15 - versão DEMO). Representação geométrica do experimento fatorial 2^3 sobre os comprimidos efervescentes.

O cubo gráfico da Figura 2.1 mostra como os níveis dos fatores afetam a resposta. E como o experimento é do tipo “quanto menor melhor”, isto é, quanto menos tempo o comprimido demorar para dissolver melhor será seu desempenho; a combinação dos fatores que produzem essa resposta é $A^+B^+C^-$, ou seja, com 100 ml de água a uma temperatura de 14,5 °C, o comprimido da marca Bio - C terá o menor tempo de dissolução. Os procedimentos que se seguem são uma confirmação analítica desta interpretação gráfica.

Deseja-se agora estimar os efeitos dos fatores. Isto é feito de um modo bastante simples. Lembrando que, no caso dos fatoriais 2^k o efeito é a mudança provocada na resposta quando o fator muda do nível “-” para o nível “+”, a própria Tabela 2.2 fornece os meios de calcular estes efeitos. Considere, por exemplo, o fator A . O efeito principal de A sobre a resposta observada é o resultado da diferença entre as médias das provas em que A está no nível alto e daquelas em que A se encontra em seu nível baixo. Desta forma, temos:

$$\text{Efeito}_A = \bar{y}_{A^+} - \bar{y}_{A^-} = \frac{1}{2^{(k-1)}n} [a + ab + ac + abc - (1) - b - c - bc] = -9,875$$

em que n é o número de réplicas do experimento. O mesmo raciocínio pode ser utilizado para calcular os efeitos principais de B e C , bem como os efeitos das interações AB , AC , BC e ABC . A Tabela 2.3 mostra as estimativas destes efeitos, calculadas no pacote estatístico MINITAB15 - versão DEMO.

Tabela 2.3: Estimativas dos efeitos dos fatores.

Efeito	Estimativa	t	Valor-p
A	-9,87	-3,97	0,004
B	-28,71	-11,55	0,000
C	26,22	10,55	0,000
AB	2,15	0,86	0,412
AC	1,16	0,47	0,653
BC	-14,58	-5,86	0,000
ABC	-1,16	-0,47	0,653

Como se pode ver, há $2^k - 1$ efeitos a estimar. Isto pode ser claramente entendido por meio do Binômio de Newton:

$$\begin{aligned} 2^k &= (1 + 1)^k = \binom{k}{0} + \binom{k}{1} + \binom{k}{2} + \dots + \binom{k}{k} \\ 2^k &= 1 + \binom{k}{1} + \binom{k}{2} + \dots + \binom{k}{k} \\ 2^k - 1 &= \binom{k}{1} + \binom{k}{2} + \dots + \binom{k}{k} \end{aligned}$$

do qual $2^k - 1$ é a quantidade total das estimativas dos efeitos principais e de todas as possíveis interações entre eles. A unidade retirada da quantidade 2^k refere-se à média geral das observações.

Duas propriedades básicas dos fatoriais 2^k são o balanço (equilíbrio) e a ortogonalidade (WU; HAMADA, 2000). O balanço é o fato de que cada coluna dos efeitos está perfeitamente balanceada com respeito a todas as outras colunas. Isso pode ser visto na Tabela 2.2. Escolhendo qualquer uma das colunas dos efeitos e atentando para o grupo de sinais nela presentes, nota-se que para cada sinal “+” existe um sinal “-” correspondente. A ortogonalidade assegura que cada efeito estimado não é afetado pelas magnitudes e sinais dos outros (BOX; HUNTER; HUNTER, 2005).

Uma importante definição no estudo dos experimentos fatoriais é a de **contraste**. Os contrastes são funções lineares das médias dos tratamentos (KUEHL, 2000), cujas estimativas são calculadas a partir dos dados observados, pela combinação linear das médias. Desta forma, um contraste C é dado por:

$$C = \sum_{i=1}^k t_i \bar{y}_i = t_1 \bar{y}_1 + \dots + t_t \bar{y}_k. \quad (2.1)$$

em que $\sum_{i=1}^k t_i = 0$, com t_i conhecido, e \bar{y}_i é a média do i -ésimo grupo de tratamento. Assim, da Equação 2.1, percebe-se que a expressão $[a+ab+ac+abc-(1)-b-c-bc]$ usada no cálculo do efeito do fator A e, conseqüentemente, todas as diferentes combinações desses tratamentos usadas nas estimativas dos demais efeitos, são contrastes. Além disso, de acordo com o parágrafo anterior, estes são também **contrastos ortogonais**, uma vez que suas estimativas são não correlacionadas (DEAN; VOSS, 1999).

O próximo passo agora é saber qual dos efeitos calculados é significativo, ou seja,

quais deles diferem estatisticamente de zero. Como o experimento é replicado, o teste formal para o julgamento dos efeitos, que pode ser visto na Tabela 2.3 e na Tabela 2.4, é a ANOVA, que apontou como significativos os efeitos A , B , C e BC . Os gráficos de probabilidade normal e semi-normal, mostrados na Figura 2.2, também indicam este mesmo resultado. Como a interação BC é significativa, não se deve considerar estes fatores isoladamente, pois há fortes evidências de interação entre eles. Assim, atenção deve ser dada ao efeito principal A e à interação BC . Vale destacar aqui também o fato de que, considerando que o experimento foi aleatorizado, não há necessidade de suposições sobre a distribuição dos dados, pois a distribuição de referência gerada pela aleatorização fornece informação para fazer inferência. A distribuição *t-student* e a distribuição *F* se aproximam suficientemente bem da distribuição de aleatorização (VIVACQUA; PINHO, 2008); por isso, para efeito de cálculo e simplicidade, elas são usadas nos testes descritos na Tabela 2.3 e na Tabela 2.4, respectivamente, ao invés da distribuição de aleatorização. Uma confirmação visual do que foi apontado pelo teste da ANOVA pode ser vista na Figura 2.3 e na Figura 2.4.

Tabela 2.4: Tabela da ANOVA para o experimento com os comprimidos de vitamina C.

FONTES	GL	SQ	QM	F	VALOR-P
Efeitos principais	3	25754,8	8584,93	86,76	0,000
Interações de segunda ordem	3	3494,5	1164,82	11,77	0,003
Interações de terceira ordem	1	21,6	21,62	0,22	0,653
Resíduos	8	791,6	98,96		
Total	15	30062,5			

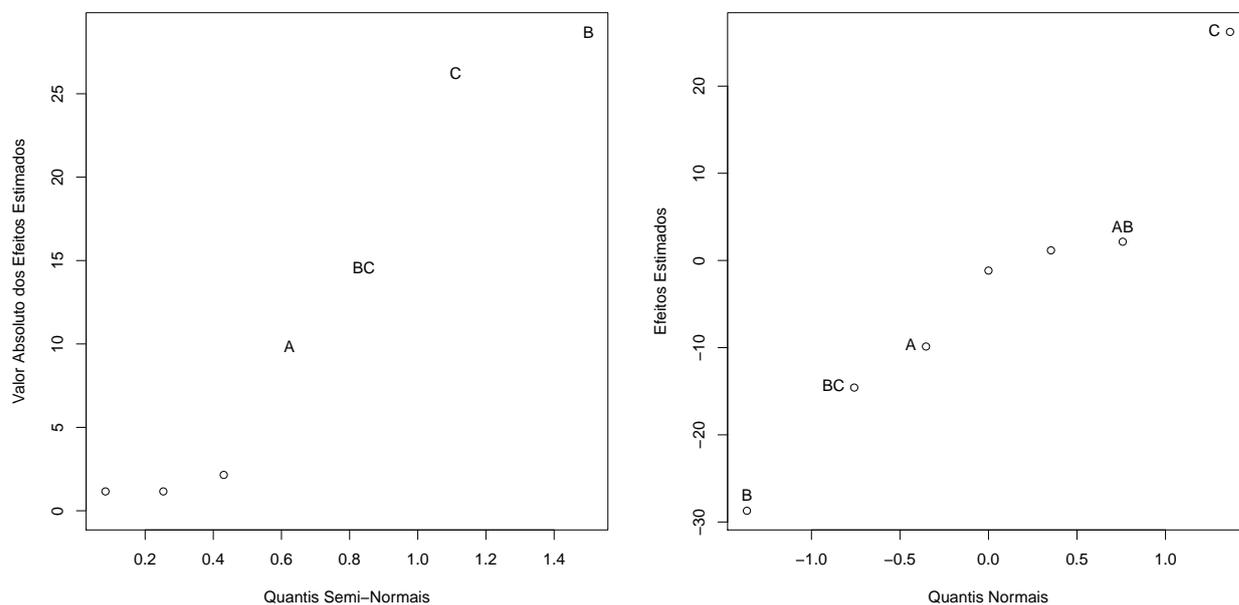


Figura 2.2: Gráficos de Probabilidade para o Exemplo 2.1.

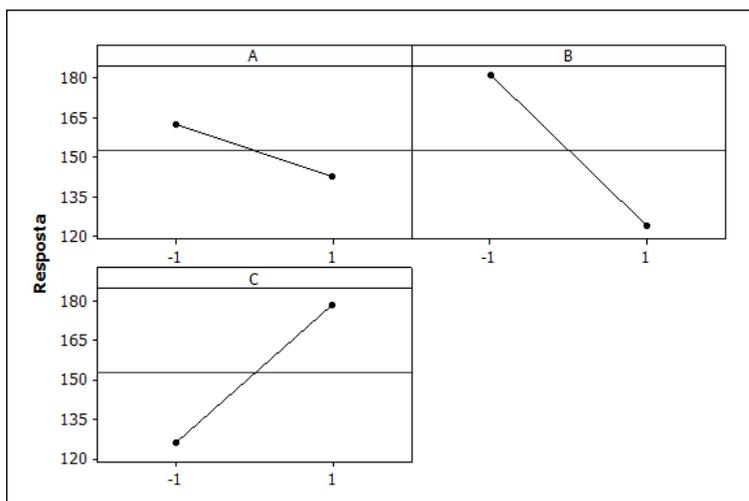


Figura 2.3: Gráfico dos Efeitos Principais. Representação geométrica de como os efeitos principais afetam o tempo de dissolução dos comprimidos efervescentes de vitamina C.

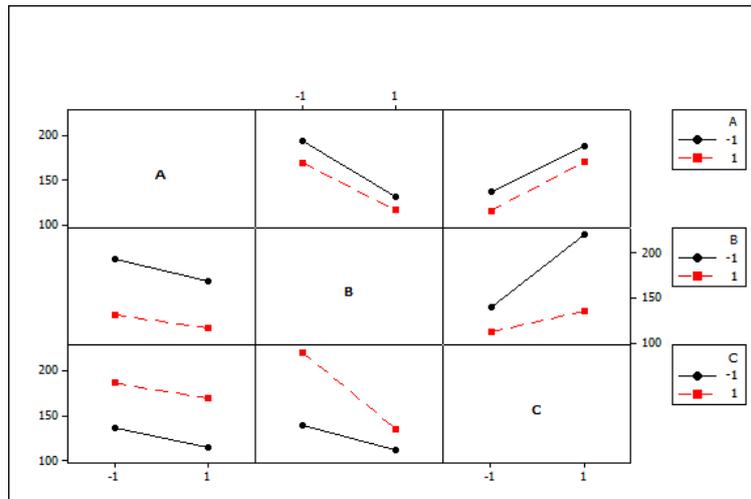


Figura 2.4: **Gráfico das Interações de Segunda Ordem.** Representação geométrica de como as interações de dois fatores afetam o tempo de dissolução dos comprimidos efervescentes de vitamina C.

Dos gráficos mostrados na Figura 2.4 percebe-se que não há interação entre os efeitos A e B , e nem tampouco entre os efeitos A e C , uma vez que as retas que representam as combinações destes fatores são paralelas; já entre os fatores B e C as retas não são paralelas, sendo isto uma evidência de que a interação BC é significativa. Desta forma, as informações contidas neste gráfico, bem como no gráfico da Figura 2.3, são uma confirmação gráfica dos resultados apontados pela ANOVA: os efeitos A e BC são significativos neste experimento. Desta forma, o comprimido da marca Bio-C, colocado num recipiente com 100 ml de água a $14,5^{\circ}\text{C}$, terá o menor tempo de dissolução.

Este exemplo foi bastante simples, e teve apenas um objetivo didático: apresentar idéias essenciais dos experimentos fatoriais 2^k . Há muito mais a ser dito, mas isso foge do objetivo deste trabalho. Para maiores aprofundamentos sobre os fatoriais 2^k completos, tais como a construção da ANOVA, o modelo de regressão adjacente a este tipo de experimento, os casos de experimentos com blocos, dentre outros; pode-se consultar a bibliografia indicada na introdução deste capítulo.

2.2.2 Experimentos Fatoriais 2^{k-p}

A situação descrita acima muitas vezes é chamada de experimento fatorial 2^k completo, pois todos os tratamentos foram utilizados no experimento. Como já afirmado anteriormente, nos experimentos fatoriais, à medida que o número de fatores aumenta, o número total de tratamentos utilizados também aumenta. Por exemplo, num fatorial 2^6 é necessário aplicar 64 tratamentos para analisar os efeitos dos fatores, dentre os quais, apenas 6 são os efeitos principais, 15 referem-se às interações de segunda ordem, 20 às interações de terceira ordem, 15 a interações de quarta ordem, 6 às interações de quinta ordem e apenas um refere-se à interação de todos os fatores. E nos casos em que o experimento é replicado, o número de provas pode ser bem maior. Esta seção aborda o tipo de planejamento no qual é possível obter informações úteis utilizando um número de unidades experimentais menor que o número total de tratamentos. Estes são conhecidos como experimentos fatoriais 2^k fracionados ou simplesmente experimentos fatoriais 2^{k-p} ; e são amplamente utilizados na experimentação industrial (LOUGHIN, 1998; ANTONY, 1999; YE; HAMADA, 2000; LAWSON, 2008).

De acordo com o Princípio da Ordenação Hierárquica considera-se, em geral, que interações de alta ordem possuem valores próximos de zero e não apresentam importância prática (BIRNBAUM, 1959). Desta forma, estimá-las passa a não ser prioridade no experimento e torna-se desnecessário, então, utilizar todos os tratamentos de um fatorial completo, em especial quando muitos fatores estão sendo estudados. Assim, o pesquisador pode utilizar apenas uma fração do experimento completo para obter informações dos efeitos principais e das interações de baixa ordem. Por exemplo, considere um fatorial 2^4 do qual se deseja obter uma fração $\frac{1}{2}$. Isso resultará num 2^{4-1} com oito tratamentos a considerar, ao invés dos 16 originais. A idéia é usar a interação $ABCD$ para gerar a fração, utilizando a usual notação $I = ABCD$, chamada de **relação definidora**. I representa a identidade (coluna de 1s).

Assim, por exemplo, para encontrar o efeito confundido com o fator D e que servirá para gerar o plano fatorial, ou seja, determinar os oito tratamentos a serem executados; o procedimento é o seguinte⁴:

⁴Tendo em mente as seguintes observações: (i) a operação de multiplicação é definida do modo usual; (ii) aplica-se a regra do cancelamento; (iii) como $(-1)^2 = (1)^2 = 1$, vale a definição de que $A^2 = B^2 = C^2 = \dots = I$.

$$\begin{aligned}
I &= ABCD \times D \\
D \times I &= ABCD \times D \\
D &= ABCD^2 \\
D &= ABCI \\
D &= ABC
\end{aligned}$$

Desta forma, deve-se fazer $D = ABC$ (mas poderia também fazer $-D = ABC$) e analisar o experimento como se fosse um fatorial completo 2^3 . Uma ilustração é dada na Tabela 2.5.

Tabela 2.5: Construção de um experimento fatorial fracionado 2^{4-1} .

Tratamento	<i>A</i>	<i>B</i>	<i>C</i>	$D = ABC$	Resposta
(1)	-	-	-	-	y_1
<i>a</i>	+	-	-	+	y_2
<i>b</i>	-	+	-	+	y_3
<i>ab</i>	+	+	-	-	y_4
<i>c</i>	-	-	+	+	y_5
<i>ac</i>	+	-	+	-	y_6
<i>bc</i>	-	+	+	-	y_7
<i>abc</i>	+	+	+	+	y_8

Pela Tabela 2.5, pode-se perceber que estimar o efeito principal D será o mesmo que estimar o efeito da interação ABC . Em outras palavras, o efeito D na verdade será o resultado da soma do efeito D com a interação ABC . Neste caso, diz-se que estes efeitos estão **confundidos**. O inteiro sistema de confundimento pode ser visto em resumo na relação definidora. A multiplicação de ambos os lados da “equação” $I = ABCD$ por qualquer letra ou conjunto de letras no grupo dos efeitos produz um par de efeitos que são aqueles medidos juntos ou confundidos. Seguindo este mesmo raciocínio, pode-se identificar todas as relações de confundimento para este exemplo, através da seguinte configuração:

$$\begin{aligned}
\ell_A &= A + BCD \\
\ell_B &= B + ACD \\
\ell_C &= C + ABD \\
\ell_D &= D + ABC \\
\ell_{AB} &= AB + CD \\
\ell_{AC} &= AC + BD
\end{aligned}$$

$$\ell_{AD} = AD + BC$$

Percebe-se que todos os efeitos principais estão confundidos com interações de terceira ordem. Esta situação é ideal, pois, assumindo que as interações de terceira e quarta ordem não são estatisticamente significativas, as estimativas ℓ_A , ℓ_B , ℓ_C , ℓ_D podem ser usadas para estimar os efeitos principais de A , B , C e D . Isso deixa claro a importância da escolha do **gerador** da fração. Um bom gerador seria aquele que não “criaria” uma relação de confundimento entre efeitos principais, nem entre efeitos provavelmente ativos (tais como as interações de baixa ordem).

As referências citadas no início deste capítulo trazem informações mais aprofundadas sobre estes planos, e discutem outros tópicos relacionados, tais como frações maiores, a questão da resolução de uma fração, os modelos relacionados, a ANOVA, a blocagem, dentre outros.

2.3 Experimentos Fatoriais 2^k e 2^{k-p} Não Replicados

Como já comentado, nos experimentos fatoriais 2^k , quanto maior o número de fatores, maior será também a quantidade de tratamentos considerados. Por conta disso, em muitas situações, realizar réplicas de um experimento pode ser inviável, do ponto de vista do tempo e dos recursos disponíveis. Sendo assim, alguns experimentos são realizados sem replicação, isto é, há apenas uma única prova para cada combinação de níveis dos fatores. Experimentos com uma única réplica são chamados de **fatoriais não replicados**. Estes são muito utilizados no setor industrial, com o intuito principal de reduzir os custos por causa da economia que eles oferecem na quantidade de provas, tendo em vista o grande número de fatores utilizados nestes experimentos (LOEPPKY; SITTER, 2002; VOSS; WANG, 1999). Os experimentos fatoriais não replicados são os arranjos experimentais utilizados nesta pesquisa.

A principal consequência de não ter réplica é o fato de não haver grau de liberdade para estimar o erro experimental. Nesta situação, os efeitos ativos não podem ser identificados usando as estatísticas usuais t ou F , pois para a construção da ANOVA, a parcela referente ao erro experimental é de fundamental importância. Sem a medida do erro, é impossível construir tal teste padrão.

Ao ser confrontado com a situação de um plano não replicado, o investigador tem

três escolhas a fazer (BOX; HUNTER; HUNTER, 2005; MONTGOMERY, 2001):

- (1) considerar as interações de alta ordem como sendo não ativas, e daí utilizá-las para estimar o erro. Ao optar por tal escolha o investigador corre o risco de considerar como não ativa uma interação que realmente é ativa. Ao comentar sobre este popular método, Birnbaum (1959) atenta para o fato de que, embora em alguns experimentos as interações de alta ordem são inativas, a experiência em outras situações tem mostrado que isso nem sempre é verdade, o que pode causar um viés na estimativa de σ^2 .
- (2) utilizar os métodos e testes formais propostos para a análise desta situação (BOX; MEYER, 1986; LENTH, 1989); tais como os descritos na Seção 2.4.
- (3) construir gráficos de probabilidade (normal ou semi-normal) para julgar a significância dos efeitos. Esta é uma opção bastante usual, como se verá no Capítulo 4. Tais gráficos podem ajudar na avaliação da significância dos efeitos, como também em fornecer uma estimativa para o erro experimental.

Em vista de sua grande utilização e facilidade de construção, os gráficos de probabilidade normal serão o objeto de pesquisa deste trabalho. Realizar-se-á uma análise comparativa destes dois métodos a fim de verificar seu desempenho como instrumentos de análise nos planos fatoriais não replicados.

2.4 Métodos Propostos para Análise de Experimentos Fatoriais Não Replicados

Os gráficos de probabilidade normal e semi-normal ainda são os métodos mais utilizados na análise dos experimentos fatoriais não replicados (MONTGOMERY, 2001). Sua construção e fundamentação está detalhada no Capítulo 3. No entanto, a análise gráfica é subjetiva pois a decisão é feita visualmente, de acordo com a percepção e interpretação daquele que analisa o gráfico. Desta forma, visando eliminar a subjetividade dos métodos gráficos, muitos procedimentos formais têm sido propostos para quantificar o que pode ser apenas visualizado. Esta seção apresenta uma síntese de alguns dos principais métodos propostos para análise de experimentos fatoriais não replicados, encontrados na pesquisa bibliográfica.

Método de Daniel (1959) - *guard-rails*. De acordo com Daniel (1959), quando um ponto é marcado fora da reta do gráfico semi-normal, isto significa que o efeito que este representa é significativo. Quanto mais distante um ponto estiver da reta, mais forte será esta indicação. No entanto, como a interpretação de “distante” pode ser subjetiva, Daniel desenvolveu também um teste formal para avaliar a significância dos efeitos. Este teste serve para determinar as “faixas de segurança” (ou o que Daniel chamou de *guard-rails*) no gráfico semi-normal, que nada mais são que curvas construídas em torno da reta traçada no gráfico. Inicialmente, é determinado, dentre os valores absolutos ordenados dos efeitos estimados, $|y_1| < |y_2| < \dots < |y_n|$, aquele cuja posição j é o primeiro inteiro que satisfaz a condição $j \geq 0,683(n + 1)$. Em seguida, são calculados valores críticos⁵ para o teste dos efeitos, com três níveis de significância: $\alpha = 0,05$, $\alpha = 0,20$ e $\alpha = 0,40$. A idéia do teste é a seguinte: inicialmente, o efeito com maior valor absoluto é comparado com os valores críticos; se este for significativo, então o efeito com o segundo maior valor absoluto é comparado com seus respectivos valores críticos, e assim por diante. O procedimento termina quando é encontrado o primeiro efeito não significativo, ou quando todos os valores absolutos ordenados dos efeitos, determinados previamente por Zahn (1975a), cujas posições forem maiores que o j anteriormente definido, tiverem sido examinados. Daí, linhas (*guard-rails*) são marcadas no gráfico semi-normal. Daniel sugeriu construir três linhas, uma para cada um dos três diferentes níveis de α . De acordo com o teste, um efeito é considerado significativo quando seu ponto correspondente é marcado acima de qualquer uma dessas “faixas de segurança”. Daniel comenta que pode ser conveniente construir o gráfico semi-normal com as faixas de confiança marcadas afim de se julgar a significância dos efeitos. Para informações mais detalhadas sobre esse teste consultar o artigo de Daniel (1959), bem como Birnbaum (1959); Zahn (1975a, 1975b); Taylor (1994); Olguin e Fearn (1997).

Método de Box e Meyer (1986). A técnica envolve calcular uma probabilidade a posteriori de que todos os efeitos sejam ativos, com base numa informação a priori resumida em dois parâmetros. Os resultados obtidos são apresentados num gráfico especial, popularmente conhecido como um *Bayes plot* ou gráfico de Bayes. Este nada mais é do que um gráfico de barras das probabilidades a posteriori calculadas, que

⁵Daniel (1959) calculou inicialmente tais valores, mas Zahn (1975a, 1975b) identificou alguns erros em tais cálculos e os corrigiu.

contém geralmente duas linhas horizontais marcando os valores 0,5 e 1,0 destas probabilidades, para ajudar na leitura dos resultados. Um efeito será considerado “ativo” ou estatisticamente significativo quando sua probabilidade a posteriori for maior que 0,5. Uma vez construído, o gráfico de Bayes é então apresentado juntamente com o gráfico normal, lado a lado, para reforçar ou confirmar os resultados da avaliação. O método de Box e Meyer (1986) trata-se, portanto, de uma aproximação bayesiana baseada no princípio da esparsidade dos efeitos.

Método de Lenth (1989). Trata-se de um método simples, alternativo ao método de Box e Meyer (1986) para uma análise formal de fatoriais não replicados. Também é baseado no princípio da esparsidade dos efeitos. Lenth definiu um pseudo erro padrão (*pseudo standard error*) PSE dos contrastes como sendo $PSE = 1,5 \cdot \text{mediana}_{|c_j| < 2,5s_0} |c_j|$, em que $|c_j|$ são as estimativas dos efeitos e $s_0 = 1,5 \cdot \text{mediana}_{|c_j| < 2,5s_0} |c_j|$. Definiu também uma margem de erro $ME = t_{(0,975;d)} \cdot PSE$ e uma margem de erro simultânea (*simultaneous margin of error*) $SME = t_{(\gamma;d)} \cdot PSE$, em que $d = n/3$, $\gamma = (1 + 0,95^{1/n})/2$ e n é o número de observações. Os efeitos são, então, apresentados num gráfico, com uma estrutura semelhante ao gráfico de Bayes ou um gráfico de barras, no qual seus valores numéricos (com os sinais correspondentes) são exibidos e os limites horizontais $\pm ME$ e $\pm SME$ são adicionados ao gráfico. A análise é feita do seguinte modo: um efeito cujo valor ultrapassa SME é claramente significativo; aquele que não ultrapassa ME não pode ser considerado significativo; e um efeito cujo valor está entre SME e ME encontra-se na região de incerteza, isto é, ele pode ou não ser significativo. Três vantagens principais deste novo método, de acordo com Lenth (1989), em comparação com o método Box-Meyer podem ser citadas. Primeiro, os resultados são dados em termos das unidades originais de medida. Esta associação direta com os dados torna a análise mais fácil de explicar, pois no método de Box-Meyer os resultados são dados em termos de probabilidades. Segundo, é possível avaliar o tamanho e a significância dos efeitos olhando apenas para um único gráfico ao invés de dois, o gráfico de Bayes e o gráfico normal, como acontece no método Box-Meyer. Terceiro, uma vantagem adicional (não tão relevante nesta era dos computadores) é que os cálculos requeridos são fáceis, podendo ser feitos à mão, considerando que o método Box-Meyer requer o desenvolvimento de *softwares* especializados. Fazendo jus ao título do artigo, percebe-se então, que o método de Lenth provê realmente uma

análise rápida e fácil do experimento.

Método de Dong (1993). Similar ao método de Lenth (1989), Dong (1993) propôs usar $s_1 = \sqrt{n^{-1} \sum_{|c_j| < 2,5s_0} c_j^2}$ no lugar do PSE e $\gamma = (1 + 0,98^{1/n})/2$. Segundo ele, seu método é melhor e mais simples de usar que o método de Lenth (1989).

Método de Lawson, Grimshaw e Burt (1998). Trata-se de um método para julgar a significância dos efeitos, que é tanto numérico quanto gráfico: tem a simplicidade computacional do método de Lenth (1989) com a interpretação gráfica e o poder aumentado do método de Loh⁶. O método pode ser descrito em quatro passos: (1) inicialmente são encontrados os efeitos significativos como no método de Lenth (1989): são aqueles cujo valor absoluto é maior que $2,5 \cdot s_0$; (2) a reta de mínimos quadrados, sem intercepto, será ajustada ao modelo $|c|_{(j)} = \hat{\beta}_1 \cdot \Phi_o^{-1}(\hat{p}_j)$, em que Φ_o é a função de distribuição acumulada (fda) da distribuição semi-normal, e construída no gráfico semi-normal; (3) é calculada a estatística do teste para o julgamento dos contrastes; (4) são construídas bandas de predição ao redor da reta ajustada: os pontos marcados fora dos limites das bandas serão os efeitos julgados significativos. Os autores optaram pelo gráfico semi-normal porque, segundo eles, a disposição dos sinais “+” e “-” é arbitrária.

Método Lenth Descendente (*Step-Down Lenth Method*) (2001). Proposto por Ye, Hamada e Wu (2001), baseia-se no método de Lenth (1989), mas com uma diferença básica: ao passo que no método de Lenth (1989) o valor da mediana no PSE é calculado após a exclusão dos efeitos claramente ativos; no Método Lenth Descendente à medida que o maior efeito é julgado significativo, ele é separado dos demais e o PSE é recalculado com base nos efeitos restantes. Daí, o maior dos efeitos restantes é analisado, e o processo continua até que o primeiro efeito não significativo seja julgado.

Estes são apenas alguns dos vários métodos propostos para a análise de fatoriais não replicados. Como se verá no Capítulo 4, o método de Lenth é o método analítico mais citado e usado nos trabalhos em que a análise subjetiva dos métodos gráficos não é utilizada. Não se pretende nesta seção listar ou comentar todas as técnicas existentes

⁶LOH, W. Y. Identification of active contrasts in unreplicated factorial experiments. *Computational Statistics and Data Analysis*. n. 14, p. 135-148, 1992. (Não encontrado) Trata-se de uma proposta gráfica para identificar efeitos ativos num gráfico de probabilidade normal (LAWSON, GRIMSHAW; BURT, 1998).

para a análise dos fatoriais não replicados. Um artigo que pode servir de consulta é o de Hamada e Balakrishnan (1998), que traz um resumo das principais técnicas usadas nos planos fatoriais 2^k não replicados, bem como compara estes métodos usando um extensivo estudo de simulação. Eles testaram o poder e a eficácia desses métodos. O método de Lenth (1989) foi recomendado por causa de sua facilidade nos cálculos, bem como por apresentar um bom desempenho.

Exemplo 2.2. (DANIEL, 1976; BOX; MEYER, 1986) Este experimento analisa a influência dos fatores A : carga, B : fluxo, C : velocidade de rotação e D : tipo de lama usada na perfuração; sobre o desempenho de uma broca. Trata-se de um fatorial 2^4 não replicado. Os efeitos dos fatores são os seguintes: $A = 0,06$, $B = 0,25$, $AB = -0,01$, $C = 0,50$, $AC = 0$, $BC = -0,02$, $ABC = 0$, $D = 0,14$, $AD = 0,03$, $BD = -0,01$, $ABD = 0,02$, $CD = -0,04$, $ACD = 0,02$, $BCD = -0,01$, $ABCD = 0,02$. Deseja-se investigar quais destes efeitos são significativos. Como ilustração, serão utilizados os seguintes métodos: gráficos de probabilidade (normal e semi-normal), o método das faixas de confiança (DANIEL, 1959), o método de Box e Meyer e o método de Lenth. A intenção é comparar os resultados que estes produzem na indicação de qual dos efeitos é significativo. A escolha destes métodos para a análise deste exemplo, como se verá no Capítulo 4, deve-se ao fato de que, com exceção do método das faixas de confiança, são os mais usados e indicados para a análise dos planos fatoriais 2^k não replicados.

Gráfico de Probabilidade Normal. De acordo com o gráfico mostrado na Figura 2.5, os efeitos *B*, *C* e *D* são claramente significativos.

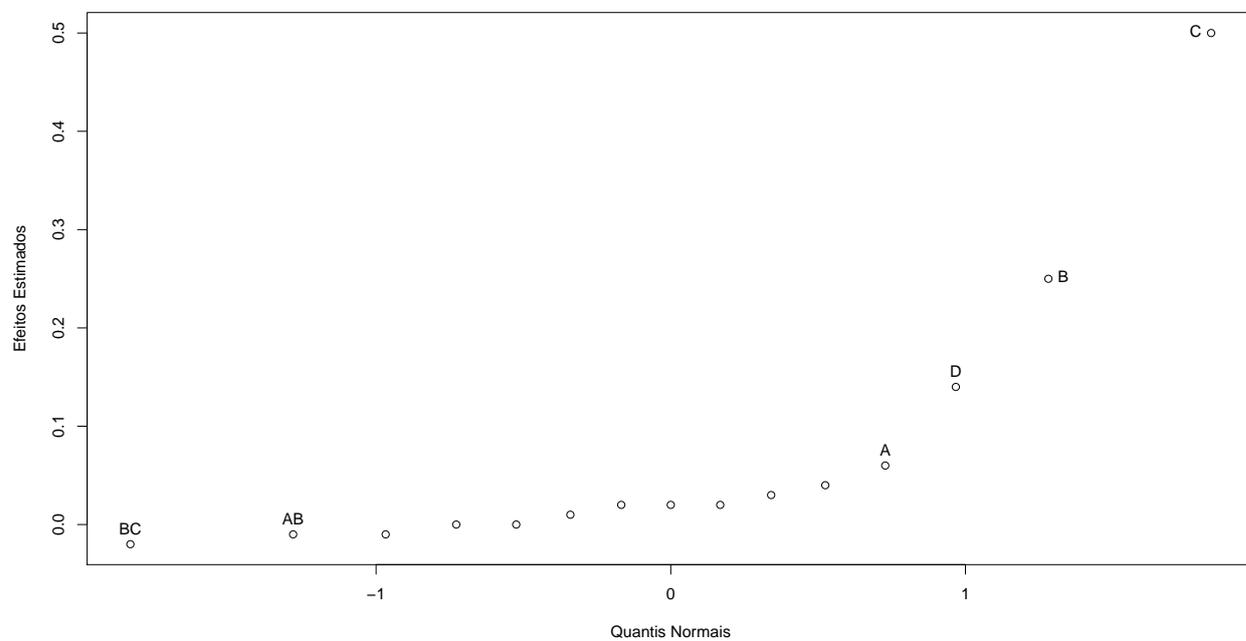


Figura 2.5: Gráfico Normal para os dados do Exemplo 2.2.

Gráfico Semi-Normal. O gráfico semi-normal mostrado na Figura 2.6 fornece evidências para se considerar apenas os efeitos B , C e D como significativos.

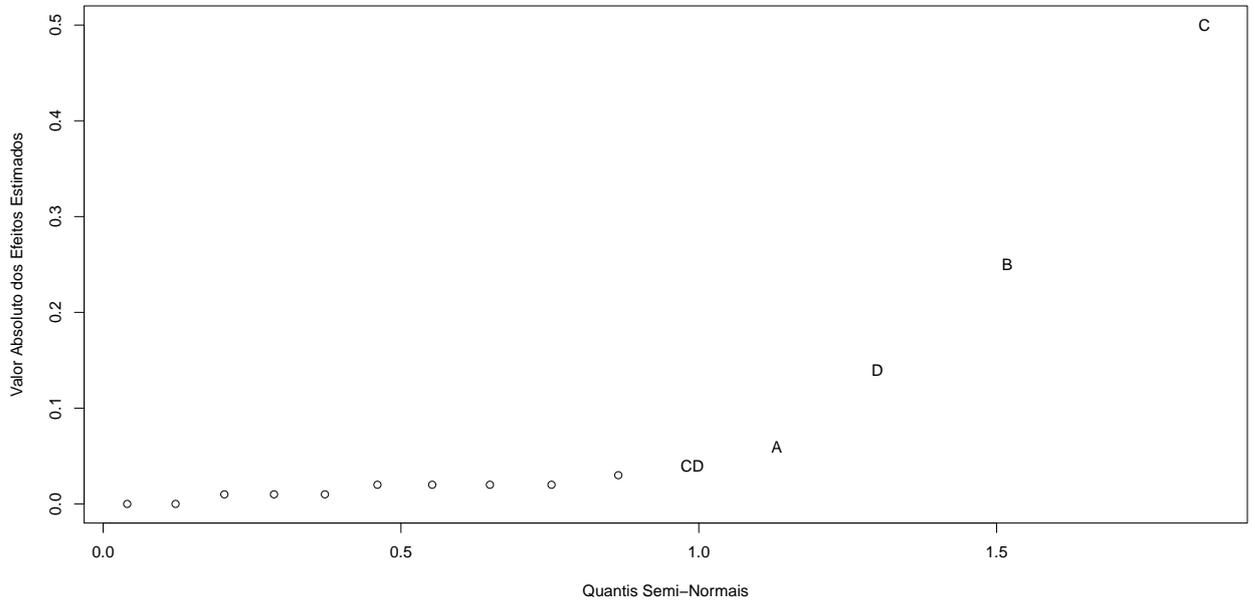


Figura 2.6: Gráfico Semi-Normal para os dados do Exemplo 2.2.

Método de Daniel (1959) - *guard-rails*. Inicialmente calcula-se os valores absolutos ordenados dos efeitos estimados, que pode ser visto na Tabela 2.6. Daí, para $n = 15$ temos que $0,683(n + 1) = 0,683(16) = 10,928$; o que faz de $|y_{11}| = 0,04$ a estimativa inicial de σ que será usada na determinação das faixas. Em seguida, calculam-se as coordenadas das faixas de confiança (*guard-rails*) de acordo com os valores críticos de Zahn (1975a), que, para o caso de $n = 15$, sugere usar para tal cálculo apenas os efeitos $|y_{12}|$, $|y_{13}|$, $|y_{14}|$ e $|y_{15}|$. Os resultados encontram-se na Tabela 2.7.

Tabela 2.6: Valores absolutos ordenados dos efeitos estimados do Exemplo 2.2.

Ordem	Efeito	Valor Absoluto
1	<i>AC</i>	0,00
2	<i>ABC</i>	0,00
3	<i>AB</i>	0,01
4	<i>BD</i>	0,01
5	<i>BCD</i>	0,01
6	<i>BC</i>	0,02
7	<i>ABC</i>	0,02
8	<i>ACD</i>	0,02
9	<i>ABCD</i>	0,02
10	<i>AD</i>	0,03
11	<i>CD</i>	0,04
12	<i>A</i>	0,06
13	<i>D</i>	0,14
14	<i>B</i>	0,25
15	<i>C</i>	0,50

Tabela 2.7: Cálculo das coordenadas das faixas de confiança para o Exemplo 2.2.

Ordem	$\alpha = 0,05$	$\alpha = 0,20$	$\alpha = 0,40$
15	$3,230 \cdot 0,04 = 0,129$	$2,470 \cdot 0,04 = 0,099$	$2,066 \cdot 0,04 = 0,083$
14	$2,840 \cdot 0,04 = 0,114$	$2,177 \cdot 0,04 = 0,087$	$1,827 \cdot 0,04 = 0,073$
13	$2,427 \cdot 0,04 = 0,097$	$1,866 \cdot 0,04 = 0,075$	$1,574 \cdot 0,04 = 0,063$
12	$2,065 \cdot 0,04 = 0,083$	$1,533 \cdot 0,04 = 0,061$	$1,298 \cdot 0,04 = 0,052$

O próximo passo agora é construir o gráfico semi-normal com as faixas de confiança marcadas. A Figura 2.7 mostra o gráfico para os dados do Exemplo 2.2. Como se pode ver pelo referido gráfico, segundo o Método de Daniel (1959) dos *guard-rails*, os efeitos *B*, *C* e *D* são julgados significativos uma vez que foram marcados acima das faixas de confiança calculadas. A importância do efeito *A* é duvidosa, uma vez que não é tão claro que ele tenha sido marcado entre as faixas de 20% e 40%, isto é, não está bem claro se o ponto *A* está fora das faixas ou entre elas.

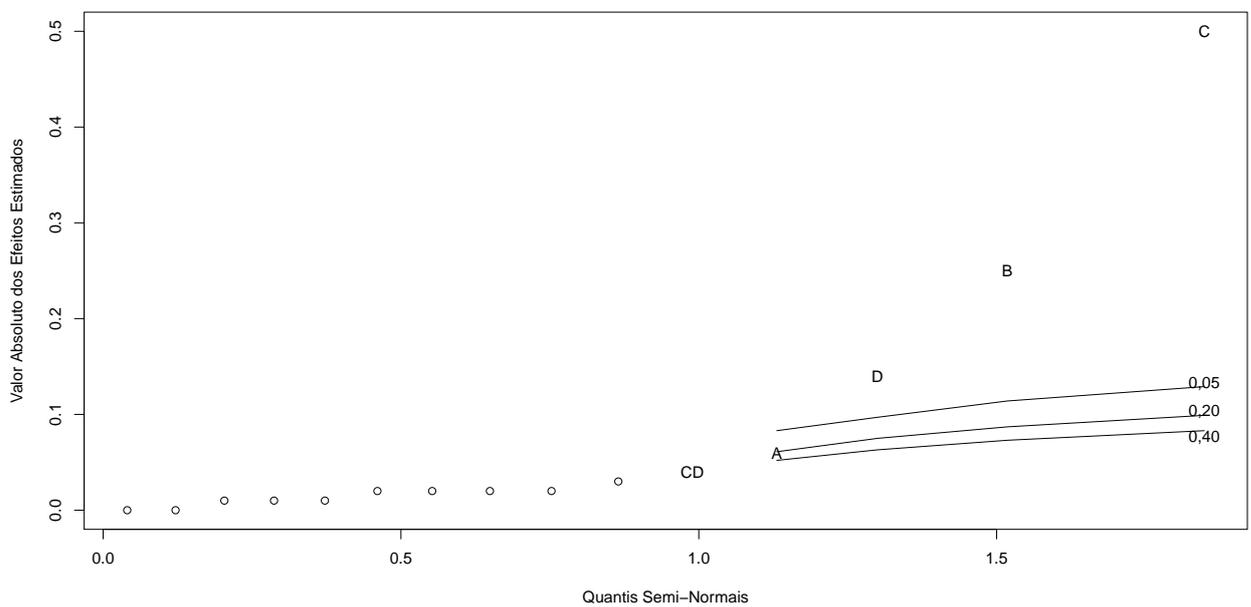


Figura 2.7: Gráfico Semi-Normal com as faixas de confiança para os dados do Exemplo 2.2.

Método de Box e Meyer (1986). Os pacotes e funções utilizados para a construção do Método de Box e Meyer (1986) bem como do Método de Lenth (1989), mencionados nos exemplos a seguir, são do R. Com o auxílio do pacote `BsMD`, no qual os dados do Exemplo 2.2 já estão disponíveis; por meio da função `BsProb` pode-se calcular as probabilidades a posteriori de todos os efeitos estimados. A Figura 2.8 mostra o gráfico de Bayes dos resultados obtidos. Por meio de tal gráfico apenas os efeitos B , C e D são julgados significativos uma vez que suas probabilidades foram maiores que 0,5. Através do método de Box e Meyer o efeito A não é significativo, esclarecendo assim a dúvida surgida na aplicação do método das faixas de segurança.

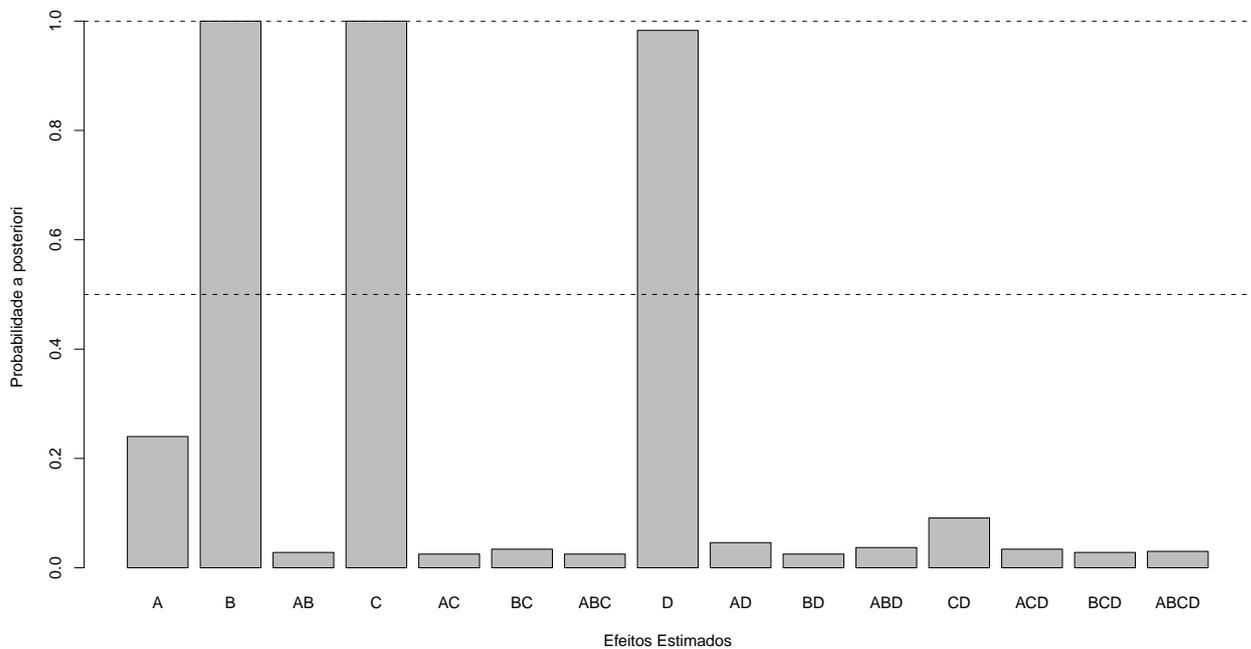


Figura 2.8: Gráfico de Bayes para os dados do Exemplo 2.2.

Método de Lenth (1989). Novamente com o auxílio do pacote BsMD pode-se realizar os cálculos necessários para a construção do gráfico resultante do método de Lenth. Tal gráfico é mostrado na Figura 2.9, do qual se pode concluir que os efeitos B e C são claramente significativos, uma vez que ultrapassaram o valor de SME . O efeito D encontra-se na região de incerteza, pois seu valor está entre SME e ME . Uma decisão a favor da significância de D seria apropriada por dois motivos: primeiro, porque seu valor está bem próximo do valor de SME e, segundo, porque o efeito D foi julgado significativo por todos os outros métodos usados acima.

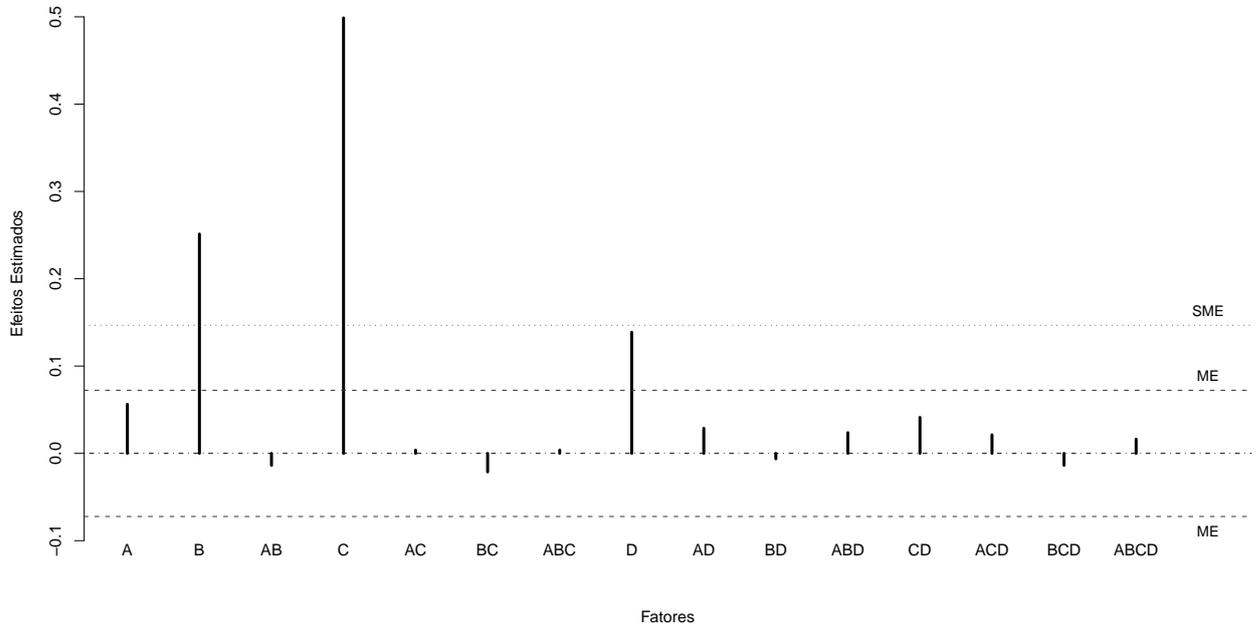


Figura 2.9: Gráfico de Lenth para os dados do Exemplo 2.2.

Os métodos utilizados para a análise do Exemplo 2.2 apontaram algumas conclusões interessantes. Primeiramente, pôde-se perceber a subjetividade das técnicas gráficas. Quando um efeito é marcado a uma distância considerável da tendência linear da maioria dos pontos, é clara sua significância. Isso ocorreu com o julgamento dos efeitos B , C e D nos gráficos normal e semi-normal. No entanto, quando essa distância não é tão grande assim, surgem dúvidas no julgamento dos efeitos: o que é “distante” para alguém, pode não ser para outro, acarretando julgamentos diferentes sobre um mesmo efeito; algo possível de acontecer com o efeito A nos gráficos normal e semi-normal. Tal situação demonstra ainda que os gráficos de probabilidade podem apontar para julgamentos diferentes, divergindo assim seus resultados.

Outra situação que este exemplo chama atenção é o fato de que mesmo as técnicas formais, analíticas, também podem divergir. Mas são extremamente úteis para esclarecer as dúvidas que surgem nos métodos gráficos, conforme discutido no parágrafo anterior; bem como podem ser melhor entendidas ou confirmadas pelo uso dos gráficos de probabilidade. Em vista disso, os resultados encontrados do Exemplo 2.2 mostram que o ideal seria utilizar as técnicas gráficas juntamente com os métodos formais; pois suas análises se complementam, contribuindo assim para um processo de inferência com decisões mais acertadas. Na verdade, é exatamente isso que Box e Meyer (1986) e Lenth (1989) recomendam. Portanto, em vista de tudo o que foi anteriormente discutido, pode-se dizer que, para o experimento da broca, do Exemplo 2.2, os candidatos a efeitos significativos são B , C e D .

2.5 Sugestões de Nomenclatura

Um grande número de termos utilizados em experimentação ainda são muito conhecidos em sua forma inglesa. Dentre os objetivos propostos para este trabalho encontra-se o de apresentar sugestões de nomenclatura em português para tais termos, no contexto dos aspectos teóricos aqui abordados. Nesta dissertação, optou-se por utilizar os termos em português, tendo em vista a padronização da linguagem em todo o texto; e também servir como um reforço ao uso de tais termos em língua materna. Portanto, é útil apresentar algumas considerações neste sentido.

Utilizando as informações contidas no glossário do *ISI - International Statistical*

Institute, cujo acesso pode ser feito através do endereço eletrônico <http://isi.cbs.nl>, como também o conhecimento construído por meio dos estudos em planejamento de experimentos; apresentam-se as seguintes sugestões de nomenclatura:

- *balanced*: balanceado
- *bias*: viés; tendência
- *biased*: viciado; tendencioso
- *box-plot*: diagrama de caixas
- *folded distributions*: distribuições dobradas
- *full-normal plot* ou *normal probability plot*: gráfico normal; gráfico de probabilidade normal; gráfico probabilístico normal
- *guard-rails*: faixas de segurança; faixas de confiança
- *half-normal distribution*: distribuição semi-normal
- *half-normal plot*: gráfico semi-normal; gráfico de probabilidade semi-normal; gráfico probabilístico semi-normal
- *half-normal probability paper*: papel de probabilidade semi-normal; papel probabilístico semi-normal
- *hidden replications*: réplicas escondidas
- *inadvertent plot-splitting*: parcelas subdivididas inadvertidas; parcelas subdivididas negligenciadas
- *normal probability paper*: papel de probabilidade normal; papel probabilístico normal
- *one-factor-at-a-time experiments*: experimentos com um fator de cada vez
- *outliers*: valores discrepantes; valores atípicos; valores aberrantes
- *probability plot*: gráfico de probabilidade; gráfico probabilístico

- *probability-probability plot* ou *p-p plot*: gráfico percentil-percentil, gráfico dos percentis ou gráfico p-p
- *quantile-quantile plot* ou *q-q plot*: gráfico quantil-quantil; gráfico dos quantis ou gráfico q-q
- *residual plot*: gráfico de resíduos
- *run*: prova
- *screening experiments*: experimentos de triagem fatorial; experimentos exploratórios
- *split-plot*: experimento em parcelas subdivididas
- *step-down Lenth method*: método Lenth descendente
- *strip-plot*: experimento em faixas
- *sub-plot factors*: fatores da subparcela
- *whole-plot factors*: fatores da parcela

Capítulo 3

Gráficos de Probabilidade Normal

Este capítulo aborda os principais conceitos, definições, construções e usos dos gráficos de probabilidade abordados nesta pesquisa: o gráfico normal e o gráfico semi-normal. Destaca-se também a controvérsia acerca do uso destes gráficos na análise dos planos fatoriais não replicados.

3.1 Os Gráficos e a Estatística

Desde a década de 1960, devido, principalmente, ao progresso computacional, tem aumentado o interesse e a pesquisa para desenvolver e aperfeiçoar métodos gráficos. De acordo com Cleveland (1987), existem três grandes áreas de pesquisa sobre esse assunto: (1) métodos: envolve os estudos de quantificação (área que aborda a informação quantitativa que deve ser mostrada num gráfico) e visualização (trata do meio visual que deve ser usado para mostrar o conjunto das informações quantitativas); (2) computação: este tópico de pesquisa trata dos algoritmos de programação dos gráficos, bem como da construção da interface com o usuário; (3) percepção gráfica: visto que as informações estão codificadas nos aspectos do gráfico, a percepção gráfica investiga o processo de decodificação destas informações. Os estudos nessa área de pesquisa geralmente são feitos por meio de experimentos em que várias pessoas são submetidas à observação dos mesmos gráficos, a fim de se verificar a percepção visual destas sobre as informações que tais gráficos fornecem. Os resultados proveem critérios para a avaliação e comparação dos vários métodos gráficos.

Os gráficos tornaram-se uma importante ferramenta estatística (WILK; GNANADESIKAN, 1964). São utilizados desde as fases iniciais da investigação, até a divul-

gação dos resultados finais de uma pesquisa. Quanto mais simples e auto-explicativo for um gráfico, no sentido de clareza de interpretação, mais eficiente será seu uso. São muitos os tipos de gráficos que existem: histograma, diagrama de caixas, gráfico de probabilidade, gráfico de dispersão e assim por diante; que são úteis aos mais diversos propósitos. Alguns desses são gráficos para exploração, ou seja, gráficos que são usados para investigar as características dos dados ou buscar a relação entre as variáveis em estudo, a fim de indicar as análises apropriadas. Há também gráficos para análise de dados, isto é, gráficos utilizados na avaliação da distribuição dos dados, na adequabilidade dos modelos, bem como na tomada de decisões. Outros gráficos resumem as análises feitas e divulgam os resultados do estudo.

Destacam-se também alguns comentários acerca do uso de gráficos como ferramentas de análise estatística: *“These comparatively simple tools are of great help in extracting the maximum amount of information from data. Graphical techniques are less formal and confining. They aid in understanding the numerous relationships reflected in the data. They help reveal departures from the assumed models and statistical distributions. They help reveal the existence of peculiar looking observations or subsets of the data. Graphical data displays often uncover features of the data that were totally unanticipated prior to the analysis. It is difficult to obtain similar information from numerical procedures . . . The relationship between graphical and numerical techniques is analogous to the relationship between a police detective and a judge. The graphical techniques are the counterpart of the detective looking for any clues to help uncover the mysteries of the data. The formal, numerical inference tools of classical statistics are the counterpart of the judge who weighs the degree of evidence contained in the clues to determine how much credence to put into them. A further advantage of graphical techniques is that they are computationally simple.”* (FEDER, 1974, p. 287) *“Probably the single most powerful tool with which the results of an experiment can be studied is a collection of plots of the raw and transformed data.”* (GERSON, 1975, p. 235) Uma vez que pode ser difícil tratar de técnicas matemáticas com pesquisadores de outras áreas, as técnicas gráficas podem ser úteis nesses casos, uma vez que estas parecem ser ferramentas de conhecimento comum às ciências em geral. (TAYLOR, 1994)

O presente trabalho tratará dos gráficos probabilísticos. O objetivo principal destes gráficos é comparar duas distribuições de probabilidade, sendo uma delas a

distribuição empírica, isto é, derivada dos dados observados no experimento. A outra distribuição é uma distribuição teórica conhecida, que está sendo avaliada se ela é adequada para descrever a distribuição dos dados disponíveis. O propósito básico destes gráficos é determinar, visualmente, se os dados observados provêm dessa dada distribuição teórica. Em caso afirmativo, isto é, se a distribuição empírica for similar à hipotetizada, o comportamento esperado será o de um conjunto de pontos se alinhando em torno de uma reta.

Os gráficos de probabilidade são um complemento informal, visual, de ajuda nos processos de inferência e muito utilizados nos arranjos fatoriais. (HEALY, 1968; WILK; GNANADESIKAN, 1968, 1970; GERSON, 1975; COX, 1984; STIRLING, 1982; HOLMGREN, 1995) Por exemplo, no caso dos fatoriais replicados, os gráficos de probabilidade podem ser usados para complementar as análises feitas pela ANOVA, bem como na análise dos resíduos, a fim de verificar a presença de algum valor discrepante. Se o experimento não tiver sido aleatorizado, estes gráficos também podem ser úteis para avaliar se houve, ou não, violação de algumas suposições do modelo. Já nos experimentos não replicados, tais gráficos são indicados para fornecer um julgamento visual, subjetivo, acerca da significância dos efeitos, assim como para fornecer uma estimativa do erro experimental. Dois tipos principais de tais gráficos, objetos desta pesquisa, são o gráfico normal e o semi-normal. As Seções seguintes discutem em maiores detalhes estas duas ferramentas.

3.2 Gráficos Probabilísticos em uma Família de Localização-Escala

Considere um conjunto de observações y_1, y_2, \dots, y_n . Deseja-se averiguar se estas observações podem ser tomadas como realizações de variáveis aleatórias Y_1, Y_2, \dots, Y_n que são independentes e identicamente distribuídas (i.i.d.) com função de distribuição acumulada F . Suponha que F forma uma família de localização-escala, ou seja, para todo $y \in \mathbb{R}$,

$$F(y) = G\left(\frac{y - \mu}{\sigma}\right), \quad (3.1)$$

em que G denota uma função de distribuição acumulada contínua que não depende de parâmetros desconhecidos e que tenha inversa, $\mu \in \mathbb{R}$ é o parâmetro de localização e $\sigma \in \mathbb{R}_+^*$ é o parâmetro de escala. Se $y_{(1)} \leq y_{(2)} \leq \dots \leq y_{(n)}$ denota os valores ordenados da amostra y_1, y_2, \dots, y_n , então segue que

$$p_j = F(y_{(j)}) = G\left(\frac{y_{(j)} - \mu}{\sigma}\right), \quad (3.2)$$

para todo $j = 1, 2, \dots, n$. Assim, temos que

$$G^{-1}(p_j) = G^{-1}\left(G\left(\frac{y_{(j)} - \mu}{\sigma}\right)\right) = \frac{y_{(j)} - \mu}{\sigma}, \quad (3.3)$$

o que implica que

$$y_{(j)} = \mu + \sigma \cdot G^{-1}(p_j) \quad (3.4)$$

para todo $j = 1, 2, \dots, n$. Portanto, se a fda F representa uma distribuição adequada para as observações y_1, y_2, \dots, y_n , então espera-se que os pontos $(G^{-1}(p_j), y_{(j)})$ se alinhem em torno de uma reta cujo intercepto é μ e coeficiente angular σ , conforme deixa claro a Equação 3.4. Um gráfico assim construído é denominado gráfico probabilístico ou gráfico de probabilidade. Contudo, como F é desconhecida, não se pode calcular os valores $p_1 = F(y_{(1)}), p_2 = F(y_{(2)}), \dots, p_n = F(y_{(n)})$. Um procedimento que pode ser aplicado para contornar este problema consiste na estimação da fda F pela função de distribuição empírica da amostra y_1, y_2, \dots, y_n , cuja definição é dada por:

$$\hat{F}_n(y) = \begin{cases} 0, & \text{se } y < Y_{(1)} \\ \frac{j}{n}, & \text{se } Y_{(j)} \leq y < Y_{(j+1)} \\ 1, & \text{se } y \geq Y_{(n)}. \end{cases} \quad \text{com } j = 1, 2, \dots, n-1 \quad (3.5)$$

para todo $y \in \mathbb{R}$.

Este estimador de F tem propriedades estatísticas bastante desejáveis, como por

exemplo¹:

i)

$$E \left[\widehat{F}_n(y) \right] = F(y),$$

para todo $y \in \mathbb{R}$. O que significa que $\widehat{F}_n(y)$ é um estimador não viciado para $F(y)$.

ii)

$$\text{Var} \left[\widehat{F}_n(y) \right] = \frac{F(y)[1 - F(y)]}{n}$$

Consequentemente, $\widehat{F}_n(y)$ tem uma variância pequena quando n é suficientemente grande (BLOM, 1989).

iii) $\widehat{F}_n(y)$ é consistente para $F(y)$, ou seja,

$$\widehat{F}_n(y) \xrightarrow{p} F(y), \quad n \rightarrow \infty$$

iv)

$$\frac{\sqrt{n} \left[\widehat{F}_n(y) - F(y) \right]}{\sqrt{F(y)[1 - F(y)]}} \xrightarrow{d} Z \sim N(0, 1), \quad n \rightarrow \infty$$

v) **(Teorema de Glivenko-Cantelli)** $\widehat{F}_n(y)$ converge uniformemente para $F(y)$, isto é, para $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} P \left(\sup_{y \in \mathbb{R}} |\widehat{F}_n(y) - F(y)| > \epsilon \right) = 0$$

Assim, o estimador correspondente aos valores p_j na Equação 3.2 é

$$\widehat{p}_j = \widehat{F}_n(y_{(j)}) = \frac{j}{n}, \quad (3.6)$$

¹Para maiores detalhes acerca dessas propriedades, bem como de suas demonstrações, consultar Rohatgi (1976).

para todo $j = 1, 2, \dots, n$. Contudo, para estimar os valores de $G^{-1}(p_j)$ em 3.3 por $G^{-1}(\hat{p}_j)$ da Equação 3.6 ocorre a seguinte limitação quando $j = n$

$$G^{-1}(\hat{p}_n) = G^{-1}\left(\frac{n}{n}\right) = G^{-1}(1) = +\infty$$

No intuito de resolver esta situação, algumas pequenas modificações são sugeridas para a Equação 3.6. Por exemplo, de acordo com Harter (1961), Blom² propôs uma aproximação para a j -ésima estatística de ordem normal (o j -ésimo menor desvio normal) de uma amostra de tamanho n por meio da seguinte relação:

$$\hat{p}_j = \frac{j - \alpha}{n - 2\alpha + 1}, \quad \text{com } j = 1, 2, \dots, n \quad (3.7)$$

Utilizando valores de $j = 1, 2, \dots, 200$ com $n = 25, 50, 100, 200, 400$, Blom tabulou os valores de α e percebeu que estes aumentam à medida que n aumenta, e supôs que α sempre se encontra no intervalo $(0, 33; 0, 50)$. Ele então sugeriu o uso de $\alpha = \frac{3}{8}$ como um valor apropriado. Desta forma, com o critério de Blom, temos:

$$\hat{p}_j = \frac{j - \frac{3}{8}}{n + \frac{1}{4}} = \frac{j - 0,375}{n + 0,25}, \quad \text{com } j = 1, 2, \dots, n \quad (3.8)$$

que possui boas propriedades: produz um insignificante viés na estimação de σ ; parece ter um mínimo erro quadrático médio para $\hat{\sigma}$; quando as observações são, de fato, normalmente distribuídas, produz pontos no gráfico que se ajustam aproximadamente a uma reta, com pequenos desvios (GERSON, 1975). Apesar disso, Harter (1961) comenta que este valor de α não é apropriado para pequenos valores de n .

A partir da Equação 3.7, outras modificações para \hat{p}_j , além daquela sugerida por Blom, também foram propostas. Por exemplo, para $\alpha = 0$, temos

$$\hat{p}_j = \frac{j}{n + 1}, \quad \text{com } j = 1, 2, \dots, n \quad (3.9)$$

²BLOM, Gunnar. *Statistical estimates and transformed beta-variables*. New York: John Wiley and Sons, Inc., 1958.

e para $\alpha = 0,5$

$$\hat{p}_j = \frac{j - 0,5}{n}, \quad \text{com } j = 1, 2, \dots, n \quad (3.10)$$

A expressão de \hat{p}_j na Equação 3.9 é citada por alguns autores como sendo uma forma apropriada para ser usada no gráfico dos percentis³ por prover boas aproximações para os valores médios de $\hat{F}_n(y)$ (STIRLING, 1982; GAN; KOEHLER, 1990); embora segundo Gerson (1975), tal expressão produz um estimador viciado para σ . A forma descrita na Equação 3.10 é a mais utilizada para a estimativa do \hat{p}_j (DANIEL, 1956; WILK; GNANADESIKAN; HUYETT, 1962; GERSON, 1975; STIRLING, 1982); por prover as melhores estimativas possíveis para os parâmetros da distribuição amostral (BLOM, 1989; MEEKER; ESCOBAR, 1998). Embora as expressões acima obtidas para $\alpha = 0$ e $\alpha = 0,5$ sejam as formas de \hat{p}_j mais conhecidas e utilizadas (ANSCOMBE; TUKEY, 1963), Harter (1961) afirma que nenhuma delas está correta; vindo a preferir a aproximação de Blom, conforme descrita na Equação 3.8. Por outro lado, Wilk e Gnanadesikan (1968) comentam que tais expressões não serão muito diferentes uma da outra, exceto em situações em que o tamanho da amostra seja muito pequeno. Gan e Koehler (1990) afirmam que, geralmente, os gráficos de probabilidade não são visivelmente afetados pela escolha do \hat{p}_j . Como se pode ver destas poucas referências, a escolha da forma apropriada do \hat{p}_j parece ser uma questão subjetiva, e tem sido um assunto de contínua investigação (WILK; GNANADESIKAN; HUYETT, 1962).

Para a execução dos exemplos e cenários apresentados nesta dissertação, bem como a construção dos gráficos de probabilidade propostos, será utilizado o pacote estatístico livre R na versão R 2.10.0. Portanto, cabe aqui uma observação acerca de qual destas formas do \hat{p}_j esta versão do R usa. O gráfico normal é construído através da função `qqnorm()`, na qual, os quantis utilizados no gráfico são calculados através do comando `qqnorm(ppoints(n))` em que `ppoints()` (*plotting points* ou *probability points*) é a função que calcula os valores de \hat{p}_j , assim definida:

³A ser discutido na próxima seção.

```

> ppoints
function (n, a = ifelse(n <= 10, 3/8, 1/2))
{
  if (length(n) > 1)
    n <- length(n)
  if (n > 0)
    (1L:n - a)/(n + 1 - 2 * a)
  else numeric(0L)
}

```

Como se pode ver, a expressão geral para o \hat{p}_j usada na versão R 2.10.0 na construção do gráfico normal é aquela descrita na Equação 3.7, cujo valor de α depende do valor de n por meio de duas condições: (1) para $n \leq 10$ teremos $\alpha = \frac{3}{8}$ e, portanto, a forma do \hat{p}_j utilizada será aquela do critério de Blom, conforme descrita na Equação 3.8; (2) para $n > 10$ teremos $\alpha = 0,5$, e o \hat{p}_j utilizado será o da forma da Equação 3.10. A função G na qual os quantis serão calculados é, pelo comando `qnorm(ppoints(n))`, a fda da distribuição normal padrão.

Por sua vez, para a construção do gráfico semi-normal utilizar-se-á a função `halfnorm()` disponível no R através do pacote `faraway`. Nesta função, os quantis usados no gráfico são calculados através do comando `qnorm((n + 1:n)/(2 * n + 1))` no qual $\hat{p}_j = \frac{j+n}{2n+1}$, uma forma não destacada acima, nem tampouco citada ou discutida entre os artigos encontrados⁴. Tal forma também difere da proposta original de Daniel (1959) para o gráfico semi-normal, na qual os quantis do gráfico eram calculados a partir da forma $\hat{p}_j = \frac{j-0,5}{n}$. Algumas simulações foram feitas para testar essas duas expressões de \hat{p}_j na construção do gráfico semi-normal, e não foram percebidas diferenças que pudessem comprometer a comparação dos gráficos. Por isso, não descartou-se o uso da função `halfnorm()`.

As considerações feitas nesta seção aplicam-se às formas gráficas que serão abordadas na Seção 3.3 como também na Seção 3.4, uma vez que a distribuição normal e a distribuição semi-normal são membros da família de localização-escala (PEWSEY, 2002).

⁴Ainda no pacote `faraway` há uma função chamada `qqnorm1()` que também constrói o gráfico normal. Uma grande diferença desta função para a já citada função `qqnorm()` é o uso da forma $\hat{p}_j = \frac{j}{n+1}$, conforme descrita na Equação 3.9. Outra diferença é que na função `qqnorm1()`, no lugar dos pontos, são grafados os números correspondentes às posições das estatísticas de ordem da amostra. Esta não é uma função muito conhecida nem muito utilizada para a construção do gráfico normal.

3.3 Gráfico de Probabilidade Normal

Seja X uma variável aleatória com distribuição normal padrão, isto é, $X \sim N(0, 1)$. A função de densidade de probabilidade de X é dada por:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty < z < \infty$$

e sua função de distribuição acumulada é dada pela expressão:

$$\Phi(z) = P(X \leq z) = \int_{-\infty}^z \phi(x) dx,$$

para todo $-\infty < z < \infty$.

O quantil p_j da distribuição normal pode ser definido como (MEEKER; ESCOBAR, 1998)⁵:

$$y_j = \mu + \sigma \cdot \Phi^{-1}(p_j) \tag{3.11}$$

em que $\Phi^{-1}(p_j) = z_j$ é o quantil p_j da distribuição normal padrão.

Para construir o gráfico de probabilidade normal, temos que

$$F(y) = \Phi\left(\frac{y - \mu}{\sigma}\right),$$

ou seja, a fda Φ da distribuição normal padrão faz o papel da fda G na Equação 5.1. A idéia central de um gráfico de probabilidade é comparar duas funções de distribuição. No caso do gráfico de probabilidade normal, ou simplesmente gráfico normal, o objetivo é verificar se os valores y_1, y_2, \dots, y_n , observações de uma variável aleatória Y , são provenientes da distribuição normal $\Phi(z)$ ⁶. Em outras palavras, deseja-se investigar se

⁵A função quantil assim definida provê um ponto de partida conveniente para encontrar a transformação necessária para linearizar a fda Φ ; a fim de que esta seja plotada como uma linha reta e não com sua forma padrão já conhecida (MEEKER; ESCOBAR, 1998).

⁶É importante destacar aqui que F e $\Phi(z)$ têm o mesmo suporte. O suporte é o conjunto $A(x)$ no qual a função assume seus valores. No caso da normal, $A(x) = \mathbb{R}$.

Y tem distribuição normal. Existem dois tipos principais de gráficos normais: o gráfico dos quantis e o gráfico dos percentis.

O gráfico dos quantis é usado mais frequentemente⁷ (STIRLING, 1982; COX, 1984), e como o próprio nome afirma, a comparação gráfica das duas distribuições é feita por se plotar os quantis de uma distribuição contra os quantis da outra. Sua construção é bastante simples. O procedimento é feito do seguinte modo: inicialmente, os valores y_j observados (considerados, nesse contexto gráfico, como os quantis da distribuição amostral F da Equação 5.1) são ordenados; colocados na ordem crescente⁸ de seus valores: $y_{(1)} \leq y_{(2)} \leq \dots \leq y_{(n)}$. Depois, é calculada a proporção acumulada da amostra, p_j , geralmente por uma daquelas formas discutidas na Seção 3.2; cujo estimador é \hat{p}_j . Assim, o gráfico normal é construído com os pontos $(\Phi^{-1}(\hat{p}_j), y_{(j)})$ ⁹, com $j = 1, 2, \dots, n$; do qual se conclui, a partir da Equação 3.11, que, se os dados forem aproximadamente normais, os pontos marcados aparecerão como aproximadamente colineares, e o gráfico dos quantis será também uma reta, com intercepto em μ e inclinação σ . Esta propriedade é importante pois μ e σ podem ser estimados pela simples visualização gráfica. Desvios da linearidade fornecem indícios de que as duas distribuições são diferentes; significando, no caso do gráfico normal, que a distribuição normal não é uma boa explicação para o comportamento dos dados observados.

Um papel gráfico especializado no qual os valores de Φ já encontram-se marcados para certos valores de \hat{p}_j é chamado de papel de probabilidade. Como Φ é a distribuição normal padrão, tal papel é chamado de papel de probabilidade normal. Mas vale salientar que, com o desenvolvimento dos pacotes estatísticos, o uso à mão de tais instrumentos torna-se desnecessário, pois os gráficos são gerados automaticamente pelos programas de computação utilizados. A Figura 3.1 mostra um destes papéis gráficos.

O *P-P Plot* é o gráfico dos percentis de uma distribuição contra os percentis da outra. Os pontos marcados são da forma $(\hat{p}_j; y_{(j)})$. A idéia na avaliação deste gráfico

⁷A versão R 2.10.0 apresenta o gráfico dos quantis como a forma padrão para os gráficos de probabilidade. Por conta disso, esta também é a forma do gráfico normal utilizada nesta dissertação.

⁸A ordenação das observações é necessária pois, como visto na Seção 3.2, estamos estimando F pela função de distribuição empírica da amostra $\hat{F}_n(y)$. Deste modo, procedendo do menor ao maior y , obtemos uma estimativa para a inteira fda F .

⁹Esta é a forma em que os pontos são marcados na versão R 2.10.0. Nas definições encontradas sobre o gráfico normal bem como em alguns outros pacotes estatísticos, os eixos do gráfico estão invertidos, isto é, os pontos marcados são da forma $(y_{(j)}, \Phi^{-1}(\hat{p}_j))$ (BLOM, 1989; MEEKER; ESCOBAR, 1998).

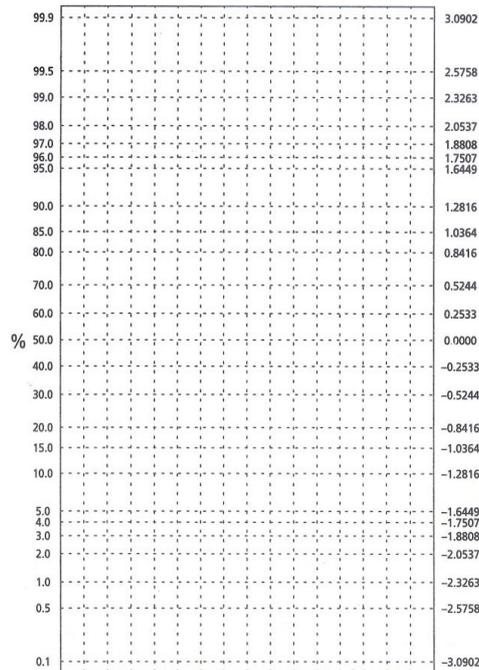


Figura 3.1: **Papel de Probabilidade Normal.** No eixo das ordenadas encontram-se as probabilidades acumuladas (à esquerda) e seus respectivos quantis (à direita). No eixo das abscissas são colocados os valores observados ordenados.

é a mesma que no gráfico dos quantis: se a amostra se “parece” com a distribuição teórica, os valores marcados terão uma tendência linear.

Alguns comentários são feitos sobre os dois tipos de gráficos normais descritos acima. Um artigo que aborda questões importantes acerca de tais gráficos é o de Wilk e Gnanadesikan (1968). Ao descrever o gráfico dos quantis, os autores comentam, por exemplo, que uma propriedade básica deste gráfico é o que eles chamam de invariância linear: se Y for uma função linear de X o gráfico dos quantis ainda será uma reta, com possíveis mudanças no intercepto e na inclinação (isso foi discutido na Seção 3.2 para o caso da família localização-escala). E é exatamente esta propriedade que, segundo os autores, tem feito do gráfico dos quantis uma ferramenta tão atraente e valiosa. Mas uma ressalva é feita: o gráfico dos quantis enfatiza mais a comparação nas caudas das distribuições; ao passo que sua análise da região central delas é obscura. Um razão apontada para isso é o fato de o quantil ser uma função que varia rapidamente nas caudas e muito lentamente no “meio” ou na região central das densidades. Os autores também descrevem algumas diferenças entre estas duas formas gráficas: o gráfico dos percentis não permanecerá linear no caso de diferenças de localização-escala; diferenças

no “interior” das distribuições são usualmente mais aparentes no gráfico dos percentis do que no gráfico dos quantis; o gráfico dos percentis pode ser usado nos casos multivariados, enquanto o gráfico dos quantis parece não ter nenhum uso nestas situações; o gráfico dos percentis também é usado para apresentar os resultados do poder de um teste estatístico. O artigo ainda observa um caso especial em que o gráfico dos quantis e o gráfico dos percentis serão idênticos: quando $X, Y \sim U(0, 1)$. E apesar de suas vantagens, “*techniques for using, generalizing, and interpreting the general P-P plot formulation require more investigation, including estimation procedures and robustness properties in preliminary data standardization methods.*” (WILK; GNANADESIKAN, 1968, p. 8).

De acordo com Stirling (1982), há um problema com a interpretação dos gráficos de probabilidade devido ao fato de que sucessivas observações ordenadas, $y_{(j)}$ e $y_{(j+1)}$, são altamente correlacionadas e, portanto, segundo ele, variações aleatórias nos dados resultarão numa curva aparentemente sistemática nos gráficos. Sendo assim, uma ajuda para a interpretação da linearidade dos gráficos de probabilidade seria construir, no próprio gráfico, intervalos de confiança simultâneos. O referido artigo dedica-se a descrever e exemplificar tal procedimento.

Gan e Koehler (1990) também citam algumas características peculiares desses dois tipos de gráficos de probabilidade normal: a maior razão para a popularidade do gráfico dos quantis é que ele não requer a estimação de parâmetros para avaliar a adequabilidade do modelo; transformações lineares nos dados originais provocarão mudanças no intercepto e na inclinação da reta ajustada no gráfico dos quantis, refletindo assim as mudanças na localização e na escala¹⁰, embora o coeficiente de correlação permaneça inalterado; o gráfico dos percentis possui a propriedade denominada *location and scale invariance*, denotando que tal gráfico permanece invariante com respeito a transformações lineares, ou seja, tais transformações não alterarão as posições dos pontos no gráfico, que sempre mostrará a mesma reta, com intercepto na origem e inclinação de 45°, independentemente do conjunto de dados, o que é útil para avaliar a qualidade de

¹⁰A esta propriedade os autores chamaram de “*location and scale equivariant*”, ou seja, os parâmetros de localização e escala mudam de acordo com as mudanças (transformações lineares) ocorridas nos dados, significando que, para cada conjunto de dados haverá uma reta correspondente. Tal comportamento mostra que os parâmetros de localização e escala são, no gráfico dos quantis, equivariantes, isto é, variam igual ou proporcionalmente às variações ocorridas nos dados.

diferentes modelos propostos. Os autores comentam ainda como interpretar algumas formas de desvios da linearidade no gráfico dos quantis: um padrão de S significa que a verdadeira distribuição tem caudas mais curtas ou mais achatadas que a distribuição normal hipotetizada; uma forma de S invertido significa caudas mais pesadas; uma curva convexa significa uma assimetria à direita com uma cauda direita pesada; uma curva côncava indica uma assimetria à esquerda com uma cauda direita curta.

Ainda nesse enfoque, Eric B. Holmgren (1995) contrastou o desempenho do gráfico dos quantis com o gráfico dos percentis em experimentos nos quais se comparava os efeitos de certas drogas em três diferentes situações. Em cada um dos três experimentos comentados, o efeito do tratamento era comparado com o grupo-controle, que recebia o placebo. As conclusões apresentadas foram as seguintes: (1) o gráfico dos percentis representa melhor diferentes condições experimentais que o gráfico dos quantis, pois, segundo Holmgren, este último varia muito com a variação das condições do experimento, tornando a análise um pouco mais difícil; (2) na presença de valores discrepantes, a avaliação feita pelo gráfico dos quantis é comprometida, enquanto que no gráfico dos percentis, a presença destes não obscurece a interpretação; (3) a informação do grupo-controle, necessária para a comparação dos tratamentos, não é incorporada no gráfico dos quantis.

Por ser um sensível indicador de discrepâncias na comparação entre duas distribuições, o gráfico normal tem utilidades muito além desta única tarefa. Como será mostrado no Capítulo 4, eles são usados também para verificar as suposições de modelos, na análise de resíduos, bem como em experimentos fatoriais. O presente trabalho se concentrará neste último aspecto.

O primeiro a propor o uso de gráficos de probabilidade para a análise de experimentos fatoriais não replicados foi Daniel (1959); cuja proposta inicial foi a utilização da distribuição semi-normal como distribuição de referência para o julgamento dos efeitos no gráfico, conforme será detalhado na Seção 3.4. Com o passar do tempo, porém, a distribuição normal torna-se preferida por alguns pesquisadores como a distribuição de referência a ser usada para a construção do gráfico de probabilidade no qual os efeitos serão julgados (DANIEL, 1976; BOX; MEYER, 1986)¹¹. A construção do gráfico normal no caso dos experimentos fatoriais usa os mesmos conceitos já discutidos nesta

¹¹Ver Seção 3.5.

Seção: Sejam $y_{(1)} \leq y_{(2)} \leq \dots \leq y_{(n)}$ os valores ordenados dos efeitos estimados dos fatores. O gráfico normal será o conjunto de pontos $(\Phi^{-1}(\hat{p}_j), y_{(j)})$, cujo padrão é uma reta. Um efeito é julgado significativo quando é marcado fora desta reta.

A razão por trás deste método é que os experimentos fatoriais são construídos sob a suposição de que os efeitos estimados são normalmente distribuídos, com média zero e variância constante (WU; HAMADA, 2000). O gráfico normal é, então, um teste para verificar se todos os efeitos estimados têm a mesma distribuição. Portanto, os efeitos marcados fora da reta, por não seguirem a mesma configuração dos demais pontos, são julgados como significativos. Efeitos significantes positivos são marcados acima da reta, ao passo que aqueles negativos são marcados abaixo desta.

Afim de ilustrar o uso do gráfico normal, considere o seguinte exemplo: **Exemplo 3.1. (DANIEL, 1959; 1976)** Trata-se de um fatorial 2^5 , não replicado, cujo interesse é estudar os efeitos sobre a produção resultante da variação nas concentrações de cinco componentes de um nutriente usado numa cultura de *Penicillium chrysogenum*. Os efeitos dos fatores (informados na referência) são os seguintes: $A = -190$, $B = -6$, $AB = -64$, $C = 153$, $AC = -53$, $BC = 53$, $ABC = 0$, $D = 9$, $AD = -54$, $BD = -7$, $ABD = -34$, $CD = -4$, $ACD = 33$, $BCD = -18$, $ABCD = 58$, $E = -224$, $AE = 2$, $BE = 29$, $ABE = -22$, $CE = -93$, $ACE = -58$, $BCE = 39$, $ABCE = 31$, $DE = 30$, $ADE = 21$, $BDE = 28$, $ABDE = 14$, $CDE = 12$, $ACDE = 47$, $BCDE = 16$, $ABCDE = 77$. O gráfico normal para este exemplo é mostrado na Figura 3.2. Segundo Daniel (1959, 1976), os efeitos claramente julgados significativos são A , C e E ; e **talvez** CE e $ABCDE$ tenham alguma significância. Estas conclusões podem ser notadas no gráfico da Figura 3.2, que também destaca o caráter subjetivo da análise gráfica: o que para Daniel (1959) é **provavelmente**¹² ativo (os efeitos CE e $ABCDE$), talvez não seja considerado como tal por outro pesquisador, uma vez que a distância destes efeitos à reta não é tão destacadamente grande como no caso dos efeitos A , C e E .

¹²Ele usa a expressão “*judged likely to be appreciable*”.

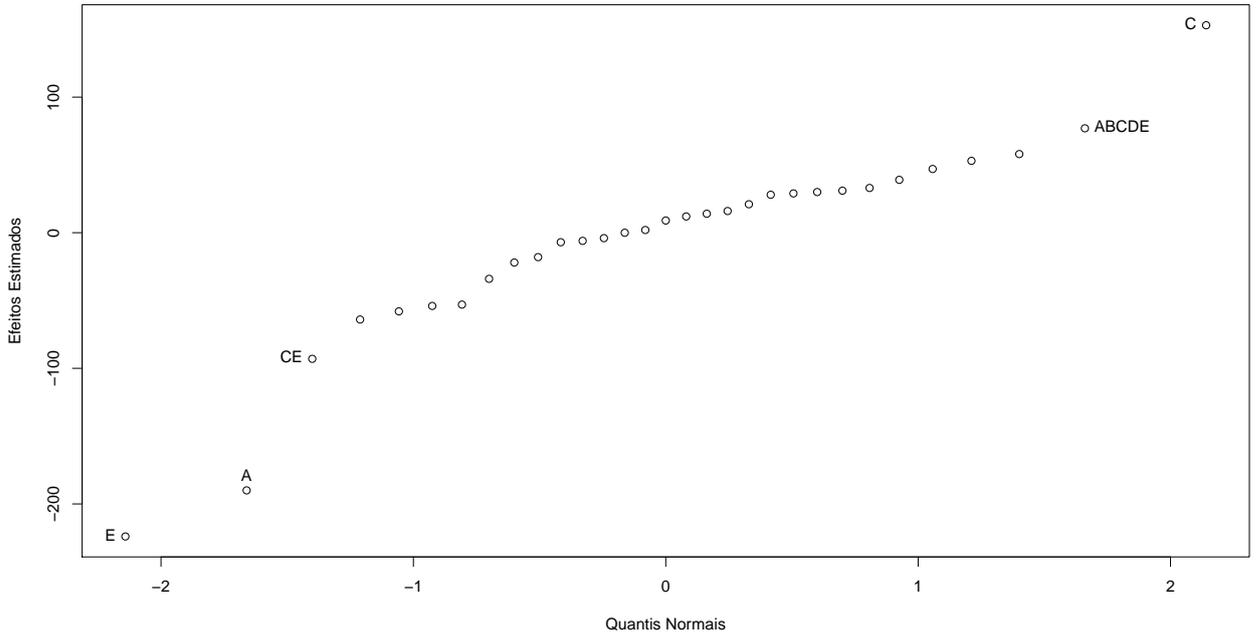


Figura 3.2: Gráfico Normal para o Exemplo 3.1.

3.4 Gráfico de Probabilidade Semi-Normal

Em novembro de 1959, Cuthbert Daniel publica um artigo intitulado *Use of Half-Normal Plots in Interpreting Factorial Two-Level Experiments*, que torna-se a primeira proposta de um método gráfico para a análise de experimentos fatoriais 2^k e 2^{k-p} não replicados: o gráfico semi-normal (AL-SHIHA; YANG, 1999). O principal enfoque do artigo é mostrar como a nova técnica pode ser utilizada para estimar o erro padrão e fazer julgamento sobre a significância dos efeitos dos fatores. Esta seção abordará a construção e os usos desse método.

A distribuição semi-normal, também chamada de “*folded normal distribution*” pelo fato de seu gráfico (ver Figura 3.3) parecer com o gráfico de uma normal “dobrado” ao meio (ELANDT, 1961), é a distribuição do valor absoluto da uma variável aleatória normal padrão. Sua função de distribuição acumulada é dada por:

$$\Phi_o(z) = 2\Phi(z) - 1,$$

em que $z > 0$ e $\Phi(z)$ é a fda da normal padrão.

Assim, se $X \sim N(0, \sigma^2)$, uma variável aleatória $X' = |X|$ tem distribuição semi-normal e sua fdp é dada por:

$$f_{X'}(x') = \begin{cases} \frac{1}{\sigma} \sqrt{\frac{2}{\pi}} e^{-\frac{x'^2}{2\sigma^2}} & , \quad \text{se } x' \geq 0; \\ 0 & , \quad \text{se } x' < 0. \end{cases}$$

com média $\sqrt{(2\sigma^2/\pi)}$ e desvio padrão σ (ELANDT, 1961; TAYLOR, 1994). Seu gráfico pode ser visto na Figura 3.3.

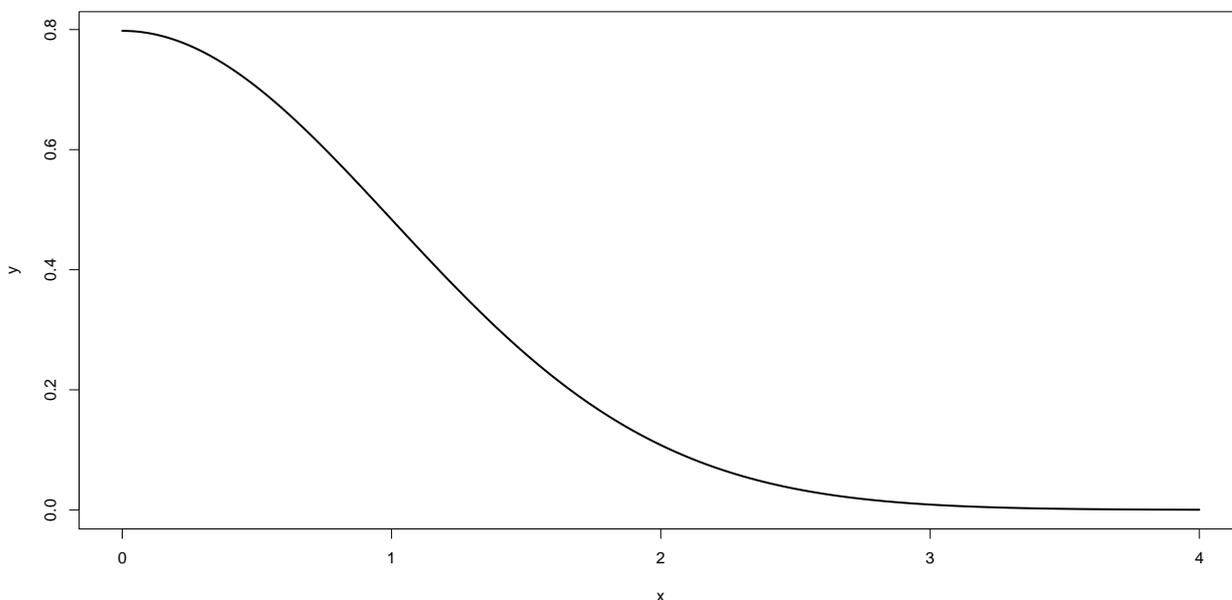


Figura 3.3: **Gráfico da Densidade da Função Semi-Normal.** A forma gráfica fornece uma idéia do motivo desta distribuição ser chamada de “**semi**”(half)-normal.

Um gráfico semi-normal em planos fatoriais é um gráfico dos valores absolutos ordenados dos efeitos estimados, $|y_1| < |y_2| < \dots < |y_n|$, contra os quantis correspondentes da distribuição semi-normal, ou seja, é o conjunto de pontos da forma $(\Phi_\sigma^{-1}(\hat{p}_j), |y_{(j)}|)$. Segundo Daniel (1959), cortando o papel de probabilidade normal pela metade, isto é, na linha 50%, e reescalando o eixo das porcentagens restantes ($P \geq 50\%$) usando a transformação $P' = 2P - 100$, obtem-se o papel de probabilidade semi-normal, que pode ser visto na Figura 3.4. A outra metade também pode ser usada, desde que se faça a seguinte substituição: $P' = 100 - 2P$.

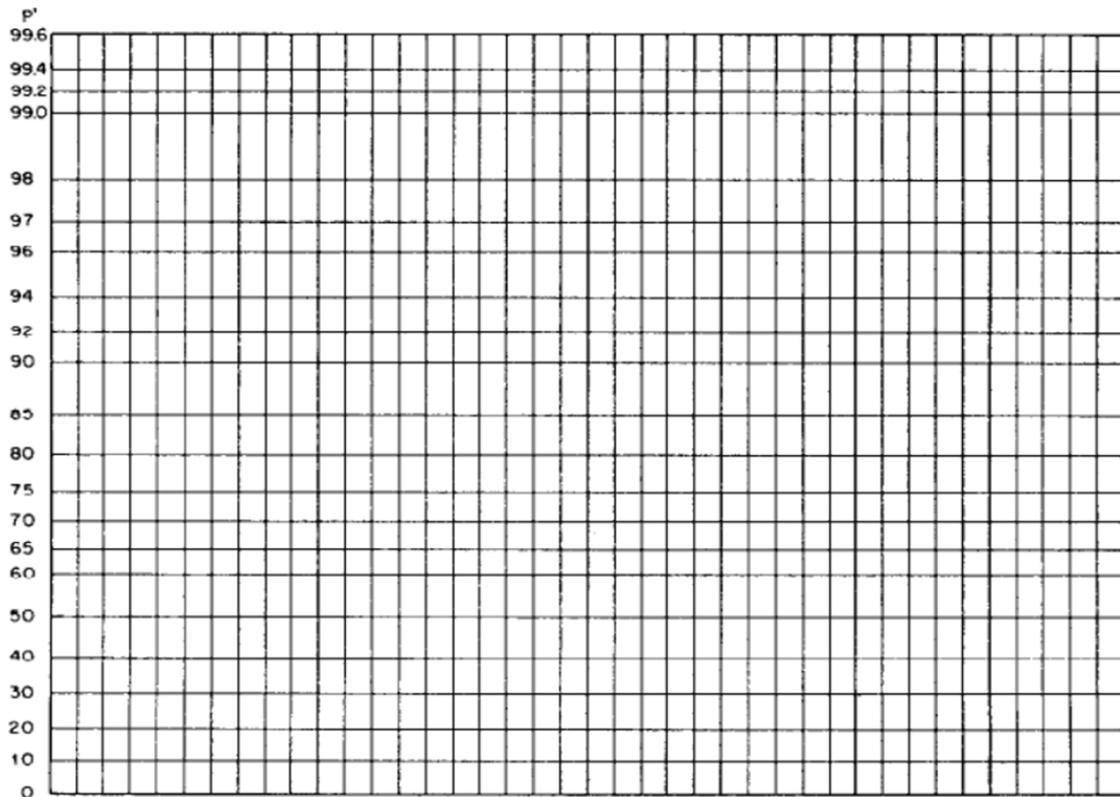


Figura 3.4: **Papel de Probabilidade Semi-Normal.** (DANIEL, 1959)

Desta forma, 50% torna-se 0%, 60% torna-se 20%, e assim por diante. Daí, os valores de P' são calculados por uma daquelas formas do \hat{p}_j discutidas na Seção 3.2 e marcados no gráfico contra os valores absolutos dos efeitos¹³. Se nenhum dos efeitos for ativo, espera-se que todos os pontos marcados sejam aproximadamente colineares, sobre os quais se pode traçar uma reta partindo da origem e com inclinação σ , que será a estimativa do erro experimental. Quando há efeitos significativos no experimento, estes são identificados pelos pontos do gráfico que apareceram fora da reta traçada pela maioria dos outros pontos. Quanto mais distante um ponto estiver dessa reta, maiores serão as evidências a favor da significância estatística do efeito que este representa. No gráfico semi-normal, todos os grandes efeitos estimados aparecem no canto superior direito da reta, numa posição acima desta (WU; HAMADA, 2000).

Ao traçar o gráfico semi-normal, os efeitos “inativos” serão aqueles com média zero (daí porque são chamados muitas vezes de efeitos nulos) e variância constante, e portanto, seria de esperar que estes “caíssem” sobre uma reta com intercepto na

¹³Em sua proposta inicial, Daniel (1959) usa a relação $P' = (j - 1/2)/n$, com $j = 1, 2, \dots, n$.

origem, e cuja inclinação representasse essa variância comum a todos eles. Um ponto marcado fora desta reta indicaria que tal efeito tem uma distribuição diferente da dos demais. Se algumas combinações de efeitos contribuem diferentemente dos demais, eles se encontram fora desta reta comum. Por isso, os efeitos que aparecem com essa configuração no gráfico são considerados estatisticamente significativos ou “não nulos”. O Teorema Central do Limite garante a convergência da distribuição semi-normal para a distribuição normal, satisfazendo assim as suposições do experimento (ELANDT, 1961; LEONE; NELSON; NOTTINGHAM, 1961). A opção pela distribuição semi-normal deve-se ao fato de Daniel ter utilizado os valores absolutos dos efeitos e não os efeitos assinalados (com sinais “-” e “+”), pois, segundo ele (DANIEL, 1956), os sinais dos efeitos são o resultado de definições arbitrárias, apenas uma notação usual.

Além de identificar os efeitos significativos, o gráfico semi-normal também pode ser usado para estimar o erro experimental. Duas maneiras podem ser destacadas da proposta original de Daniel. A primeira delas é utilizar a inclinação da reta do gráfico semi-normal como uma estimativa de σ (uma justificativa para tal raciocínio pode ser encontrada na Equação 3.4) considerando duas situações: (1) se nenhum efeito for julgado significativo, então estima-se σ diretamente pela inclinação da reta; (2) se alguns efeitos forem significantes então, estes são retirados, o gráfico é reconstruído com os efeitos restantes e σ é estimado pela inclinação da reta deste novo gráfico. Uma segunda maneira de estimar σ é usar o j -ésimo efeito para o qual o valor de $P' = (j - 1/2)/n$ (esta foi a escolha de Daniel) é o mais próximo de 0,683 ou 68,3%, considerando também as duas situações citadas anteriormente: (1) se nenhum efeito for julgado significativo fica valendo $P' = (j - 1/2)/n$; (2) se k efeitos forem julgados significantes então usa-se $P' = (j - 1/2)/(n - k)$. O gráfico semi-normal também pode ser útil em detectar (DANIEL, 1959): a presença de um ou mais valores discrepantes nos dados; restrições na aleatorização não identificadas pelo pesquisador, tais como os planos com parcelas subdivididas; e heterocedasticidade.

Daniel conclui seu artigo destacando que sua técnica possui muitas limitações, que esta não é um substituto geral para a análise de variância, e que deve-se ter cuidado com seu uso indiscriminado, pois ela não é infalível. Ele chama a atenção também para o fato de que uma linha suave nem sempre significa que nenhum fator seja significativo; da mesma forma, pontos fora da reta podem levar à conclusão de

que tais efeitos sejam reais, quando, na verdade, não são. Enfim, “*if the experimental conditions were precisely followed, if there are no defective values, if the error variance is nearly constant over all experimental conditions, and if there are only a few real effects (say fewer than $n/5$)¹⁴, then a half-normal plot often looks very nice and can be used to make decisions about effects with roughly known experiment-wise probability of type I error.*” (DANIEL, 1967, pg. 3005) Estes aspectos fazem do gráfico semi-normal uma importante ferramenta estatística de grande utilização nos experimentos fatoriais não replicados (TAYLOR, 1994; OLGUIN; FEARN, 1997; LAWSON; GRIMSHAW; BURT, 1998; VARIYATH; ABRAHAM; CHEN, 2005).

Exemplo 3.2. Considere os dados do Exemplo 3.1. O gráfico semi-normal dos efeitos é mostrado na Figura 3.5. As interpretações feitas por ocasião do Exemplo 3.1 aplicam-se igualmente no exemplo atual.

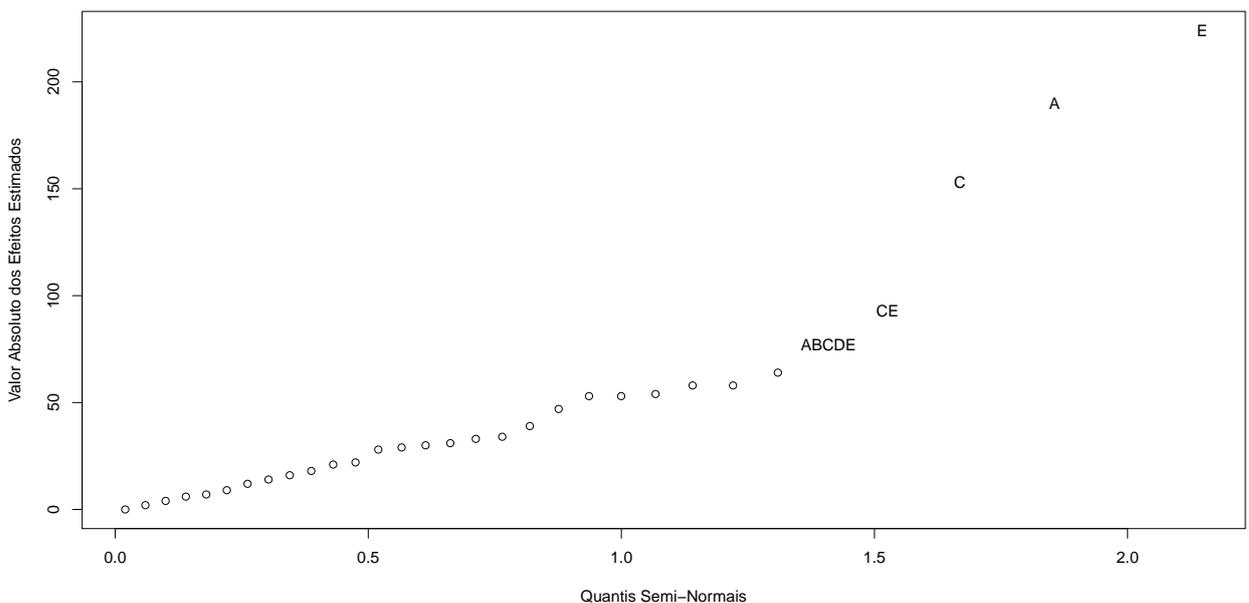


Figura 3.5: Gráfico Semi-Normal para o Exemplo 3.2.

¹⁴O método gráfico de Daniel é construído sob a suposição da esparsidade dos efeitos: apenas uma pequena proporção dos efeitos é significativamente diferente de zero (BOX; MEYER, 1986).

Novamente, a dúvida acerca da significância dos efeitos CE e $ABCDE$ também está presente no gráfico semi-normal, destacando a subjetividade desta técnica. Como discutido no Capítulo 2, diante de tais incertezas, o uso de um método formal pode ser de ajuda nas interpretações. Sendo assim, continuando a análise do Exemplo 3.2, aplicar-se-á agora o método das faixas de segurança, proposto pelo próprio Daniel.

Método de Daniel (1959) das faixas de segurança, aplicado aos dados do Exemplo 3.2. Inicialmente calcula-se os valores absolutos ordenados dos efeitos estimados, que pode ser visto na Tabela 3.1. Daí, para $n = 31$ temos que $0,683(n+1) = 0,683(32) = 21,856$; o que faz de $|y_{22}| = 53$ a estimativa inicial de σ que será usada na determinação das faixas. Em seguida, determina-se as coordenadas das faixas de acordo com os valores críticos de Zahn (1975a), que, para o caso de $n = 31$, sugere usar para tal cálculo apenas os efeitos $|y_{27}|$, $|y_{28}|$, $|y_{29}|$, $|y_{30}|$ e $|y_{31}|$. Os resultados encontram-se na Tabela 3.2.

Tabela 3.1: Valores absolutos ordenados dos efeitos estimados do Exemplo 3.2.

Ordem	Efeito	Valor Absoluto
1	<i>ABC</i>	0
2	<i>AE</i>	2
3	<i>CD</i>	4
4	<i>B</i>	6
5	<i>BD</i>	7
6	<i>D</i>	9
7	<i>CDE</i>	12
8	<i>ABDE</i>	14
9	<i>BCDE</i>	16
10	<i>BCD</i>	18
11	<i>ADE</i>	21
12	<i>ABE</i>	22
13	<i>BDE</i>	28
14	<i>BE</i>	29
15	<i>DE</i>	30
16	<i>ABCE</i>	31
17	<i>ACD</i>	33
18	<i>ABD</i>	34
19	<i>BCE</i>	39
20	<i>ACDE</i>	47
21	<i>BC</i>	53
22	<i>AC</i>	53
23	<i>AD</i>	54
24	<i>ACE</i>	58
25	<i>ABCD</i>	58
26	<i>AB</i>	64
27	<i>ABCDE</i>	77
28	<i>CE</i>	93
29	<i>C</i>	153
30	<i>A</i>	190
31	<i>E</i>	224

Tabela 3.2: Cálculo das coordenadas das faixas de confiança para o Exemplo 3.2.

Ordem	$\alpha = 0,05$	$\alpha = 0,20$	$\alpha = 0,40$
31	$3,351 \cdot 53 = 177,6$	$2,730 \cdot 53 = 144,7$	$2,372 \cdot 53 = 125,7$
30	$3,173 \cdot 53 = 168,2$	$2,586 \cdot 53 = 137,1$	$2,247 \cdot 53 = 119,1$
29	$2,990 \cdot 53 = 158,5$	$2,439 \cdot 53 = 129,3$	$2,123 \cdot 53 = 112,5$
28	$2,807 \cdot 53 = 148,8$	$2,288 \cdot 53 = 121,3$	$1,992 \cdot 53 = 105,6$
27	$2,615 \cdot 53 = 138,6$	$2,160 \cdot 53 = 114,5$	$1,857 \cdot 53 = 98,4$

O próximo passo agora é construir o gráfico semi-normal com as faixas de confiança marcadas, mostrado na Figura 3.6. Como se pode ver pelo referido gráfico, segundo o Método de Daniel das faixas de segurança, os efeitos A , C e E são julgados significativos, uma vez que foram marcados acima de alguma das faixas calculadas. Os efeitos CE e $ABCDE$ são, por este método, claramente insignificantes.

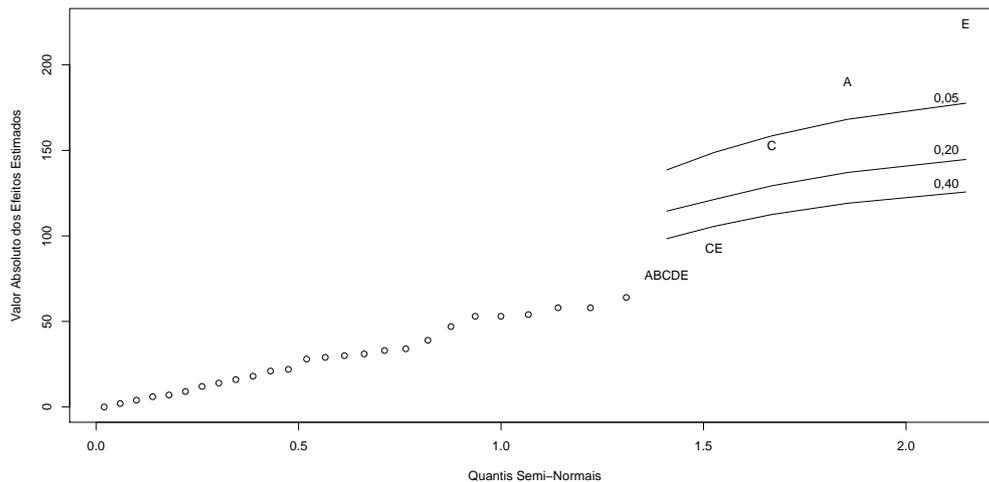


Figura 3.6: Gráfico Semi-Normal com as faixas de confiança para os dados do Exemplo 3.2.

3.5 A Controvérsia

Apesar de ter sido o idealizador e grande apoiador do gráfico semi-normal para a análise de experimentos fatoriais não replicados, Cuthber Daniel acaba mudando de idéia. Com a publicação de seu livro *Applications of Statistics to Industrial Experimentation* em 1976, Daniel deixou clara sua preferência pelo gráfico normal: todos os gráficos utilizados para as análises dos fatoriais não replicados são gráficos normais. Daniel explica o motivo de sua nova escolha: *“I have of course made half-normal plots [Daniel, 1959] of the contrast-sums from each of the 2^5 experiments discussed in this chapter. To my great chagrin none of the peculiarities discovered above is reflected in these plots, nor have any other notable irregularities been found. The reasons are not far to seek. The defects found are all strongly sign dependent, and all are properties of subsets of the data set which are obscured in the half-normal plots by overaggregation.*

The signed contrasts in standard order have more information in them than do the unsigned contrasts ordered by magnitude.” (DANIEL, 1976, página 149.)

As características a que Daniel se refere são, principalmente, a presença de valores discrepantes nos dados e sua influência no cálculo dos efeitos; a avaliação da significância dos efeitos quando estes forem pequenos ou em número considerável; e também a avaliação da violação da suposição de homocedasticidade. Daniel também percebeu que o gráfico semi-normal era insensível a tais características no sentido de esconder tais situações. Realmente uma grande mudança se comparada com suas palavras citadas anteriormente neste capítulo, de que o gráfico semi-normal era sensível e útil em detectar exatamente estas situações, e que o uso dos sinais era uma questão meramente arbitrária.

Esta mudança deu início a uma controvérsia sobre o uso dos gráficos de probabilidade normal na análise dos planos fatoriais não replicados. Ao passo que alguns pesquisadores optam pelo gráfico semi-normal, outros utilizam o gráfico normal. Um considerável aumento pela preferência do gráfico normal deve-se ao artigo de Box e Meyer (1986) intitulado *An Analysis for Unreplicated Fractional Factorials*. Ali os autores utilizam este gráfico com os mesmos objetivos do gráfico semi-normal: avaliar a significância dos efeitos e estimar o erro em experimentos fatoriais não replicados. Este artigo ainda é citado hoje como referência e justificativa para o uso do gráfico normal nos experimentos fatoriais, como pode ser visto no Capítulo 4 desta dissertação. Apesar disso, como se verá mais adiante, o gráfico semi-normal ainda é a ferramenta mais utilizada para a avaliação da significância dos efeitos nos arranjos fatoriais não replicados.

A mudança de opinião de Daniel teve outros alcances também. De acordo com Mark J. Anderson¹⁵, no manuscrito original do livro *Statistics for Experimenters - Design, Innovation, and Discovery*, considerado um dos livros de referência na área de planejamento de experimentos, os autores G. E. P. Box, J. Stuart Hunter e William G. Hunter, haviam escolhido o gráfico semi-normal como ferramenta padrão para a análise dos fatoriais não replicados; mas que, depois de comentários de Daniel a favor do gráfico normal, os autores optaram por usar exclusivamente esta forma gráfica na

¹⁵Ver www.statease.com/news/faqalert6-09.html

versão impressa. Ao ser indagado sobre esta afirmação¹⁶, J. Stuart Hunter confirmou sua veracidade, embora lamentasse o fato de não lembrar as razões de Daniel para esta mudança.

Como se já não bastasse a controvérsia criada pela mudança de opinião Daniel em 1976; Mark J. Anderson, acima mencionado, afirma algo realmente intrigante: Daniel “*reversed himself back*” e veio a favorecer novamente o gráfico semi-normal! O Professor Hunter disse não recordar do julgamento final de Daniel sobre o gráfico semi-normal. Perguntado sobre qual sua preferência pessoal, ele respondeu que, se tivesse que escolher, escolheria o gráfico normal; mas deixou claro que isso não significa que ele sempre evitaria ou nunca usaria o gráfico semi-normal.

Embora não se tenha encontrado outra referência que comprovasse, ou não, esta segunda mudança de opinião de Daniel, pretende-se destacar aqui a controvérsia sobre a escolha entre o gráfico normal e o gráfico semi-normal. Como se verá no Capítulo 4, além da subjetividade própria das técnicas gráficas, o uso de tais ferramentas também é subjetivo na maioria das vezes; dependendo apenas da preferência daquele que a usa. Daí surgiu a motivação deste trabalho: contribuir com critérios mais objetivos, se possível, para ajudar na tomada de decisão sobre qual dos dois métodos gráficos escolher. Ora, o fato de o próprio inventor do gráfico semi-normal ter mudado de idéia algumas vezes sobre o desempenho de seu método, por si só é um bom motivo para se analisar mais profundamente essa questão.

Este capítulo é encerrado com a resposta final do Professor Hunter às indagações a ele feitas, parte da qual encontra-se na epígrafe desta dissertação. “*This puts me in mind of one of George Box’s famous quotes . . . “All models are wrong, some are more useful than others.” A paraphrase might be . . . “All plots are useful, some are less useful than others.” Cordially, Stu Hunter*”. O objetivo desta pesquisa é justamente mostrar, em que situações (se houver alguma), qual gráfico é “mais útil que o outro”.

¹⁶A Professora Orientadora deste trabalho, Carla Almeida Vivacqua, entrou em contato com o Professor J. Stuart Hunter via e-mail, em 25/01/2009.

Capítulo 4

Levantamento Bibliográfico

Este capítulo apresenta uma síntese do levantamento bibliográfico feito entre os artigos publicados na revista *Technometrics*, bem como em outros periódicos. Por meio deste levantamento pretende-se construir o referencial teórico deste trabalho, como também traçar o foco da pesquisa.

4.1 Levantamento Bibliográfico na *Technometrics*

Lançada em fevereiro de 1959, sob a liderança de J. Stuart Hunter como editor, a *Technometrics* estabeleceu-se, no último meio século, como uma das principais revistas especializadas na pesquisa estatística relacionada com as ciências físicas e engenharias, como também um veículo importante para apresentar métodos e estudos de caso pertinentes ao uso da estatística na indústria e tecnologia (STEINBERG; BISGAARD, 2008). Esta foi uma das razões que motivaram a escolha da *Technometrics* como fonte inicial da pesquisa bibliográfica retratada neste capítulo. Outro fator decisivo para tal escolha foi o fato de Cuthbert Daniel ter publicado seu artigo sobre o gráfico semi-normal neste periódico, no quarto número do primeiro volume, isto é, em novembro de 1959. Visto ser este o artigo base desta pesquisa, e pelo fato da técnica desenvolvida por Daniel ter completado seus cinquenta anos em 2009, deseja-se fazer uma busca dentro da *Technometrics* a fim de verificar e justificar, se possível, o seu uso (bem como o do gráfico normal) na pesquisa estatística. Não foi pretensão traçar um “estado da arte” sobre o assunto, mas fornecer uma referência sobre o uso dos gráficos de probabilidade normal do ponto de vista dos artigos publicados nestes cinquenta anos da *Technometrics*.

Ao todo, foram pesquisados 2162 artigos, publicados durante o período de fevereiro de 1959 a agosto de 2008. Para cada um destes, a metodologia de pesquisa é a seguinte: inicialmente verifica-se se o artigo trata de algum experimento ou não; para isto, utiliza-se as seguintes palavras de busca: *experiment, design, plan*. Em caso afirmativo, verifica-se qual o plano experimental utilizado. Por fim, em todos os artigos, busca-se informações sobre o tipo gráfico utilizado: gráfico normal e/ou gráfico semi-normal, ou nenhum dos dois. Como os artigos pesquisados estão todos em formato *PDF (Portable Document Format)*, as palavras de busca para cada item, conforme se pode observar nas Tabelas 4.1, 4.2 e 4.3 abaixo, são digitadas no campo de pesquisa do documento, e os resultados anotados numa planilha eletrônica, na qual contém, dentre outras coisas, a referência do artigo e os códigos correspondentes aos resultados da pesquisa. Vale salientar que as categorias consideradas no processo de busca não são mutuamente exclusivas, ou seja, é possível que mais de uma delas seja aplicável a um mesmo artigo. O modelo de busca utilizado para as pesquisas é o seguinte:

Tabela 4.1: Tipo de artigo publicado.

CÓDIGO	EXPERIMENTO	BUSCA
1	sim	<i>experiment, design, plan</i>
0	não	
Fora do escopo	<i>Book Reviews</i>	

Tabela 4.2: Plano experimental utilizado.

CÓDIGO	MÉTODO UTILIZADO	BUSCA
1	Fatorial completo	<i>full/factorial</i>
2	Fatorial fracionado	<i>fractional/factorial</i>
3	Taguchi	<i>taguchi</i>
4	Mistura	<i>mixture</i>
5	Parcelas subdivididas e variantes	<i>split/strip/block/lot/plot</i>
6	Experimento ótimo	<i>optimal</i>
7	Superfície de resposta	<i>response/surface</i>

Tabela 4.3: Tipo de gráfico utilizado.

CÓDIGO	TIPO GRÁFICO	BUSCA
0	Nenhum	
1	Normal	<i>normal probability plot/normal/full-normal/daniel</i>
2	Semi-Normal	<i>half-normal/daniel</i>
3	Ambos	<i>normal/full-normal/half-normal/daniel</i>

De todos os artigos pesquisados, 514 (24%) trataram de alguma forma de experimento, ao passo que 1648 (76%) não. Notou-se ainda que apenas 109 (5%) artigos usaram ou comentaram sobre os gráficos de probabilidade (normal e/ou semi-normal) e, destes, 40 (37%) não eram aplicações de planos experimentais (ou seja, estão no grupo dos 1648 já citados).

Quanto ao plano experimental utilizado, a Figura 4.1 mostra que, dentre os 514 artigos que trataram de alguma forma de experimento, 220 (43%) dedicaram-se aos experimentos fatoriais, com a seguinte configuração mostrada na Tabela 4.4:

Tabela 4.4: Artigos que se dedicaram aos experimentos fatoriais.

Plano fatorial utilizado	Usos
Fatorial Completo	110(50%)
Fatorial Fracionado	77(35%)
Ambos	33(15%)
Total	220

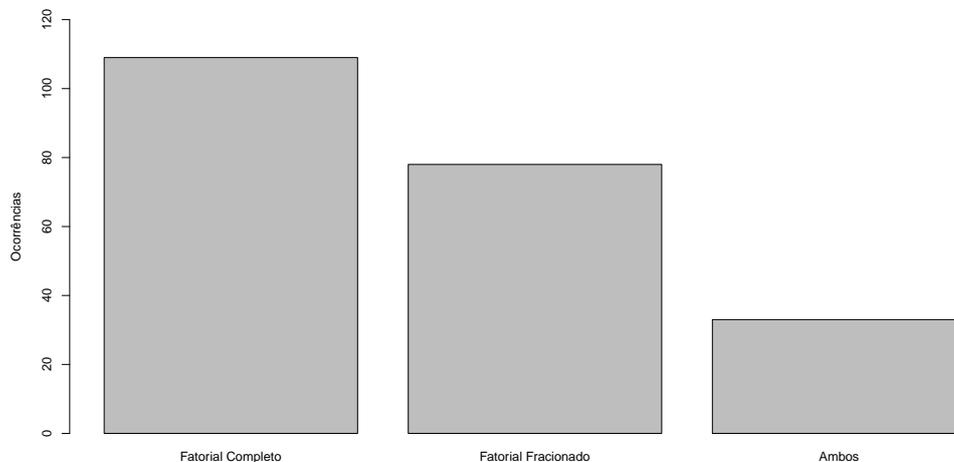


Figura 4.1: Planos Fatoriais Utilizados.

Considerando os 109 artigos que trataram de gráficos de probabilidade normal, constataram-se os seguintes resultados, mostrados nas Tabelas 4.5, 4.6, 4.7; e ilustrados nas Figuras 4.2, 4.3, 4.4 e 4.5¹.

¹Por uma questão de estética textual e aproveitamento de espaço nas páginas, algumas expressões foram abreviadas ou resumidas nos eixos dos gráficos. Explicando as abreviaturas: Ef. = avaliação da significância dos efeitos; Val. Discrep. = detectar valores discrepantes; Normal. = verificar a suposição de normalidade dos dados; Erro = estimar o erro experimental.

Tabela 4.5: Artigos que trataram de gráficos de probabilidade.

Tipo de gráfico	Usos
Gráfico Normal	53(49%)
Gráfico Semi-Normal	48(44%)
Ambos	8(7%)
Total	109

Tabela 4.6: Artigos que apenas citam ou recomendam as técnicas gráficas.

Tipo de gráfico	Efeitos	<i>Outliers</i>	Normalidade	Erro	Total
Gráfico Normal	4	1	9	0	14
Gráfico Semi-Normal	12	3	0	4	19
Ambos	5	0	0	0	5
Total	21	4	9	4	38

Tabela 4.7: Artigos que usam as técnicas gráficas.

Tipo de gráfico	Efeitos	<i>Outliers</i>	Normalidade	Efeitos e <i>outliers</i>	Total
Gráfico Normal	10	2	25	2	39
Gráfico Semi-Normal	23	4	0	2	29
Ambos	1	1	0	1	3
Total	34	7	25	5	71

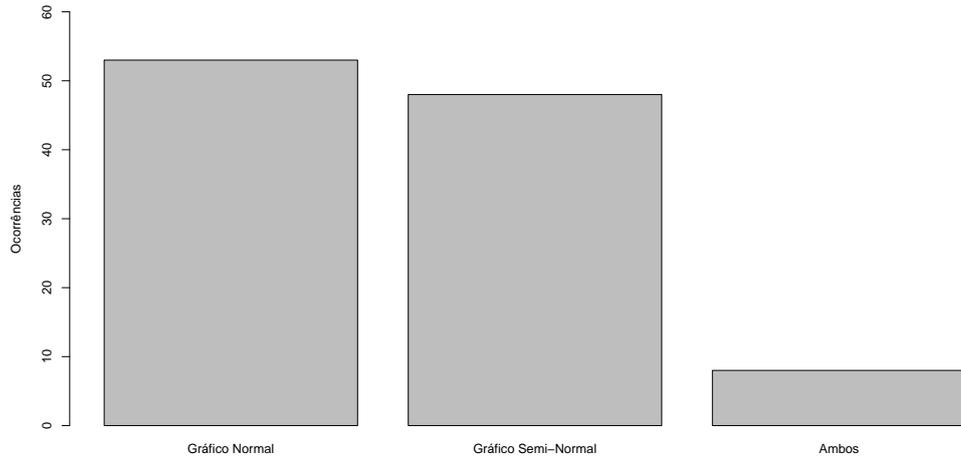


Figura 4.2: Tipo de Gráfico Utilizado.

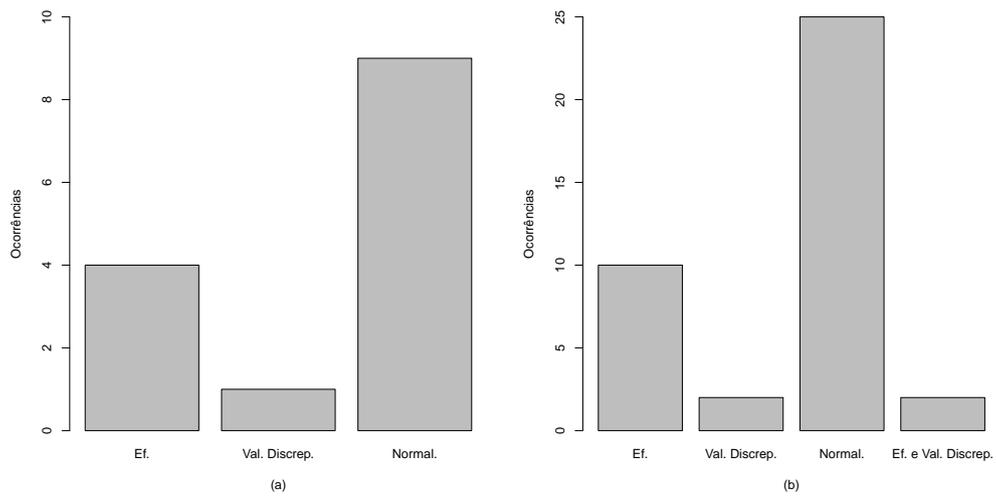


Figura 4.3: Abordagem do Gráfico Normal. (a) Artigos que apenas citam o gráfico normal. (b) Artigos que usam o gráfico normal.

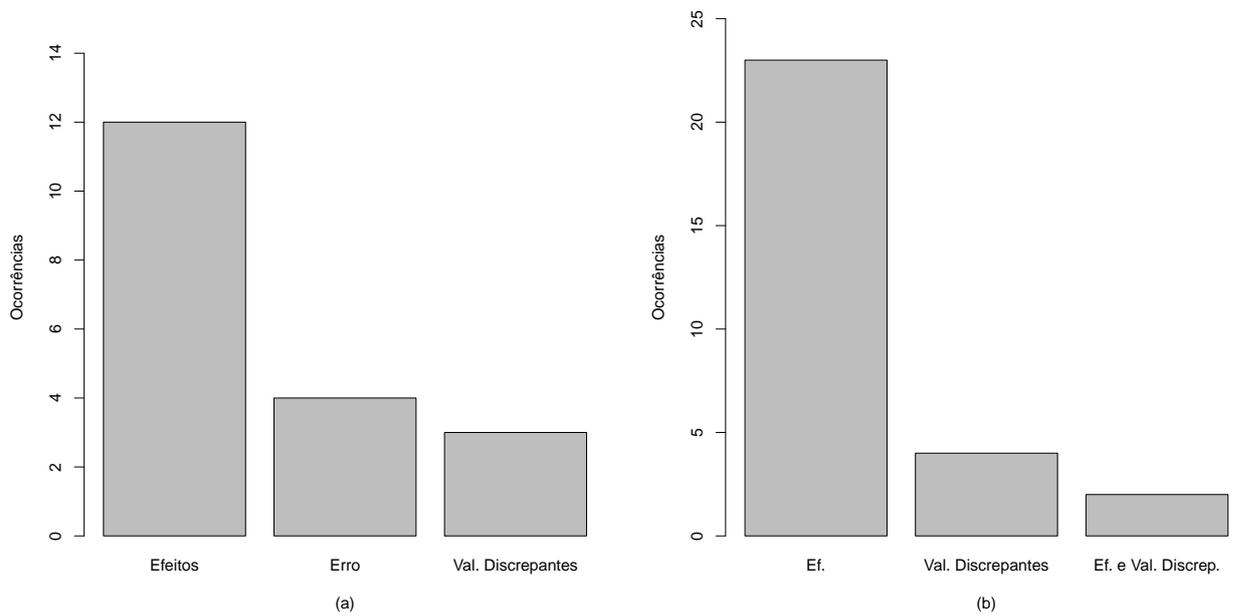


Figura 4.4: **Abordagem do Gráfico Semi-Normal.** (a) Artigos que apenas citam o gráfico semi-normal. (b) Artigos que usam o gráfico semi-normal.

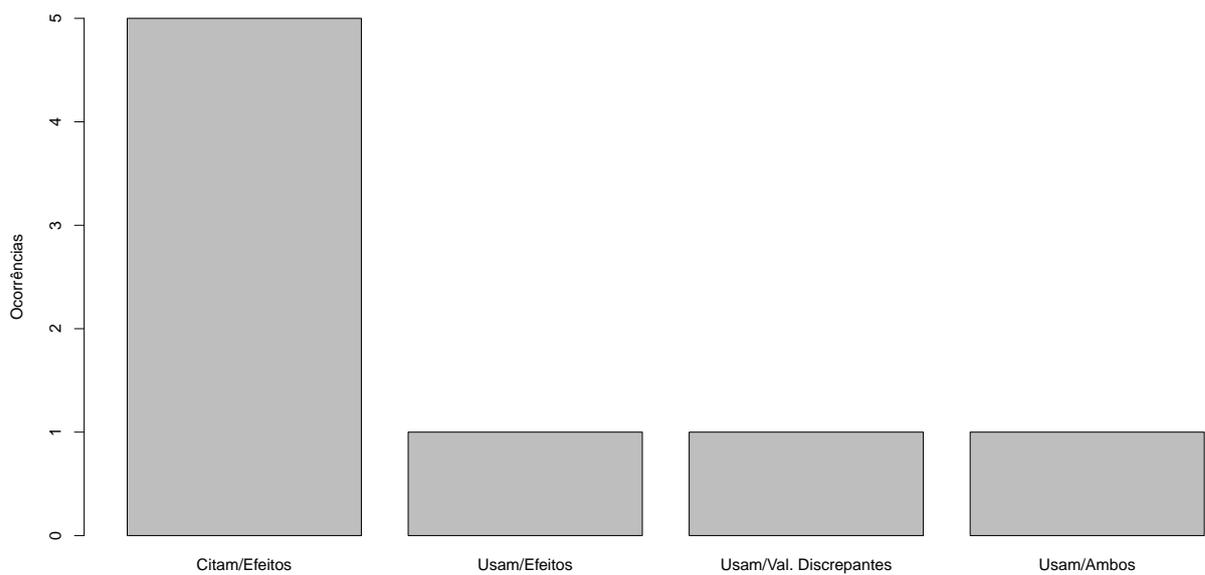


Figura 4.5: **Citações e Usos dos Dois Métodos Gráficos.**

4.1.1 Artigos Seleccionados da *Technometrics*

Como observado na seção anterior, apenas 109 artigos pesquisados tratam de algum tipo de gráfico de probabilidade normal. Alguns aspectos importantes destes artigos seleccionados são comentados nesta seção, pois tratam de conhecimentos indispensáveis à construção do referencial teórico deste trabalho. Destaca-se a seguir alguns deles, comentados em ordem cronológica. Os artigos são escolhidos tendo em vista sua contribuição para os objetivos do presente trabalho.

O primeiro artigo publicado na *Technometrics* a tratar de experimentos fatoriais e gráfico semi-normal é o *Partial Duplication of Factorial Experiments* (DYKSTRA, 1959). Dykstra comenta que um modo costumeiro de determinar a estimativa do erro em experimentos fatoriais não replicados é por utilizar as interações de alta ordem, consideradas sem efeito significativo. Daí aponta também outra maneira: utilizar o gráfico semi-normal, proposto por Daniel. É digno de nota que este artigo comenta o uso do gráfico semi-normal antes mesmo desta técnica ter sido publicada por Daniel na *Technometrics*. Dykstra faz alusão ao artigo *Fractional replication in industrial research* publicado em *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 1956, de autoria de Cuthbert Daniel. Neste artigo, Daniel faz sua primeira referência ao gráfico semi-normal. Num dado momento do texto, ao comentar sobre a avaliação da significância dos efeitos, ele sugere que um método apropriado seria traçar um gráfico dos valores absolutos destes efeitos contra os correspondentes quantis da distribuição semi-normal; o que ele chamou de gráfico semi-normal. Uma justificativa para a escolha dos valores absolutos é que, segundo Daniel, uma vez que os sinais dos efeitos são o resultado de definições arbitrárias dos dois níveis de cada fator, os valores absolutos da coleção completa dos contrastes contêm todas as informações disponíveis sobre o desvio padrão da distribuição. Assim, uma vez construído o gráfico, desvios da reta indicam que um fator pode ser julgado significativo. Daí, a estimativa do erro experimental será a inclinação da reta construída apenas pelos efeitos julgados não significativos. Nenhum outro detalhe acerca de sua construção, definição e propriedades é dado neste artigo. Trata-se apenas de uma alusão de Daniel à técnica que somente três anos mais tarde seria formalizada.

Como não poderia deixar de ser, destaca-se também aquele que serviu de base

para todo este trabalho: o artigo de Cuthbert Daniel, de novembro de 1959, intitulado *Use of Half-Normal Plots in Interpreting Factorial Two-Level Experiments*; no qual ele desenvolve uma nova ferramenta estatística para interpretar os experimentos fatoriais 2^k , completos e fracionados, não replicados: o gráfico semi-normal. Segundo Daniel, sua técnica é útil para avaliar a significância dos efeitos, fornecer uma estimativa para o erro experimental, detectar valores discrepantes nos dados, bem como identificar restrições na aleatorização. Além de apresentar uma técnica gráfica, o artigo sugere um teste formal para verificar a presença de efeitos significativos nos dados. Como o interesse desta pesquisa é na técnica gráfica, esse teste não será detalhado aqui². Uma consideração mais detalhada desse artigo e da técnica gráfica proposta encontra-se no Capítulo 3 desta dissertação.

Outro artigo de Daniel que também merece destaque é o *Locating Outliers in Factorial Experiments*, publicado em maio de 1960. Neste artigo, Daniel desenvolve um teste formal para localizar e corrigir valores discrepantes em experimentos fatoriais não replicados. Dentre outras coisas, afirma que o gráfico semi-normal pode ser utilizado para a identificação de valores discrepantes nestes experimentos. Tal gráfico, segundo Daniel, ainda tem a vantagem de que o reconhecimento da presença de um valor discrepante não depende de quais efeitos são julgados significativos. Apesar destas afirmações, não há nenhum exemplo do uso do gráfico semi-normal no artigo.

Outro destaque encontra-se no artigo de Anscombe e Tukey (1963). Além de suas importantes contribuições para a análise de resíduos, tais como a formulação de testes para detecção, rejeição e transformação de valores discrepantes; os autores foram os primeiros, dentre todos os artigos pesquisados neste periódico, a se referirem ao gráfico normal como *FUNOP: Full Normal Plot*. Mais outros cinco artigos também utilizam esta expressão ao referirem-se ao gráfico normal: Pasternack e Liuzzi (1965); Box e Meyer (1986); Flack e Flores (1989); Freeny e Landwehr (1995); Venter e Steel (1996). Esta expressão é importante pois, com o desenvolvimento do gráfico semi-normal (*half-normal plot*), serve para distinguir estes dois métodos gráficos. No artigo de Anscombe e Tukey (1963) o gráfico normal é apenas citado como um método para verificar a suposição de normalidade dos resíduos.

Esboçando os usos dos gráficos de probabilidade normal nos artigos publicados

²Ver o Capítulo 2 desta dissertação.

pela *Technometrics* no período de 1959 até meados da década de 80, pode-se descrever o seguinte cenário: Iniciando pelo gráfico semi-normal, as citações e usos sobre esta ferramenta gráfica envolviam a avaliação dos efeitos (HUNTER, 1964; GORMAN; TOMAN, 1966; MARGOLIN, 1967); a estimação do erro em experimentos não replicados (ADDELMAN, 1962; 1964); e a detecção de valores discrepantes nos dados (BRADU; HAWKINS, 1982; ATKINSON, 1983; COOK; WANG, 1983). Dentre alguns pontos abordados, pode-se comentar o fato do gráfico ter uma interpretação subjetiva (ADDELMAN, 1961) e de ser mais sensível na presença de poucos efeitos reais (WEBB, 1968). Neste período, também foi feita uma adaptação da técnica para o caso das tabelas de contingência multidimensionais (COX; LAUH, 1967); bem como algumas modificações nos valores críticos dos quantis do gráfico semi-normal (ZAHN, 1975a; 1975b). Sobre o gráfico normal, dos 12 artigos selecionados neste período que abordavam esta técnica, nove deles destacavam seu uso para verificar a suposição de normalidade dos dados, ao passo que os outros três frisavam os seguintes aspectos: analisar a significância dos efeitos, estimar a variância do erro e verificar a presença de valores discrepantes nos dados, em planos fatoriais não replicados (MARGOLIN, 1976); verificar a significância dos efeitos (LABRECQUE, 1977); detectar valores discrepantes na análise de resíduos em modelos de regressão (BECKMAN; COOK, 1983). Como se pode ver, raros foram os usos do gráfico normal em experimentos fatoriais. Mas esta configuração estava para mudar.

Tendo em vista a subjetividade dos procedimentos gráficos, Box e Meyer (1986) publicam a proposta de um método mais formal, analítico, para a análise de experimentos fatoriais não replicados. Neste artigo, intitulado *An Analysis for Unreplicated Fractional Factorials*, os autores apresentam um teste formal, um método iterativo bayesiano proposto para identificar os efeitos significativos em planos fatoriais fracionados não replicados. Trata-se de uma proposta realmente inovadora, cujos detalhes principais são abordados no Capítulo 2 deste trabalho. O ponto de interesse neste artigo, porém, é outro. Este também é o primeiro artigo na *Technometrics* a tratar da controvérsia acerca da opinião de Daniel sobre o uso do gráfico semi-normal. Os autores comentam que, embora tenha dado primazia ao gráfico semi-normal em seu artigo original de 1959; Daniel, por ocasião da publicação de seu livro em 1976³, mudou de

³*Applications of Statistics to Industrial Experimentation*. Para outras considerações a respeito

idéia, e optou pelo uso do gráfico normal, usando como justificativa sua maior habilidade para revelar discrepâncias nas suposições. Em vista disso, Box e Meyer utilizam o gráfico normal em seu artigo com o objetivo de avaliar a significância dos efeitos em experimentos fatoriais não replicados⁴. Outro aspecto importante a ser destacado no artigo de Box e Meyer é o fato deste tratar do princípio da esparsidade ou escassez dos efeitos (“*sparsity effects*”). Na verdade, são os autores que criam esta expressão para explicar a hipótese de que uma grande proporção da variação de um determinado processo é explicada por um pequeno número de variáveis, isto é, apenas uma pequena proporção de fatores tem efeitos realmente significativos.

É possível afirmar que este artigo (BOX; MEYER, 1986) tornou-se, pelo menos dentro da *Technometrics*, um divisor de águas no que tange ao uso dos gráficos de probabilidade normal para a análise de experimentos fatoriais não replicados. De um lado, muitos artigos publicados começam a utilizar o gráfico normal como ferramenta padrão para avaliar a significância dos efeitos em arranjos fatoriais, citando o artigo de Box e Meyer como justificativa para tal escolha (10, num total de 23 artigos, fizeram isso); de outro, alguns trabalhos ainda optam pela escolha do gráfico semi-normal para tal objetivo. Além da subjetividade inerente ao uso das técnicas gráficas; a própria escolha de qual método usar, se o gráfico normal ou o semi-normal, na maioria das vezes, também é feita de modo subjetivo, sem qualquer justificativa aparente. Esta foi a motivação principal do presente trabalho: fornecer subsídios para que esta escolha seja a menos subjetiva possível.

Continuando com a análise dos artigos selecionados, vale a pena ressaltar também o primeiro artigo publicado na *Technometrics* que trata do uso comparativo das duas técnicas gráficas em questão: *Using Simulated Envelopes in the Evaluation of Normal Probability Plots of Regression Residuals* (FLACK; FLORES, 1989). Neste, os autores investigam o método de diagnóstico por envelopes para avaliar o desempenho dos gráficos de probabilidade normal utilizados na análise de resíduos de modelos de regressão; e optam pela comparação entre o gráfico semi-normal e o gráfico normal. Embora

desta mudança de opinião, ver o Capítulo 3 desta dissertação.

⁴Embora os autores tenham apresentado um teste formal para avaliar a significância dos efeitos; este, segundo os mesmos, não é um substituto das técnicas gráficas, mas uma ferramenta útil a ser utilizada em conjunto com elas. “*Plotting is always valuable and in particular can suggest model inadequacies.*” (BOX; MEYER, 1986, p. 12)

suas conclusões não sejam diretamente absorvidas por este trabalho, visto tratar-se de um foco de pesquisa diferente do apresentado aqui⁵, destaca-se um comentário feito na conclusão do artigo. Após fazerem referência ao artigo que serviu de base para seu trabalho⁶, os autores comentam: “*Atkinson (1981) stated that the half-normal plot can highlight extreme values more effectively than full-normal plots. When compared to the simulated-envelope boundaries for the data sets here that have outlying residuals, the full-normal plots exclude residuals more frequently than the half-normal plots. In one data example here, the half-normal plot masks irregularities that are evident in the full-normal plot - an argument for suggesting the plotting of both types for data analyses rather than one or the other.*” (FLACK; FLORES, 1989, p. 224) Embora Atkinson (1981) tenha indicado apenas o gráfico semi-normal, Flack e Flores indicam o uso simultâneo das duas técnicas gráficas para a análise dos resíduos em modelos de regressão.

Ainda em 1989, um artigo também se destaca por sua valorosa contribuição no estudo dos planos fatoriais não replicados. Trata-se do trabalho de Russel V. Lenth intitulado *Quick and Easy Analysis of Unreplicated Factorials*. Neste artigo, Lenth apresenta um método formal, uma técnica alternativa àquela desenvolvida por Box e Meyer (1986), para a análise de experimentos fatoriais não replicados⁷. Sua proposta consiste numa simples fórmula para estimar o erro padrão dos efeitos e, a partir desta, avaliar aqueles que são realmente ativos. O método de Lenth está entre os mais citados, comentados e usados por artigos que tratam da análise de experimentos fatoriais não replicados, tanto entre aqueles publicados na *Technometrics* como fora dela. Para maiores detalhes sobre esta técnica, consultar o Capítulo 2 da presente dissertação.

Alguns outros comentários e observações feitas acerca do tema abordado neste trabalho também podem ser destacados: O gráfico semi-normal serve para estimar a variância do erro e fazer julgamentos subjetivos sobre a realidade dos efeitos observados, e embora não seja um substituto geral para a análise de variância pode ser

⁵Não se desenvolverá neste trabalho cenários de análise de resíduos de modelos de regressão. As aplicações serão feitas para arranjos fatoriais não replicados.

⁶Atkinson, A. C. Two Graphical Displays for Outlying and Influential Observations in Regression. *Biometrika*, v. 68, pp. 13-20, 1981.

⁷Assim como Box e Meyer, Lenth não descarta o uso dos gráficos de probabilidade normal para a análise dos fatoriais não replicados; chegando até mesmo a recomendar tais gráficos por causa de seus “valiosos diagnósticos”. Ele até comenta suas utilidades; embora frise sua principal desvantagem: a subjetividade da interpretação. Lenth não usa tais gráficos em seu método.

muito útil para fazer avaliações grosseiras acerca de efeitos e interações (ADDELMAN, 1961). Uma limitação do gráfico semi-normal é que este pode ser usado para fazer julgamento sobre os efeitos apenas se uma pequena proporção destes for significativa; isto porque, nesta situação, a sensibilidade do gráfico não estará comprometida (HOLMS; BERRETTONI, 1969). Na presença de valores discrepantes nos dados, o gráfico semi-normal pode fornecer informações mais detalhadas e específicas em algumas situações; mostrando-se mais sensível e mais informativo que certas estatísticas (GENTLEMAN; WILK, 1975). Utilizar as interações de alta ordem para estimar a variância do erro experimental não é um procedimento satisfatório; já o procedimento gráfico de Daniel tem a vantagem de não requerer nenhuma identificação prévia de contrastes inertes, estes são identificados automaticamente no gráfico (BOX; MEYER, 1986). Os recursos gráficos têm uma interpretação subjetiva; o que vem explicar as diferentes conclusões que podem ser tiradas por diferentes pesquisadores sobre uma mesma avaliação gráfica (FLACK; FLORES, 1989).

Estes foram alguns dos principais destaques a serem mencionados. A próxima seção trata do segundo estágio do trabalho de busca: a pesquisa por artigos publicados fora da *Technometrics*.

4.2 Levantamento Bibliográfico em Outros Periódicos

Com o intuito de tornar o levantamento bibliográfico o mais abrangente possível, foi feita uma busca por artigos em periódicos fora da *Technometrics*. Para tal, foi escolhido o “*ISI Web of Knowledge*” por se tratar de um conhecido instrumento de busca acadêmico. O objetivo era encontrar o maior número possível de artigos que citam os dois principais trabalhos de Daniel usados nesta dissertação: *Use of Half-Normal Plots in Interpreting Factorial Two-Level Experiments* de 1959, que é o artigo chave do presente trabalho; e *Locating Outliers in Factorial Experiments* de 1960. Assim como na seção anterior, os resultados obtidos são analisados para se construir um referencial teórico sobre o uso dos gráficos de probabilidade normal na análise de experimentos fatoriais. Tal análise também serve de diretriz para a elaboração da metodologia deste trabalho.

A busca acima referida foi realizada em 25 de janeiro de 2009 através do endereço eletrônico *www.pcs.isiknowledge.com.silk.library*. As palavras utilizadas para a busca foram “*Daniel Citations 1959-1960*”. Foram apresentados 391 artigos como resultado. Um pouco mais da metade destes, isto é 212, vieram acompanhados de seus respectivos resumos e palavras-chave. Isto foi extremamente útil para a agilidade dessa fase de busca, pois os artigos selecionados para uma análise mais profunda e detalhada, tendo em vista sua relação com o tema do presente trabalho, foram escolhidos através da leitura destes resumos. Para os outros 179 artigos cujos resumos não apareceram, a busca teve que ser feita um a um, nos sites de seus respectivos periódicos de origem; para só depois selecionar aqueles significativamente importantes para a construção deste referencial teórico.

Nesta fase de procura pelos artigos selecionados, constatou-se um problema: nem todos os artigos ou periódicos estavam disponíveis no portal Periódicos Capes. Dos 391 artigos apontados, 96 (25%) não puderam ser encontrados. Portanto, os resultados discutidos nesta seção referem-se aos 295 (75%) artigos disponíveis. Dentre estes, 43 (16%) são artigos publicados na *Technometrics*; já estando então computados e analisados na seção anterior.

Como se pode ver na Figura 4.6, dentre os 295 artigos disponíveis, 238 (81%) abordavam alguma forma de experimento fatorial e 138 (47%) trataram ou comentaram acerca dos tipos de gráficos de probabilidade discutidos nesta dissertação.

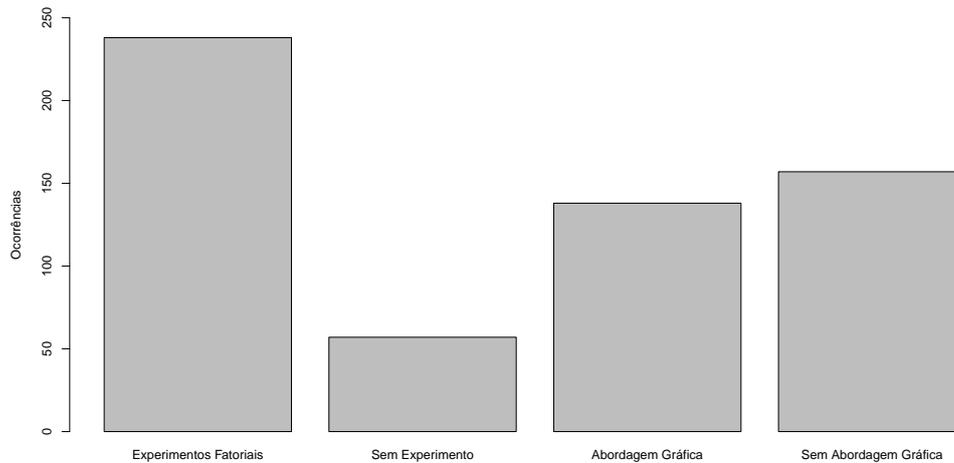


Figura 4.6: **Distribuição do uso de Experimentos e de Gráficos de Probabilidade nos artigos pesquisados.**

A Figura 4.7 mostra que, dentre os 138 artigos que trataram dos gráficos de probabilidade, verificou-se que 16% referiram-se ao gráfico normal; 78% referiram-se ao gráfico semi-normal; e 6% abordaram as duas técnicas.

No que tange à abordagem das formas gráficas nestes artigos, pode-se traçar a seguinte configuração: (1) gráfico normal: dentre os 22 artigos que trataram do gráfico normal, verificou-se que 5% destes apenas citam ou recomendam o uso desta técnica para a avaliação da significância dos efeitos; os outros 95% restantes usam este gráfico, sendo que 24% para a verificação da normalidade dos dados, 10% para a análise de resíduos, e 66% para avaliar a significância dos efeitos dos fatores; (2) gráfico semi-normal: dos 107 artigos que trataram desta forma gráfica, 51% deles apenas comentam o uso da técnica (sendo 75% para a avaliação da significância dos efeitos, 5% para a estimação do erro, 2% para testar a normalidade dos dados, e 18% para detectar valores discrepantes) ao passo que 49% usam este gráfico (dos quais 87% para estimar a significância dos efeitos, 4% para verificar a suposição de normalidade dos dados, 5% para a detecção de valores discrepantes, e 4% para avaliar os efeitos e detectar valores discrepantes); (3) ambos: dos 9 artigos que destacaram as duas técnicas simultaneamente, pôde-se perceber que 44% apenas citam os dois métodos, e em todos estes, como instrumentos de avaliação da significância dos efeitos; os 56% restantes utilizam as duas técnicas gráficas, sendo 60% para avaliar a significância dos

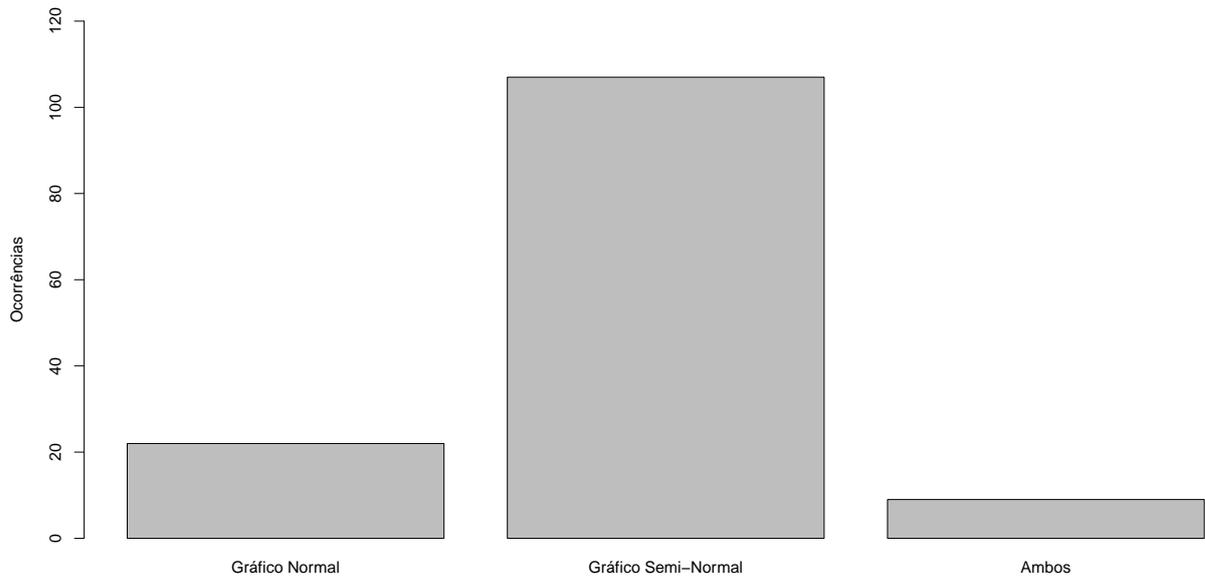


Figura 4.7: **Abordagem das Formas Gráficas.**

efeitos, 20% para a detecção de valores discrepantes e avaliação dos efeitos, e 20% na análise de resíduos. As Figuras 4.8, 4.9 e 4.10 ilustram esses resultados.

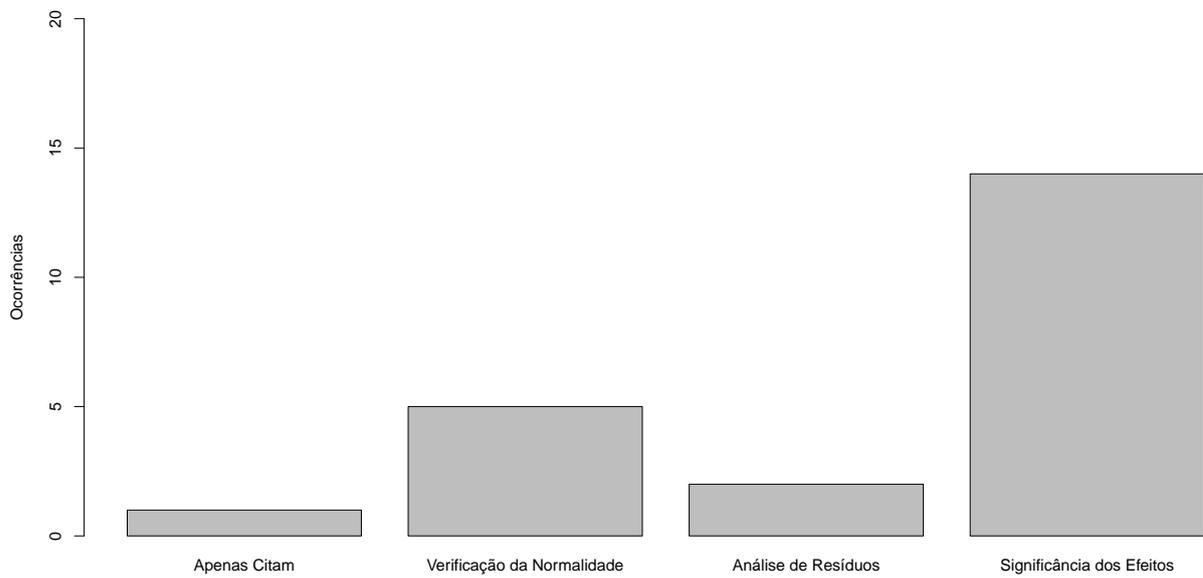


Figura 4.8: Abordagem do Gráfico Normal.

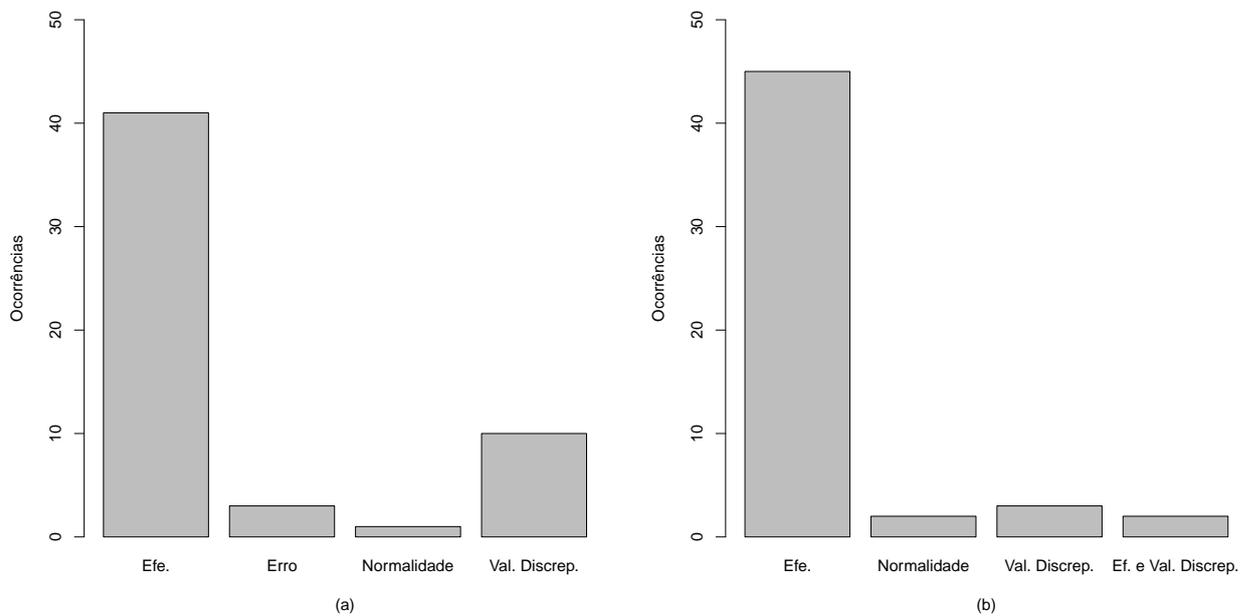


Figura 4.9: Abordagem do Gráfico Semi-Normal. (a) Artigos que Apenas Citam o Gráfico Semi-Normal. (b) Artigos que Usam o Gráfico Semi-Normal.

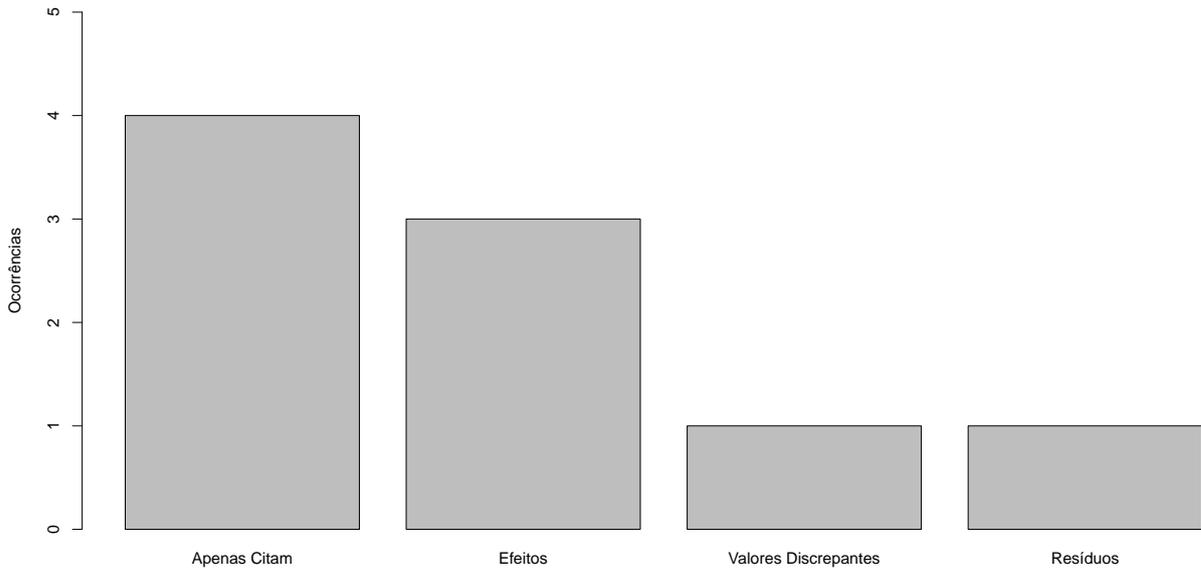


Figura 4.10: Abordagem das Duas Técnicas Gráficas.

4.2.1 Artigos Seleccionados dos Demais Periódicos

Alguns aspectos importantes dos artigos seleccionados na busca feita através do *ISI*, descrita anteriormente, são destacados nesta seção. Por exemplo, Taylor (1994), em seu artigo *Analysis of Experiments by Using Half-Normal Plots*, chama a atenção para alguns cuidados que devem ser tomados no uso e na interpretação do gráfico semi-normal, tais como a subjetividade acerca da posição da reta e dos valores para o cálculo das “faixas de segurança” (*guardrails*) no gráfico; e ainda sobre a habilidade do gráfico de detectar valores discrepantes nos dados. Uma consideração importante neste artigo é sobre a quantidade de efeitos significativos que o gráfico semi-normal é capaz de identificar. O ideal, segundo as recomendações de Daniel e Zahn comentadas por Taylor, é que este gráfico seja usado quando houver no máximo $\frac{n}{5}$ (ou 20%) efeitos significativos, em que n é o número total de efeitos estimados. Se o número de efeitos significativos for superior a essa razão, o gráfico semi-normal não irá detectá-los.

Em seu artigo *A New Look at Half-Normal Plots for Assessing the Significance of Contrasts for Unreplicated Factorials*, Jorge Olguin e Tom Fearn (1997) propõem um método formal, baseado no método de Lenth (1989) e no próprio teste desenvolvido por

Daniel (1959), para avaliar a significância dos efeitos, juntamente com o uso do gráfico semi-normal. Os autores afirmam que o gráfico semi-normal é a recomendação mais comum para avaliar a significância dos efeitos em experimentos fatoriais; lembram que sua construção baseia-se no princípio da esparsidade dos efeitos; e que seu uso pode ser estendido também para experimentos que incluem fatores com mais de dois níveis. Mas chamam a atenção para o fato de que, embora popular, esta é uma técnica subjetiva, visto tratar-se de um procedimento gráfico. Portanto, recomendam que seu uso deve ser complementado por um procedimento mais formal⁸; embora, segundo os autores, tal prática seja rara.

Outro artigo que traz considerações interessantes acerca das vantagens e limitações do uso das técnicas gráficas é o *Robust regression and outlier detection in the evaluation of robustness tests with different experimental designs*, de autoria de Edelgard Hund, D. Luc Massart e Johanna Smeyers-Verbeke (2002). Nele, os autores utilizam as duas técnicas gráficas em diferentes planos experimentais: o gráfico semi-normal para a análise de resíduos e o gráfico normal para a avaliação dos efeitos. Em ambos os casos, o objetivo é verificar o desempenho destas formas de análise em detectar a presença de valores discrepantes nos dados. Pode-se destacar alguns resultados interessantes: quanto maior o valor discrepante maior será a abertura (*gap*) em torno do zero no gráfico normal; um valor discrepante que afeta igualmente todos os efeitos não será detectado num gráfico de probabilidade normal; o gráfico semi-normal dos resíduos não é capaz de identificar múltiplos valores discrepantes, mas apenas um único e, aparentemente, se este for grande; igualmente, parece que o gráfico normal não é satisfatório para a identificação de múltiplos valores discrepantes; quando os fatores estão todos no mesmo nível, o gráfico normal nem sempre indica a presença de valores discrepantes, mesmo que estes sejam grandes; a ausência de abertura no gráfico normal não significa necessariamente que não há valores discrepantes nos dados; o gráfico normal dos efeitos não identificou a maioria dos valores discrepantes, enquanto o gráfico semi-normal dos resíduos identificou apenas valores discrepantes óbvios. Com certeza estas informações serão úteis para a elaboração dos cenários e posterior análise dos resultados neste trabalho.

⁸Na verdade, foi isso o que Daniel (1959) tentou sugerir em seu artigo. Depois de construir o gráfico semi-normal, ele desenvolveu um método analítico, formal, para testar a significância dos efeitos, que deveria ser utilizado para complementar as interpretações feitas com o uso do gráfico.

Ainda na temática sobre a identificação de valores discrepantes em experimentos fatoriais, convém destacar o artigo *SAS macros for analysis of unreplicated 2^k and 2^{k-p} designs with a possible outlier*, de John Lawson (2008). Lawson apresenta uma programação em SAS, com base no método de Daniel (1959), para a análise de experimentos fatoriais não replicados quando um único valor discrepante está presente. A presença de valores discrepantes nos dados compromete as técnicas utilizadas na identificação dos efeitos, pois as estimativas destes são tendenciosas. Com isso Lawson justifica a importância de seu trabalho, pois são poucos os métodos propostos para detectar valores discrepantes e testar a significância dos efeitos em experimentos não replicados na presença destes. Uma interessante observação feita pelo autor é que, para experimentos com mais de 16 provas, os efeitos não são grandemente afetados pela presença de valores discrepantes e, portanto, as técnicas gráficas podem ser usadas para identificar quais deles são ativos. Por outro lado, nos experimentos com até 16 provas, os valores discrepantes provocarão um “viés” em todos os efeitos calculados. Devido a isto, apenas uma análise gráfica não seria suficiente. Um método mais formal para detectar e corrigir este valor discrepante, bem como para estimar os efeitos, deveria ser utilizado. Segundo Lawson, sua proposta cumpre esse objetivo.

Destacam-se ainda outras importantes considerações: Os testes formais não são suficientes para a análise dos dados; estes devem ser acompanhados por procedimentos de avaliação informais, tais como o gráfico semi-normal, pois as informações nele resumidas facilitam as conclusões (WILK; GNANADESIKAN, 1964, 1970; GERSON, 1975). O avanço da informática permitiu que os gráficos de probabilidade normal se tornassem uma ferramenta estatística tão grandemente utilizada (EVANS, 1973); mas devido ao fato de não haver teorias que forneçam critérios de avaliação, um método gráfico pode ser avaliado utilizando-o para analisar um grande número de conjunto de dados, depois verificando que aspectos dos dados são revelados pelo gráfico e, por fim, comparar isto com outros métodos (CLEVELAND, 1987). A análise do gráfico semi-normal fica difícil quando há poucos contrastes presentes; sendo assim, a sugestão dada para que a avaliação seja mais sensível é que este gráfico deva ter no mínimo sete contrastes (SCHOEN, 1999). Uma das razões de Daniel ter recomendado o gráfico normal é que este é mais capaz de detectar anomalias nos dados que o gráfico semi-normal (SCHOEN; KAUL, 2000). Schoen e Kaul (2000) também argumentam que a

subjetividade das técnicas gráficas encontra-se no fato de decidir quão grande tem que ser um desvio da reta para que o efeito seja considerado significativo. Os gráficos de probabilidade normal são as ferramentas mais utilizadas para detectar a significância dos efeitos em planos fatoriais não replicados (AGUIRRE-TORRES; PÉREZ-TREJO, 2001; HOLM; MARK; ADOLFSSON, 2005). A escolha do tipo de gráfico utilizado é uma questão de preferência do pesquisador (DODGSON, 2003). Ainda no campo da subjetividade das técnicas gráficas, Auer e Kunert (2005) defendem a opinião de que, ao analisar um conjunto de dados, um estatístico deve ter um modelo em mente, bem como algum conhecimento dos métodos relacionados com o uso do mesmo. Este conhecimento pode ser baseado em teoria matemática ou, no caso das situações mais complicadas, em simulações. Daí criticam o uso da intuição por parte de alguns autores, devido ao fato de esta “não ser um bom guia ao tentar identificar estruturas aleatórias”.

Grande parte dos artigos selecionados trata da construção, uso e interpretação dos gráficos de probabilidade normal em situações reais ou simuladas, sobre experimentos fatoriais não replicados, nos mais diversos ramos da ciência. Outros apresentam algumas considerações pontuais e bem específicas sobre a temática em questão, que já foram abordadas nos diversos capítulos desta dissertação. Muitos destes artigos estão citados ou comentados no decorrer de todo este trabalho, nas situações exclusivas de sua abordagem, servindo também como aporte teórico e referencial para a construção dos conhecimentos. Para evitar redundância e repetições desnecessárias, estes não são destacados aqui⁹.

4.3 Síntese

A pesquisa bibliográfica realizada permite a elaboração das seguintes conclusões:

- Os planejamentos fatoriais 2^k são, de fato, muito utilizados na experimentação industrial, como também em outras áreas do conhecimento humano.
- Os gráficos de probabilidade normal ainda são a ferramenta padrão para a análise de experimentos fatoriais não replicados, embora existam outros métodos, tais

⁹Esta consideração final também vale para os artigos selecionados da *Technometrics*.

como os discutidos no Capítulo 2. A avaliação da significância dos efeitos é feita principalmente através do gráfico semi-normal, embora o uso do gráfico normal esteja em ascensão.

- O principal uso do gráfico semi-normal é, sem sombra de dúvida, para avaliar a significância dos efeitos. Em seguida, encontra-se seu uso como um instrumento para detectar valores discrepantes nos dados, como também para estimar o erro experimental. Poucos foram os trabalhos encontrados nos quais seu uso se deu nos arranjos com parcelas subdivididas não replicados (DANIEL, 1959; SCHOEN; WOLFF, 1997; SCHOEN, 1999; SITTER; LOEPPKY, 2002; SIT-TER; BINGHAM, 2003; MONTGOMERY; KULAHCI; ALMIMI, 2007). Muito menos ainda são os usos do gráfico semi-normal na análise de planos strip-plot (MILLER, 1997; MILLIKEN *et. all.*, 1998).
- Quanto ao gráfico normal, seu uso principal na experimentação ainda é para verificar a suposição de normalidade dos dados, bem como detectar valores discrepantes na análise dos resíduos, em modelos de regressão. Sua utilização na avaliação dos efeitos já é bastante difundida, e vem crescendo no decorrer dos anos. Na ocorrência dos planos com parcelas subdivididas, é a técnica gráfica mais utilizada nas análises.
- Poucos são os trabalhos em que seus autores justificam o motivo da escolha do método gráfico utilizado. Pode-se destacar alguns destes: “*The reason for doing half-normal plotting is that the sign of the contrast is, from the null viewpoint, irrelevant and one may obtain a more stable and focused display. One reason for having also the full-normal plot is the possible interest in the actual signs of the individual or groups of contrasts when in fact some may be exhibited as being non-null. Another reason is to aid in exhibiting possible distributional peculiarities, when some of the contrasts reflect real experimental effects, which may coincidentally be concealed when the distribution is ‘folded’.*” (WILK; GNANADESIKAN, 1968, p. 13) O gráfico semi-normal provê informações mais detalhadas e específicas em algumas situações. (GENTLEMAN; WILK, 1975) O gráfico semi-normal exhibe valores discrepantes grandes mais eficientemente que o gráfico normal (AT-KINSON, 1981); portanto, provê uma indicação mais forte da presença destes

valores do que o gráfico normal (ATKINSON, 1982). O gráfico semi-normal é bem mais apropriado que o próprio gráfico dos resíduos, pois é mais sensível à presença de valores discrepantes nos dados. (BRADU; HAWKINS, 1982) O gráfico normal parece preferível no caso dos experimentos não replicados; em parte porque pode ser dada uma interpretação física aos sinais de importantes contrastes, e em parte porque se os níveis estão definidos de tal forma que os efeitos principais sejam todos positivos, as interações terão um sinal interpretável; como por exemplo, uma interação positiva indica um reforço dos efeitos separados. (COX, 1984) *“Box ... preferred the use of the full normal plot, i.e., a normal probability plot of the signed contrasts. The reason is that the signs carry information that may be useful in the analysis. However, in most analyses signs are at least partially arbitrary: the same experiment with different codings for the factor levels might lead to different normal plots but only to one half-normal plot, so we consider only half-normal plots.”* (OLGUIN; FEARN, 1997, p. 450) *“We prefer the original half-normal plots (DANIEL, 1959), because these are insensitive to the arbitrariness of sign of the effects.”* (SCHOEN; WOLFF, 1997, p. 412) *“The interpretation of such a half-normal plot is often more evident than for the corresponding normal probability plot, because the data are presented in a more compressed manner.”* (HUND; MASSART; SMEYERS-VERBEKE, 2002, p. 55) A mudança de opinião de Daniel acerca do uso do gráfico semi-normal também foi usada como justificativa (SCHOEN; KAUL, 2000). Apesar dessa exposição de motivos encontrada em alguns casos, a imensa maioria dos artigos pesquisados não apresentou alguma justificativa para a escolha da técnica gráfica utilizada.

- Nenhum dos artigos encontrados apresentou uma abordagem comparativa das duas formas gráficas nos moldes da que está sendo proposta neste trabalho.

Os resultados acima obtidos vêm motivar esta dissertação e destacar sua importância. Primeiramente, a originalidade deste trabalho é evidente na proposta apresentada: comparar o desempenho do gráfico semi-normal e do gráfico normal na análise de experimentos fatoriais não replicados, sob os mais diferentes cenários em que tais experimentos podem ser planejados. Não há registro de proposta anterior semelhante a esta, pelo menos até o momento e dentre os artigos disponíveis para a pesquisa. Ou-

tro ponto importante a ser enfatizado é a relevante contribuição deste trabalho para a área de planejamento de experimentos. Uma vez que não há nenhum procedimento formal ou teste estatístico que indique “qual dos dois gráficos é melhor”, pois o uso ou não deles é, na maioria das vezes, uma questão meramente subjetiva; os resultados apresentados nesta dissertação poderão servir como diretrizes nessa tomada de decisão; uma pequena contribuição para essa questão tão controversa.

Capítulo 5

Estudo Comparativo dos Gráficos de Probabilidade Normal

Este capítulo aborda o estudo experimental proposto por este trabalho. A primeira seção trata da metodologia utilizada nos cenários. A segunda seção apresenta os resultados encontrados nos cenários construídos com os dados dos artigos selecionados do levantamento bibliográfico. A terceira seção discute os resultados obtidos de cenários simulados, sempre no enfoque da comparação das duas técnicas gráficas e seus usos em várias situações.

5.1 Metodologia

Os cenários apresentados neste capítulo são construídos com base nas indicações do levantamento bibliográfico realizado. Estes encontram-se divididos em três grupos. Cada grupo representa o aspecto sob o qual o desempenho dos gráficos é avaliado: (1) o julgamento da significância dos efeitos dos fatores; (2) a capacidade de detectar a presença de valores discrepantes (*outliers*) nos dados; (3) a identificação de restrições na aleatorização: o caso das parcelas subdivididas (*inadvertent plot-splitting*). Estes aspectos são escolhidos, pois, segundo a proposta de Daniel (1959), o gráfico semi-normal é uma ferramenta útil para estas três situações, assim como também o gráfico normal.

Inicialmente, para cada cenário apresentado são construídas as duas formas gráficas: o gráfico normal e o gráfico semi-normal. O objetivo é comparar o desempenho destas duas técnicas sob as três situações apresentadas no parágrafo anterior. Busca-se

avaliar em que condições e sob quais aspectos uma técnica gráfica é mais útil do que outra; sendo este “mais útil” no sentido de fornecer informações mais claras e evidentes das características dos dados. Todos os cenários são executados no pacote estatístico livre R, versão R 2.10.0, e envolvem dados de experimentos fatoriais 2^k e 2^{k-p} não replicados. Concentra-se nos fatoriais cujas provas são 8, 16, 32 e 64 por serem os mais utilizados em experimentação (DANIEL, 1976; DONG, 1993; LAWSON, 2008).

Embora as interpretações de instrumentos gráficos sejam subjetivas, o intuito deste trabalho é fazer com que pelo menos a escolha de qual técnica usar seja a menos subjetiva possível. Os resultados apresentados e as conclusões feitas a partir destes cenários servem de subsídio teórico para que um investigador, tendo em mente seu plano experimental e a configuração dos seus dados, saiba que método gráfico seria mais apropriado para a consecução de seus objetivos. Busca-se também contribuir para a controversa questão que envolve o uso do gráfico normal e do gráfico semi-normal, apresentada nos capítulos anteriores.

Vale ressaltar que nenhum teste formal para a análise de experimentos fatoriais não replicados é fornecido aqui. O trabalho concentra-se no uso das técnicas gráficas em tais experimentos que, conforme já comentado muitas vezes nesse texto, são ferramentas não formais e, portanto, subjetivas nas informações que fornecem; devendo ser utilizadas, sempre que possível, em associação com algum teste. Os leitores interessados em estudar os testes formais que podem ser usados para a análise dos fatoriais não replicados podem consultar a bibliografia sugerida, bem como o Capítulo 2, no qual alguns desses testes são comentados.

O estudo comparativo das duas técnicas gráficas dá-se em duas etapas. Na primeira, são construídos cenários com base nos dados de experimentos encontrados em artigos selecionados do levantamento bibliográfico. Em seguida, os resultados obtidos na etapa anterior servirão de orientação para a construção de cenários simulados, com o intuito de verificar algumas características observadas.

5.2 Cenários Construídos com os Dados dos Artigos

Esta seção trata da apresentação e análise dos resultados de alguns experimentos selecionados do levantamento bibliográfico descrito no Capítulo 4. Ela está dividida em três subseções, em que cada uma delas representa uma das condições sob a qual as técnicas gráficas são comparadas e avaliadas. Para cada cenário apresentado é feito um pequeno comentário sobre o experimento realizado, e é indicado o artigo que fornece os dados, quando for o caso. Ao final de cada subseção é feita a análise dos resultados.

5.2.1 Julgamento da Significância dos Efeitos

O principal uso dos gráficos de probabilidade nos experimentos fatoriais não replicados é avaliar a significância dos efeitos (MONTGOMERY, 2001). Tanto no gráfico semi-normal quanto no gráfico normal, um efeito significativo é identificado como sendo aquele ponto que se distancia mais acentuadamente da reta pela qual os demais pontos se ajustam (BOX; MEYER, 1986; DANIEL, 1959). Como a interpretação gráfica é subjetiva, surgem alguns questionamentos: “O que seria uma distância acentuada da reta? Um ponto a certa distância da reta é suficientemente distante a ponto de ser considerado um efeito real?” Muitas pessoas podem dar respostas diferentes ao interpretarem um mesmo gráfico (CLEVELAND, 1987), uma vez que, neste caso, a visualização e interpretação de “distante” é algo particular.

Nesta Subseção, portanto, os dois métodos gráficos são comparados com o objetivo de responder à seguinte pergunta: “Embora a interpretação gráfica seja subjetiva, em qual dos dois métodos o julgamento dos efeitos é mais fácil?” Em outras palavras: “Em qual gráfico a visualização de um efeito real é mais clara? Qual deles apresenta uma distância maior, em relação à reta, do ponto que representa um efeito significativo, tornando mais clara e fácil sua identificação?” Os cenários apresentados a seguir fornecem algumas respostas. No intuito de evitar repetições desnecessárias, as apresentações gráficas seguem o seguinte padrão: quando os dois gráficos estiverem sendo comparados, estes aparecerão juntos numa mesma figura, lado a lado, com o gráfico semi-normal à esquerda e o gráfico normal à direita.

Exemplo 5.1 (DANIEL, 1959): Trata-se de um experimento fatorial 2^5 sobre a produção de penicilina. O objetivo é verificar quais os efeitos significativos. No experimento realizado, os efeitos julgados significativos foram A , C , E , e provavelmente CE e $ABCDE$ (veja os Exemplos 3.1 e 3.2 do Capítulo 3).

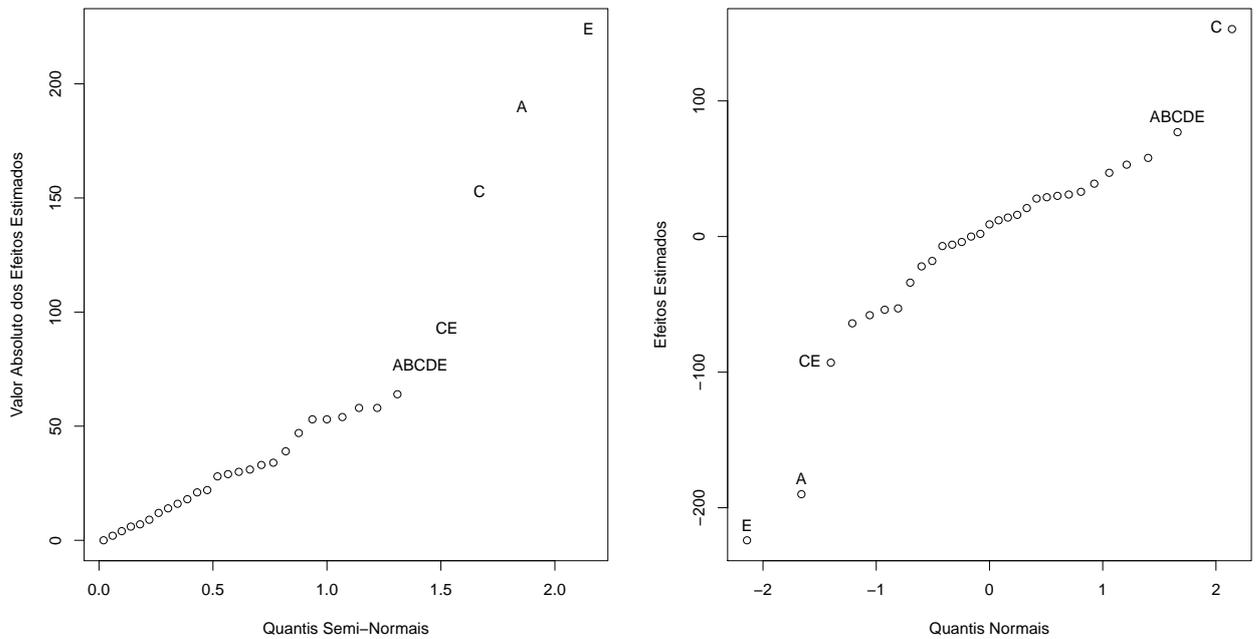


Figura 5.1: Gráficos de Probabilidade para o Exemplo 5.1.

Exemplo 5.2 (HUNTER, 1964): Trata-se de um experimento fatorial 2^4 completo. Os efeitos A e C foram julgados significativos.

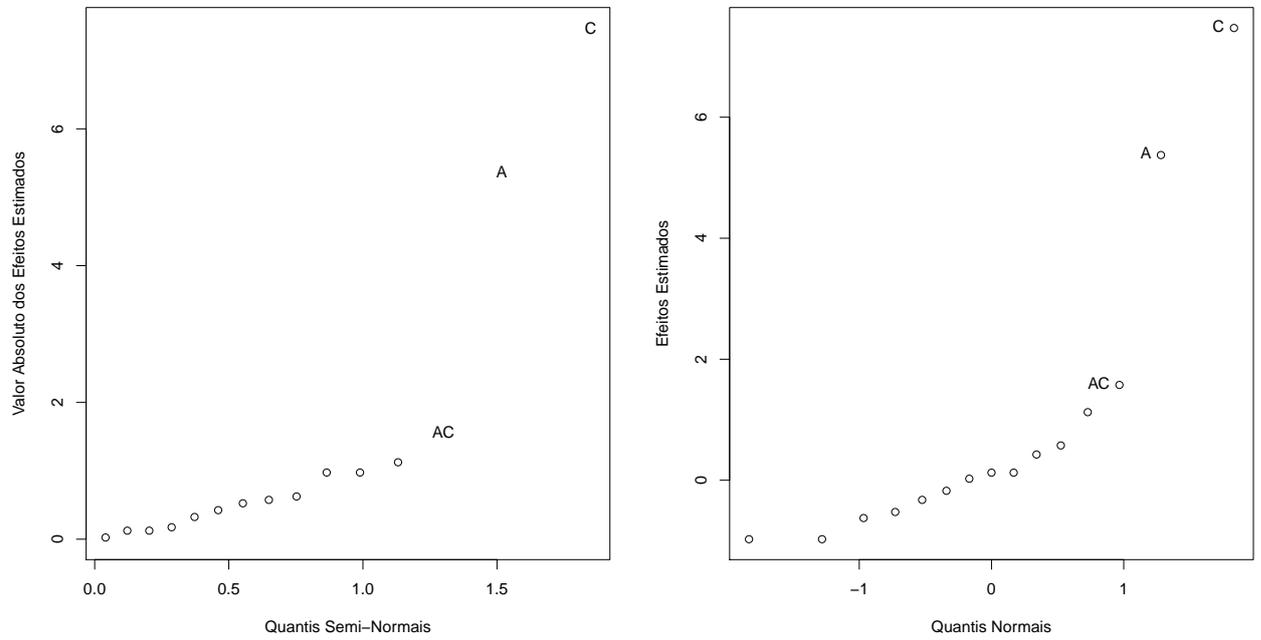


Figura 5.2: Gráficos de Probabilidade para o Exemplo 5.2.

Exemplo 5.3: Trata-se agora da fração um meio do experimento anterior (usando o mesmo conjunto de dados), isto é, 2^{4-1} , cuja relação definidora é $I = ABCD$. Os efeitos A e C ainda são considerados significantivos.

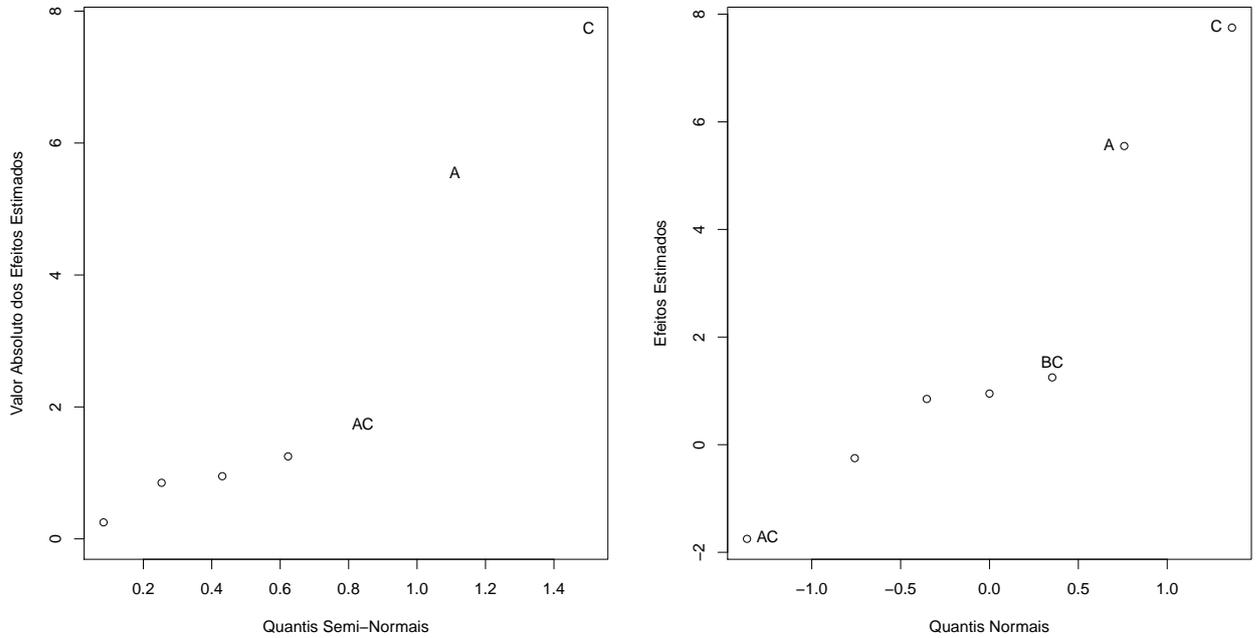


Figura 5.3: Gráficos de Probabilidade para o Exemplo 5.3.

Exemplo 5.4 (BOX; MEYER, 1986): Trata-se de um fatorial 2^4 , citado no artigo como exemplo I. O objetivo foi julgar a significância dos efeitos utilizando o método apresentado no artigo: o gráfico de Bayes. Os efeitos B e D foram julgados significativos. O efeito CD também foi julgado significativo, embora, segundo os autores, não com tanta certeza.

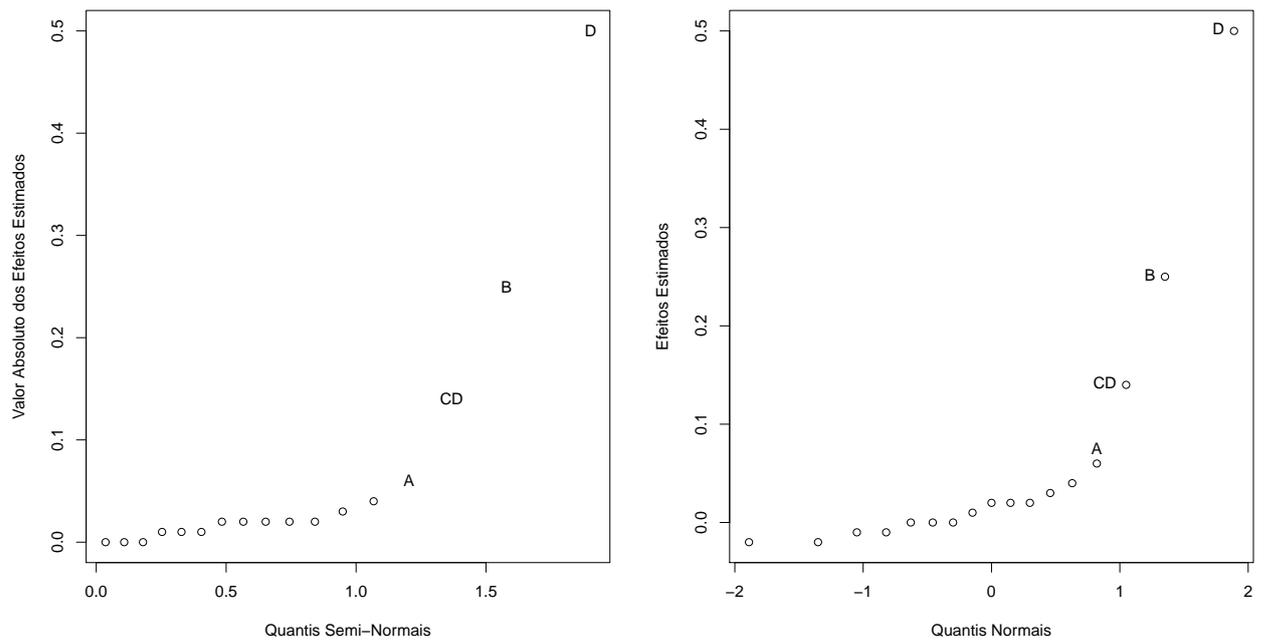


Figura 5.4: Gráficos de Probabilidade para o Exemplo 5.4.

Exemplo 5.5 (BOX; MEYER, 1986): Trata-se de um fatorial 2^{9-5} , citado no artigo como exemplo II. Os efeitos $ABCD$ e BCD foram julgados significativos pelo método proposto pelos autores.

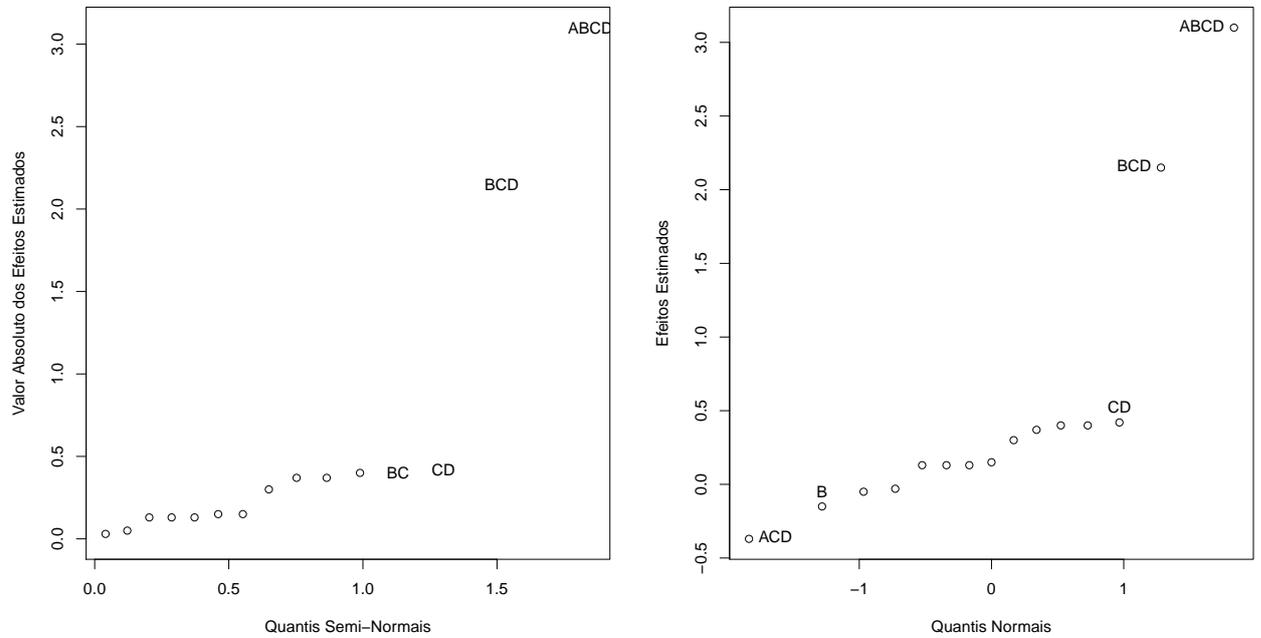


Figura 5.5: Gráficos de Probabilidade para o Exemplo 5.5.

Exemplo 5.6 (BOX; MEYER, 1986): Trata-se de um fatorial 2^{8-4} , citado no artigo como exemplo III. Os efeitos D , ABD e ACD foram julgados significativos.

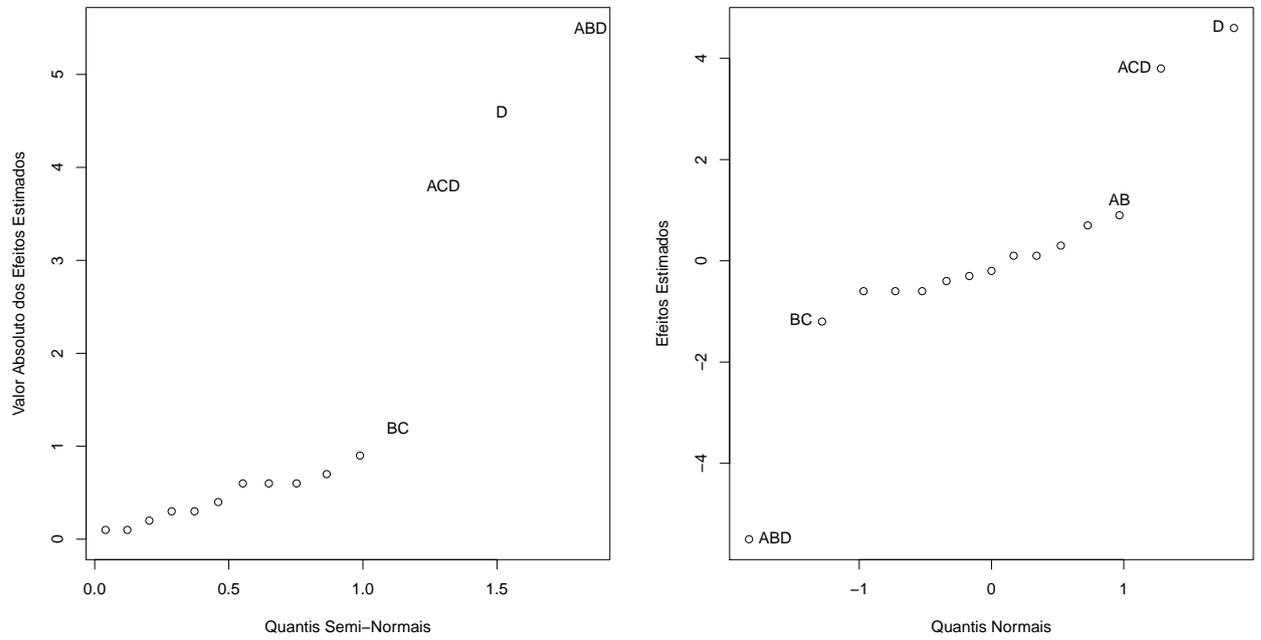


Figura 5.6: Gráficos de Probabilidade para o Exemplo 5.6.

Exemplo 5.7 (BOX; MEYER, 1986): Trata-se de um fatorial 2^4 , citado no artigo como exemplo IV. É digno de nota que de todos os exemplos apresentados neste artigo, este recebe um destaque à parte, pois, segundo os autores, ilustra uma situação mais problemática. Eles comentam que no artigo original do qual este experimento foi consultado, julgou-se, duvidosamente, que os efeitos BC e CD eram significativos. Porém, a disposição dos pontos nos gráficos abaixo, bem como a realização dos testes formais feitos por Box e Meyer, apontam baixíssimas evidências de que algum efeito seja significativo neste experimento. Os autores então concluem afirmando que é impossível tirar conclusões seguras sobre a significância dos efeitos neste experimento. Eles também não descartam a possibilidade de certos efeitos ativos ocorrerem. Para uma análise mais confiável da situação, os autores recomendam replicar o experimento, pois isto proveria uma maior precisão da estimativa dos efeitos, como também uma estimativa independente para o erro experimental.

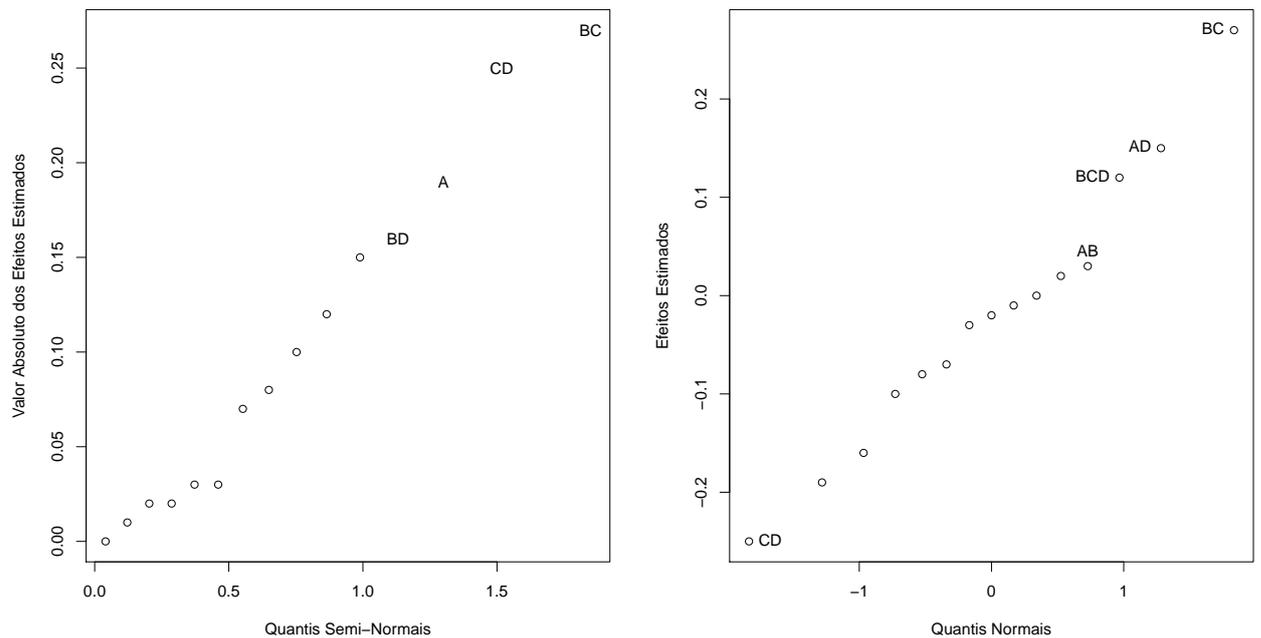


Figura 5.7: Gráficos de Probabilidade para o Exemplo 5.7.

Exemplo 5.8 (TAYLOR, 1994): Trata-se de um fatorial 2^4 , no qual os autores usam o gráfico semi-normal e o teste das faixas de segurança (DANIEL, 1959) para avaliar a significância dos efeitos. Pelo resultado do teste feito no artigo, os efeitos C e D são julgados significativos.

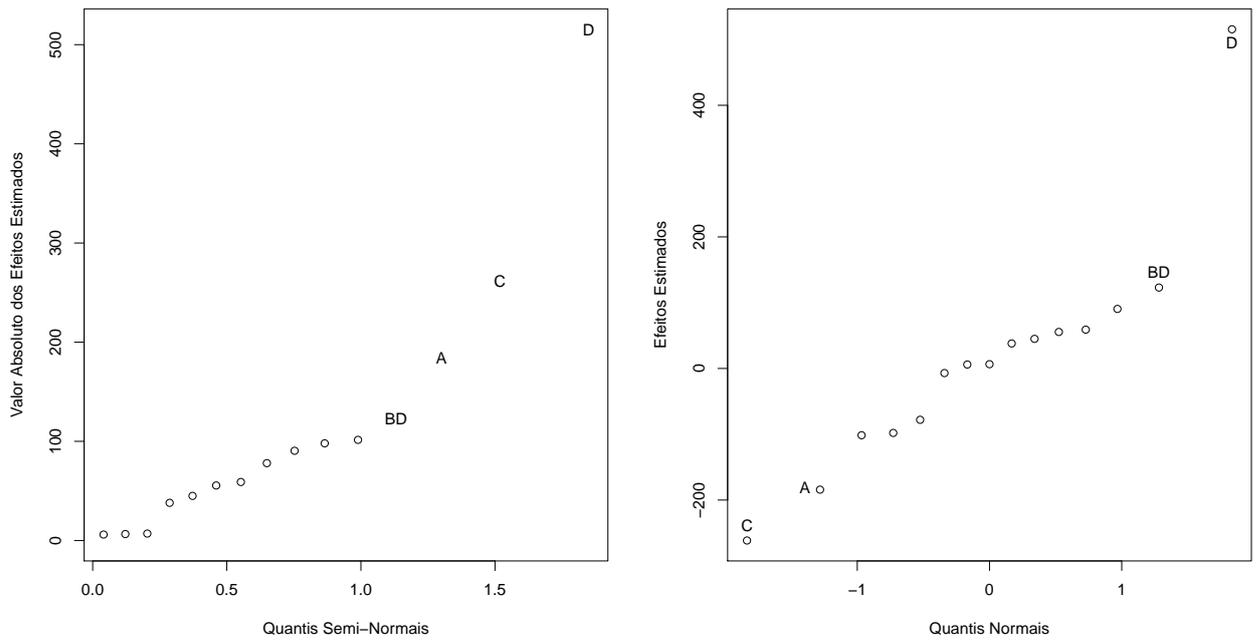


Figura 5.8: Gráficos de Probabilidade para o Exemplo 5.8.

Exemplo 5.9 (BOX, MEYER; STEINBERG, 1996): Trata-se de um fatorial 2^{8-4} com as seguinte relação definidora: $I = ABDH = ACEH = BCFH = ABCG$. Os autores usam o teste Box-Meyer para julgar a significância dos efeitos. Eles comentam que os efeitos principais C e E claramente se evidenciam como significativos; e então chamam a atenção para o fato de o outro efeito ativo não ser um efeito principal, mas sim um efeito confundido com quatro interações de dois fatores, uma vez que $AE = AE + BF + CH + DG$. Os gráficos mostrados na Figura 5.9 também apontam para esta mesma conclusão. Segundo os autores, uma análise inicial provavelmente ignoraria interações de alta ordem, tais como esta que foi considerada significativa; o que seria uma conclusão equivocada para este experimento em questão. Deve-se lembrar que, para este caso, cada efeito principal está confundido com 7 diferentes interações de terceira ordem.

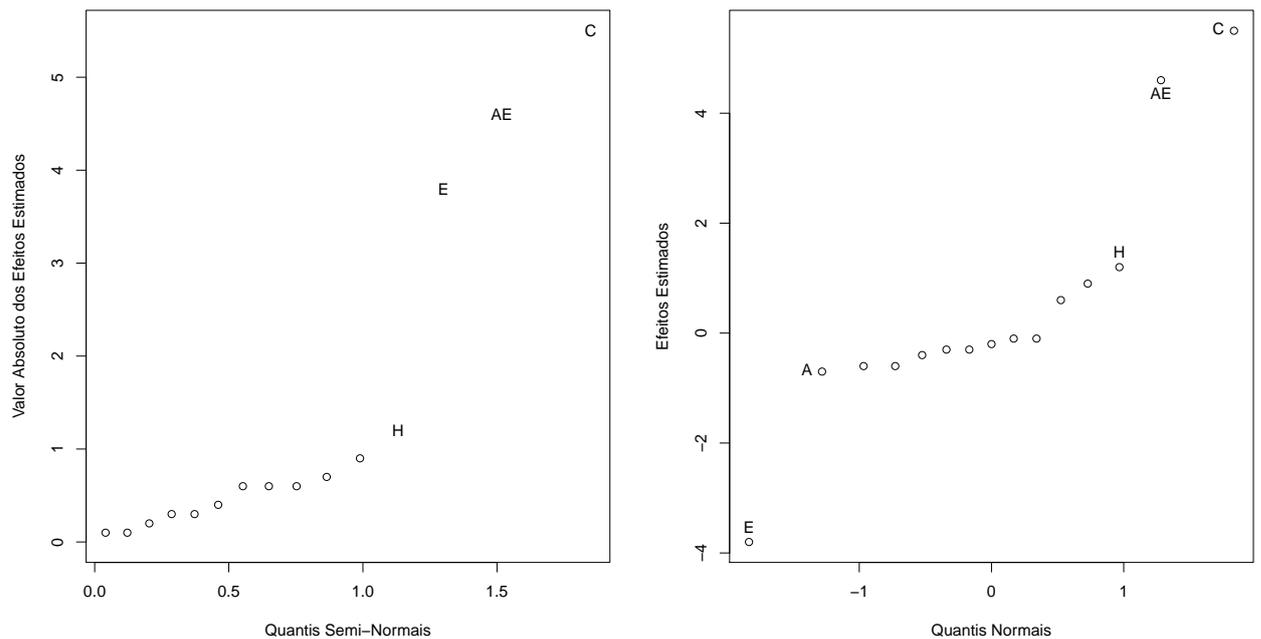


Figura 5.9: Gráficos de Probabilidade para o Exemplo 5.9.

Exemplo 5.10 (BERGMAN; HYNÉN, 1997): Trata-se de um fatorial fracionado 2^{5-1} . O efeito D é considerado significativo.

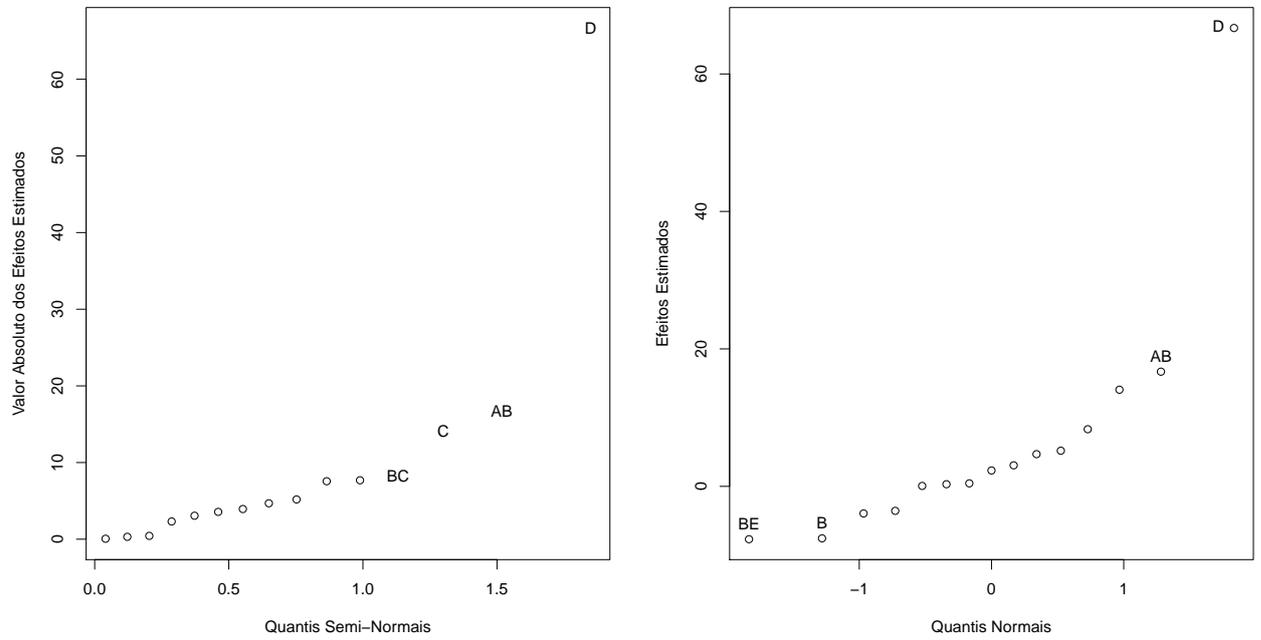


Figura 5.10: Gráficos de Probabilidade para o Exemplo 5.10.

Exemplo 5.11 (YE; HAMADA; WU, 2001): Trata-se de um fatorial 2^4 . Vale destacar algo interessante neste exemplo. Utilizando o método original de Lenth, o único efeito julgado significativo neste experimento é o efeito principal A . Por sua vez, utilizando o método proposto pelos autores no artigo, o Método Lenth Descendente, os autores encontraram cinco efeitos significativos: A , B , C , D e AC . Já nos gráficos de probabilidade mostrados na Figura 5.11, parece não haver indicação de que haja algum efeito significativo nos dados. Segundo os autores, uma vez que a diferença entre os valores absolutos dos contrastes não é muito grande, o uso do gráfico semi-normal pode não detectar qualquer efeito.

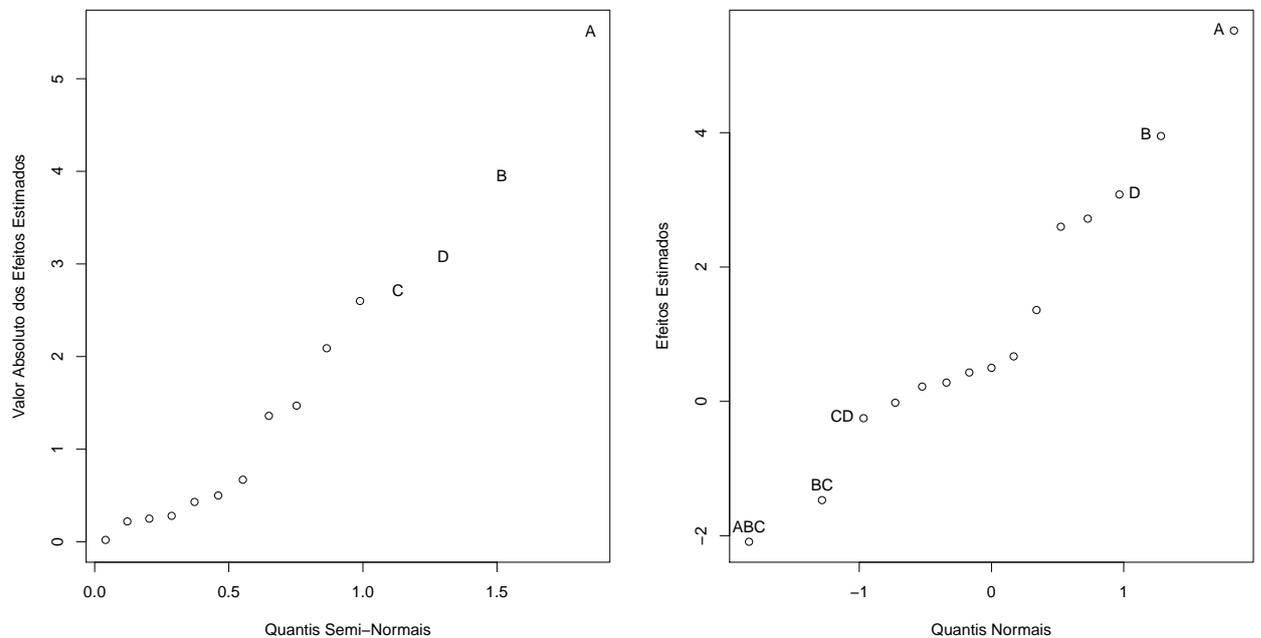


Figura 5.11: Gráficos de Probabilidade para o Exemplo 5.11.

Exemplo 5.12 (BISGAARD; VIVACQUA; PINHO, 2004): Trata-se de um fatorial 2^4 . Os gráficos com os dados originais do experimento são mostrados na Figura 5.12. Aparentemente, nenhum efeito se destaca como sendo significativo. Porém, segundo os autores, uma outra conclusão seria que os efeitos principais B e D e algumas interações de segunda ordem (AB e BD) ou de terceira ordem (ABD e BCD) sejam importantes. Para verificar esta segunda opção, os autores realizam algumas transformações nos dados.

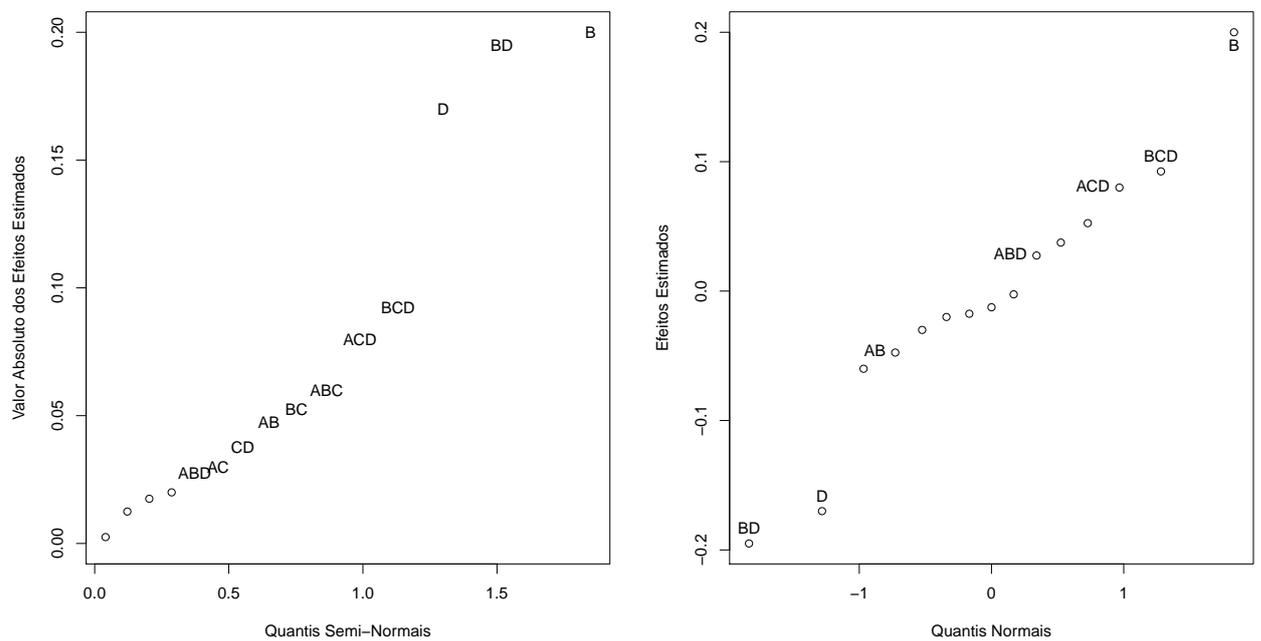


Figura 5.12: Gráficos de Probabilidade para o Exemplo 5.12.

Exemplo 5.13 (BISGAARD; VIVACQUA; PINHO, 2004): Ao realizarem a transformação inversa nos dados, três efeitos aparecem como sendo significativos: B , D , e a interação entre eles BD , como se pode ver na Figura 5.13. E neste caso, acontece algo que Daniel (1976) chama a atenção em seu livro: quando o valor absoluto de dois efeitos principais e de sua interação é aproximadamente o mesmo, então seus efeitos produzem a mesma resposta média. Para o caso em questão temos $|B| = 0,20 \approx |D| = 0,17 \approx |BD| = 0,195$. Segundo Daniel (1976), esta situação é mais claramente percebida no gráfico semi-normal, pois os valores absolutos dos efeitos ficam quase na mesma altura. Esta situação é chamada de “*critical mix*” (BISGAARD; VIVACQUA; PINHO, 2004) ou mistura crítica.

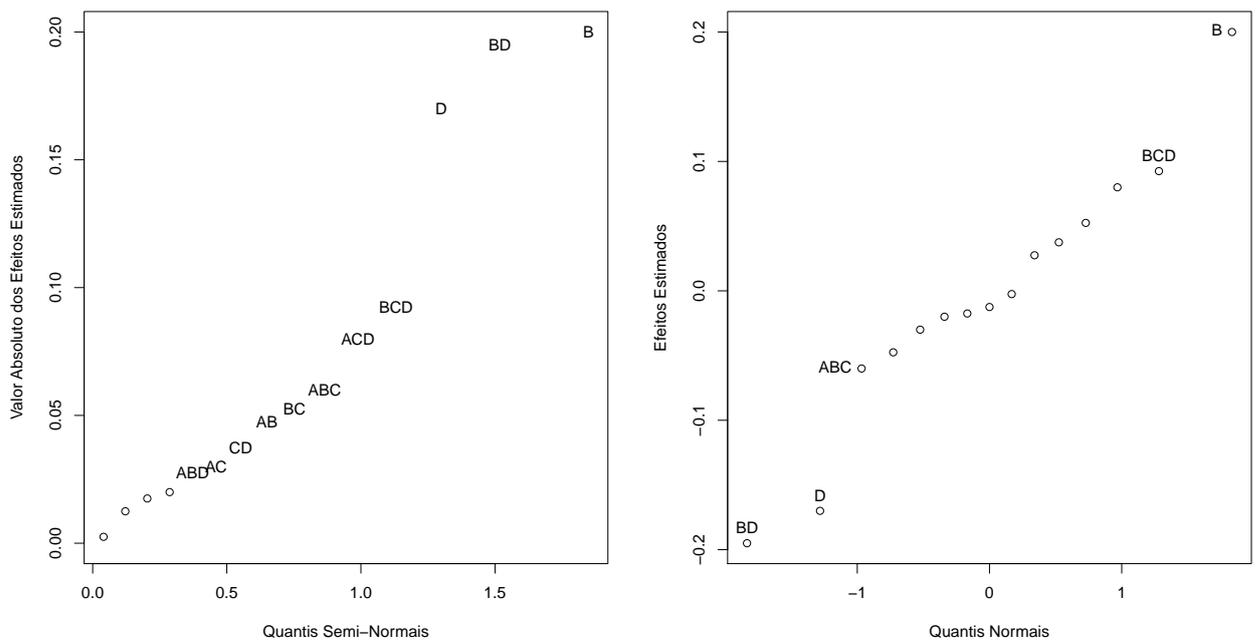


Figura 5.13: Gráficos de Probabilidade para o Exemplo 5.13.

Exemplo 5.14 (MILLER, 2005): Trata-se de um fatorial fracionado 2^{7-4} . Os efeitos A , C e AC são julgados significativos.

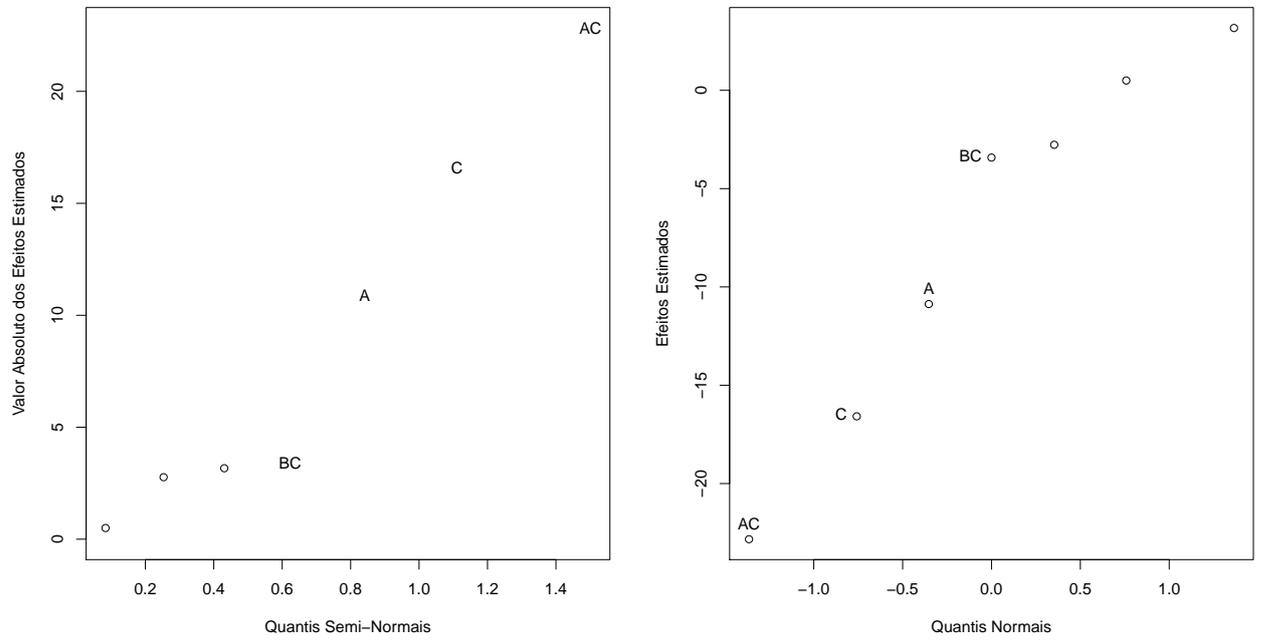


Figura 5.14: Gráficos de Probabilidade para o Exemplo 5.14.

Exemplo 5.15 (VARIYATH; ABRAHAM; CHEN, 2005): Trata-se de um fatorial 2^4 . O objetivo era julgar a significância dos efeitos comparando o gráfico semi-normal, o método de Lenth e o método proposto no artigo, que é o método Jackknife. O gráfico semi-normal e o método de Lenth apontam apenas o efeito D como sendo ativo. Mas, ao aplicarem seu método, os autores julgam os efeitos A e D como sendo significativos.

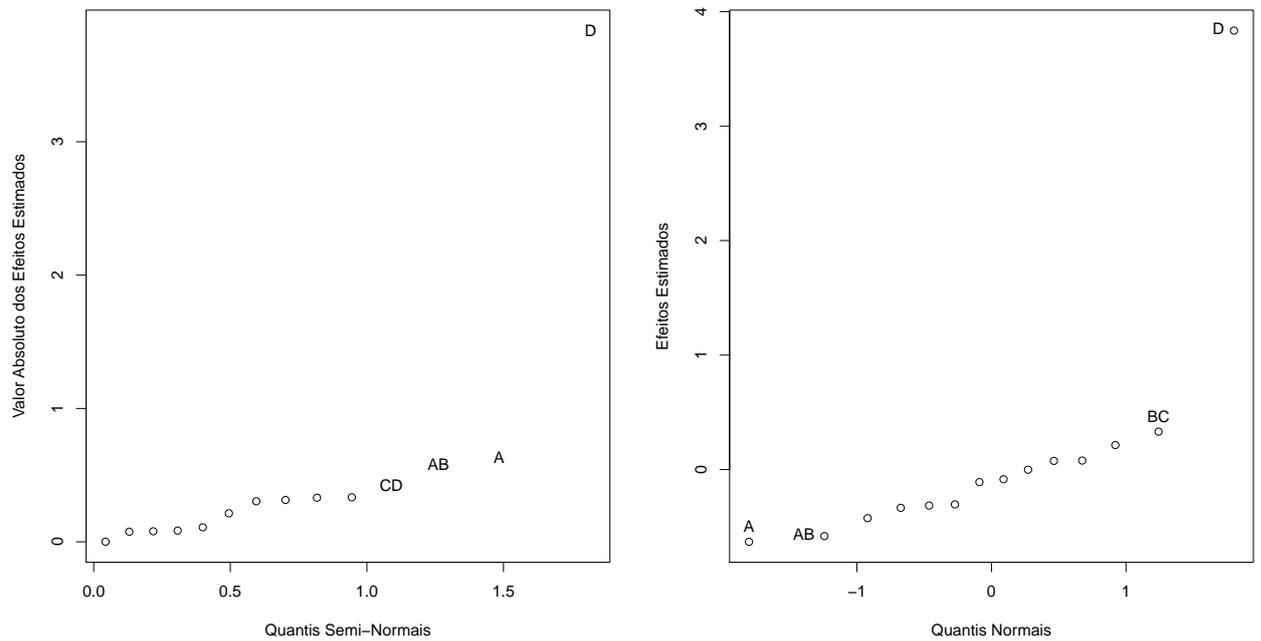


Figura 5.15: Gráficos de Probabilidade para o Exemplo 5.15.

Exemplo 5.16 (DANIEL, 1976): Trata-se de um fatorial 2^5 . Os efeitos B e ABD são julgados significativos. Este experimento é apenas uma ilustração do que foi comentado no Exemplo 5.13: a mistura crítica $|A| \approx |E| \approx |AE|$ é mais claramente percebida no gráfico semi-normal.

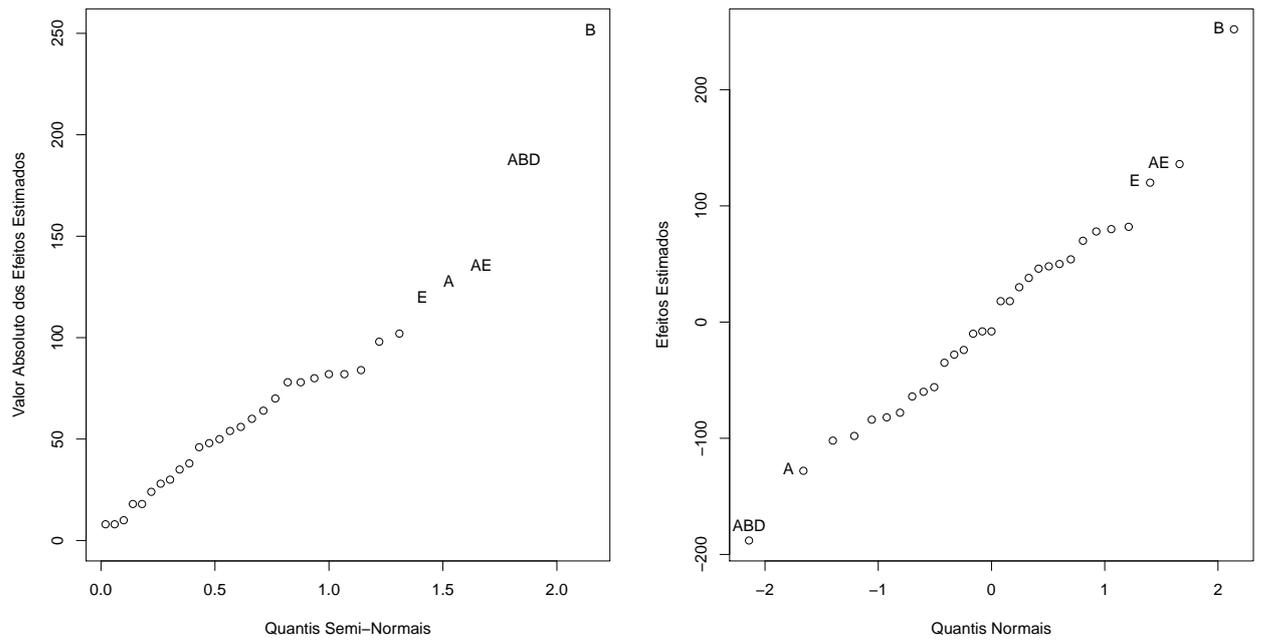


Figura 5.16: Gráficos de Probabilidade para o Exemplo 5.16.

Análise dos Resultados

A Tabela 5.1 resume alguns resultados dos exemplos desta Seção. Como se pode notar, duas importantes características dos experimentos fatoriais são evidentes nos exemplos destacados nesta Seção: (1) efeitos principais e interações de baixa ordem parecem ser mais frequentemente significativos do que interações de ordem mais alta (Princípio da Ordenação Hierárquica); (2) o número de efeitos significativos é relativamente pequeno (Princípio da Esparsidade dos Efeitos).

Tabela 5.1: Características dos experimentos abordados nesta Seção.

Exemplo	Número de provas	Efeitos ativos	Proporção de efeitos ativos
5.1	32	A, C, E	10%
5.2	16	A, C	13%
5.3	$8(2^{4-1})$	A, C	13%
5.4	16	B, D, CD	20%
5.5	16	$ABCD, BCD$	13%
5.6	16	D, ABC, ACD	20%
5.7	16	Nenhum	0%
5.8	16	C, D	13%
5.9	16	C, E, AE	20%
5.10	16	D	7%
5.11	16	Nenhum	0%
5.12	16	Nenhum	0%
5.13	16	B, D, BD	20%
5.14	$8(2^{7-4})$	A, C, AC	2%
5.15	16	D	7%
5.16	32	B, ABD	6%

No que diz respeito ao uso dos gráficos de probabilidade para o julgamento da significância dos efeitos nos exemplos acima destacados, o gráfico semi-normal apresentou um desempenho aparentemente melhor, pois, embora os efeitos julgados reais tenham sido os mesmos em ambos os gráficos, para todos os cenários apresentados nesta seção; estes são mais facilmente identificáveis no gráfico semi-normal em virtude de aparecerem mais nitidamente distantes da reta imaginária do que aqueles que aparecem no gráfico normal. De acordo com Hund, Massart e Smeyers-Verbeke (2002), a interpretação do gráfico semi-normal é frequentemente mais evidente que a do gráfico normal, pois os dados são apresentados de uma maneira mais condensada, resumida. Além disso, o gráfico semi-normal mostrou-se eficaz em mais outro aspecto: ele é mais nítido em mostrar a ocorrência da especial situação em que $|A| \approx |B| \approx |AB|$, chamada

de mistura crítica (BISGAARD; VIVACQUA; PINHO, 2004).

Ainda no aspecto do julgamento dos efeitos, é importante destacar também que o fato dos gráficos de probabilidade (normal e semi-normal) não evidenciarem a presença de algum efeito significativo, não quer dizer que o efeito não exista no problema. Nesta situação, o indicado é aprofundar a investigação do problema, talvez utilizando a replicação ou algum teste formal adequado.

5.2.2 Detecção de Valores Discrepantes nos Dados

Valores discrepantes (*outliers*) são observações que não se ajustam ao padrão dos valores produzidos pelo resto dos dados; podem ser erros, ou então observações precisas, porém inesperadas, que lançam nova luz sobre o fenômeno em estudo; ou podem ser simplesmente manifestações da variabilidade inerente aos dados (STEFANSKY, 1971). Valores discrepantes são observações geradas por mecanismos diferentes daqueles da distribuição assumida (BACON-SHONE; FUNG, 1987). Em experimentos fatoriais, um valor discrepante pode ser devido a uma única interação relativamente grande; pode ser também uma manifestação extrema do sistema de erro aleatório; ou pode ser um engano (DANIEL, 1960). Os valores discrepantes podem ser o impedimento principal para validar a interpretação de dados de experimentos não replicados (LAWSON, 2008). Uma única resposta aberrante num experimento fatorial é o bastante para perturbar todos os efeitos e interações (BOX, 1990).

Daniel (1959) acreditava que os valores discrepantes ou o que ele se refere como valores ruins (*bad values*) são o maior perigo de experimentos fatoriais não replicados, uma vez que eles “inflacionam” os contrastes, podendo fazer com que os efeitos reais não sejam detectados. Ele diz que os valores discrepantes “danificam” o experimento mais que os valores perdidos ou ausentes, pois não são facilmente detectados; e comenta também que sua ocorrência não é tão rara quanto se pensa. Segundo sua própria experiência, a chance de aparecer um valor discrepante em experimentos fatoriais está entre 1% e 10%, dependendo da complexidade da situação experimental ou da experiência do pesquisador.

A presença de um valor discrepante nos dados tem diferentes influências nos diversos efeitos estimados: alguns aumentam seus valores ao passo que outros diminuem, e ainda outros ficam com seus valores absolutos comparáveis, embora com sinais con-

trários (HUND; MASSART; SMEYERS-VERBEKE, 2002). Essa situação tem uma grande influência sobre a decisão da significância dos efeitos, uma vez que todos estão viciados e, por conta disso, o poder dos testes formais é comprometido (LAWSON, 2008). Normalmente, a presença de valores discrepantes é detectada pelo exame dos resíduos do modelo ajustado. O problema, no caso dos fatoriais não replicados, é que os resíduos são todos zero. Por isso, os gráficos de probabilidade normal são usados como ferramentas eficazes na detecção de valores discrepantes (HUND; MASSART; SMEYERS-VERBEKE, 2002).

É fácil entender a influência de um valor discrepante sobre a estimativa dos efeitos. Considere um fatorial 2^k e assumamos que a i -ésima resposta seja um valor discrepante, a qual vale agora $y_i + \Delta$ em vez de apenas y_i . Todos os efeitos são então modificados por $\frac{\Delta}{2^{(k-1)}}$ ou $-\frac{\Delta}{2^{(k-1)}}$, dependendo do sinal que aparece antes da resposta y_i , isto é, aqueles efeitos que possuem um sinal de “+” na linha do tratamento em que se encontra o valor discrepante serão $\frac{\Delta}{2^{(k-1)}}$ unidades maiores; ao passo que aqueles efeitos que possuem o sinal de “-”, serão $\frac{\Delta}{2^{(k-1)}}$ unidades menores. Por conta disso, a presença de valores discrepantes nos dados gera uma configuração padrão característica, específica para cada um dos dois tipos gráficos aqui considerados:

- **Gráfico Normal:** Se Δ é suficientemente grande, alguns efeitos são deslocados para a direita do gráfico e outros para a esquerda. Os pontos próximos a zero parecem seguir duas retas paralelas ao invés de uma; com os valores positivos sobre uma reta e os valores negativos sobre outra. Isto produz uma abertura ou uma lacuna (*gap*) no meio do gráfico; um espaço aberto em torno da origem (BOX, 1990). *“An outlier increases half of the main and interaction effects and decreases the other half . . . a single outlying result should lead to a gap around zero for the effects in the normal probability plot.”* (HUND; MASSART; SMEYERS-VERBEKE, 2002, pg. 55).
- **Gráfico Semi-Normal:** De acordo com Daniel (1960), a presença de um valor discrepante provoca um aumento no valor absoluto da média dos efeitos, fazendo com que o gráfico semi-normal tenha poucos efeitos próximos de zero. Desta forma, a reta traçada pela maioria dos pontos do gráfico não parte da origem, como deveria ser, mas aparece deslocada. Portanto, ao observar esta

forma padrão no gráfico, o pesquisador será avisado sobre a presença de um valor discrepante.

Os aspectos gráficos apontados acima são analisados nos cenários desta seção para comparar as duas técnicas gráficas, a fim de se determinar qual delas é mais útil em apontar ao investigador a presença de valores discrepantes nos dados.

Exemplo 5.17 (BOX, 1990): Trata-se de um fatorial 2^4 com um valor discrepante presente. Pode-se perceber o padrão dos *gaps* nos dois gráficos.

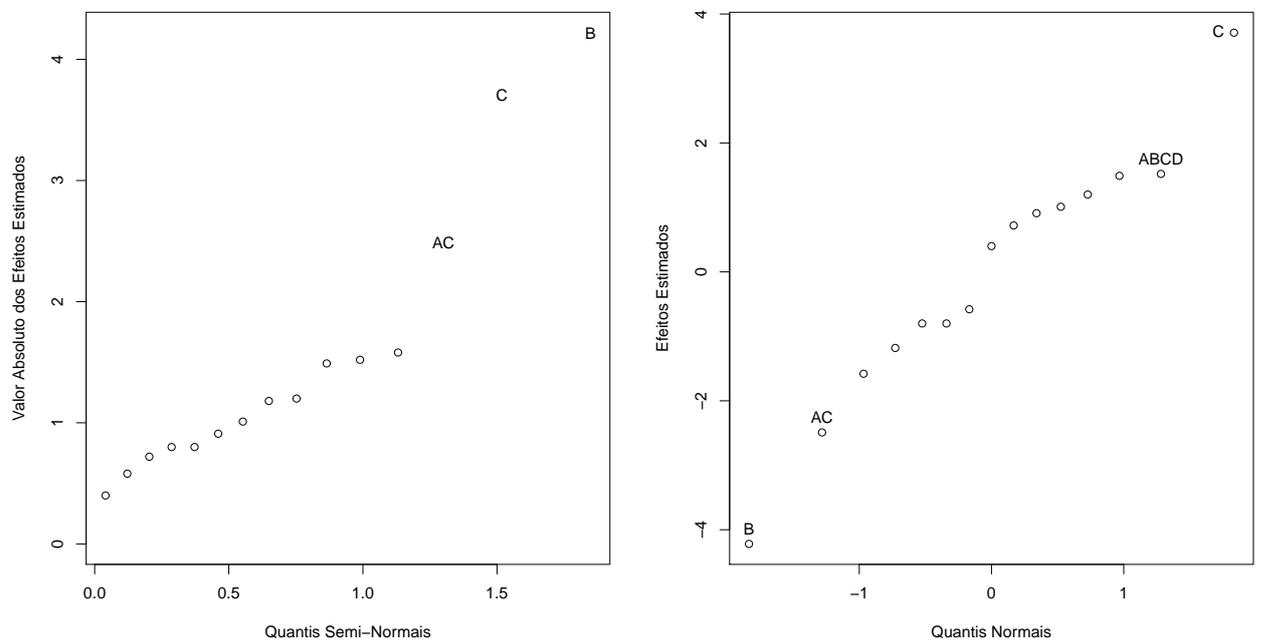


Figura 5.17: Gráficos de Probabilidade para o Exemplo 5.17.

Exemplo 5.18 (GOUPY, 2006 - exemplo 1): Trata-se de um fatorial 2^4 . O objetivo é detectar um valor discrepante nos dados, usando tanto o gráfico semi-normal quanto um teste desenvolvido pelo autor. Como se pode ver pelos gráficos da Figura 5.18, não há indicação da presença de valores discrepantes.

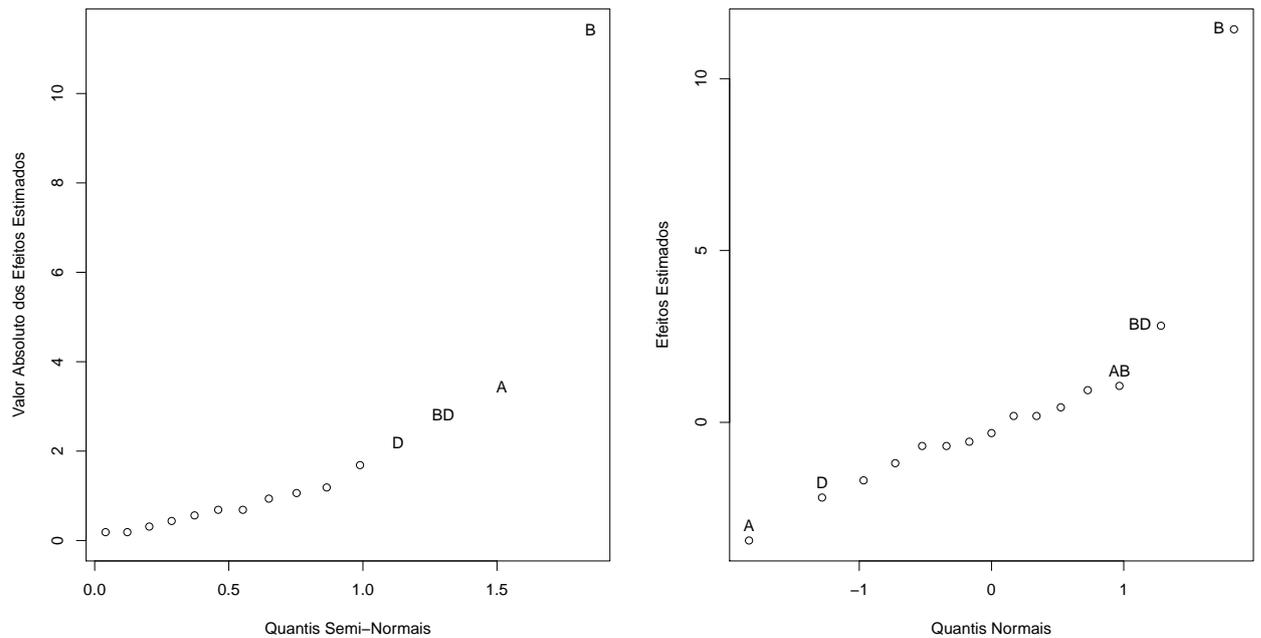


Figura 5.18: Gráficos de Probabilidade para o Exemplo 5.18.

Exemplo 5.19 (GOUPY, 2006 - exemplo 2): Desta feita tem-se um fatorial fracionado 2^{6-2} cujos geradores são: $E = ABC$; $F = BCD$. Pelos gráficos mostrados na Figura 5.19, é nítida a presença de valores discrepantes nos dados: o tamanho dos *gaps* é extremamente acentuado. Neste caso, há um valor discrepante de tamanho $10\sigma_{obs}$, conforme especificado no artigo em questão, em que σ_{obs} é o desvio padrão das observações sem o valor discrepante.

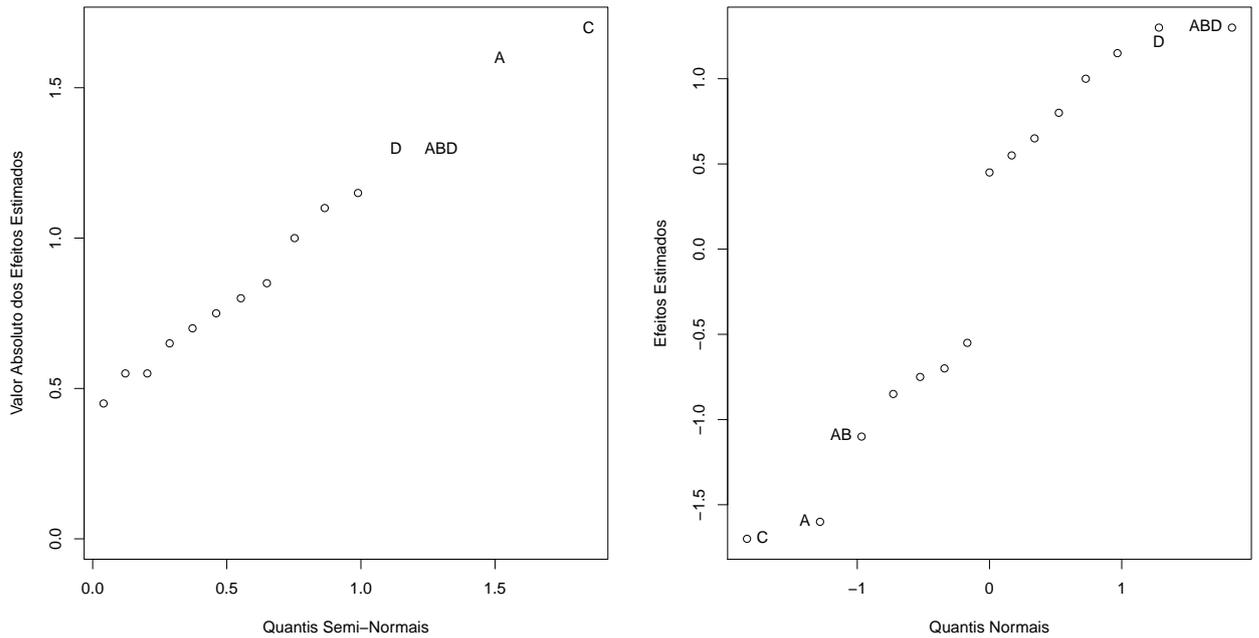


Figura 5.19: Gráficos de Probabilidade para o Exemplo 5.19.

Exemplo 5.20 (GOUPY, 2006 - exemplo 2): Para este exemplo, utiliza-se os dados do Exemplo 5.19 sem o valor discrepante, o que pode ser notado pela ausência das *gaps*. É uma ilustração da influência do valor discrepante na estimação dos efeitos. Percebe-se que até a mistura crítica (*critical mix*) $|A| \approx |D| \approx |AD|$, que neste caso pode ser claramente identificada em ambos os gráficos, é afetada pelo valor discrepante.

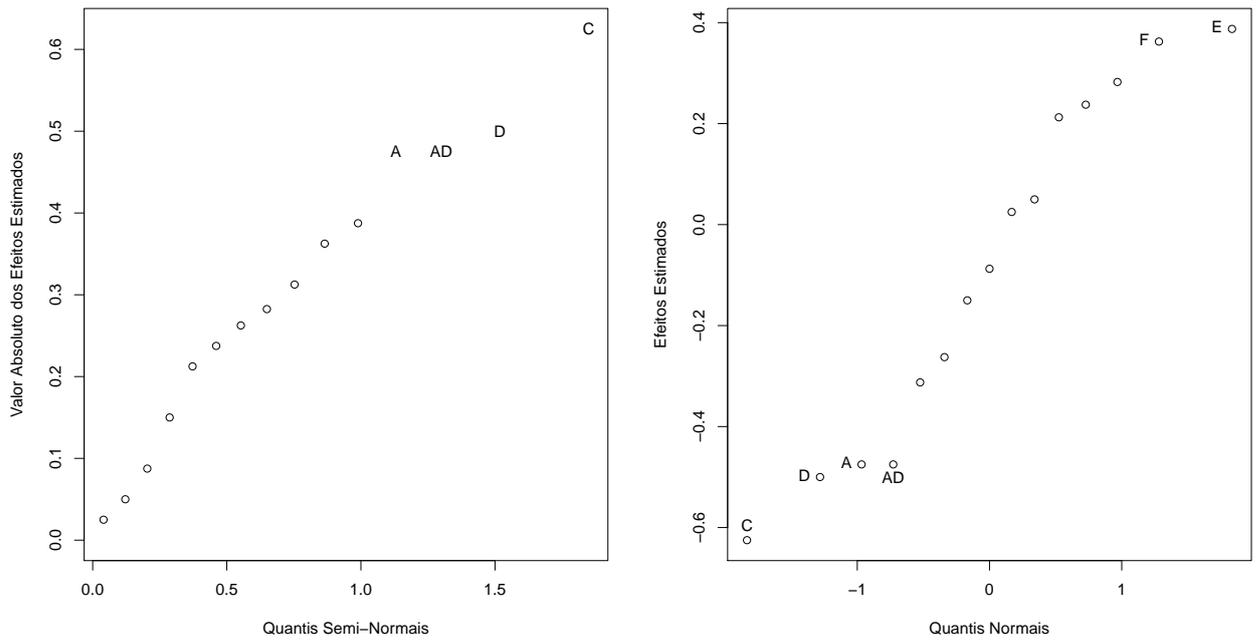


Figura 5.20: Gráficos de Probabilidade para o Exemplo 5.20.

Exemplo 5.21 (LAWSON, 2008 - exemplo 1): Trata-se de um fatorial 2^4 no qual o autor utiliza seu método proposto para identificar valores discrepantes nos dados, como também corrigi-lo, e por fim, avaliar a significância dos efeitos. O método trata de uma programação em SAS, que combina o teste desenvolvido por Daniel (1960) e por Box e Meyer (1986), que visa detectar valores discrepantes em experimentos fatoriais não replicados. A configuração dos *gaps* em ambos os gráficos é uma indicação da presença de valores discrepantes nos dados.

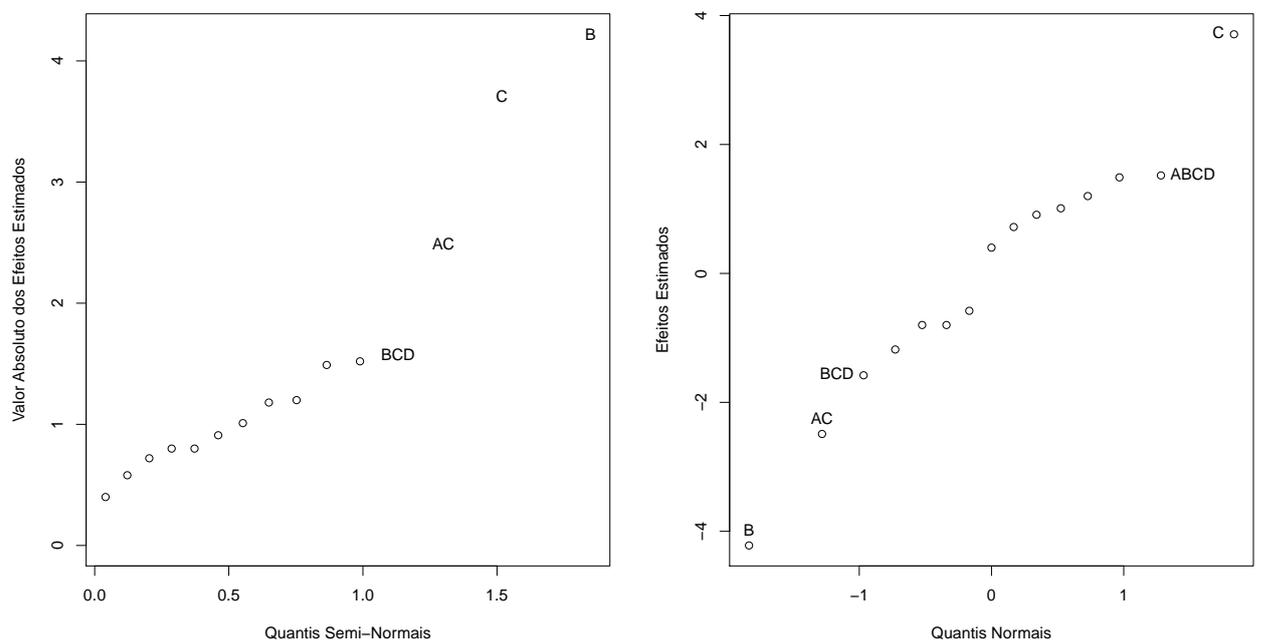


Figura 5.21: Gráficos de Probabilidade para o Exemplo 5.21.

Exemplo 5.22 (LAWSON, 2008 - exemplo 2): Trata-se de um fatorial fracionado 2^{6-3} com a relação definidora dada por $I = ABD = ACE = BCF = BCDE = ACDF = ABEF = DEF$. O teste proposto pelos autores aponta a presença de um valor discrepante. Os *gaps* observados nos gráficos da Figura 5.22, embora consideravelmente pequenos, indicam esta mesma conclusão.

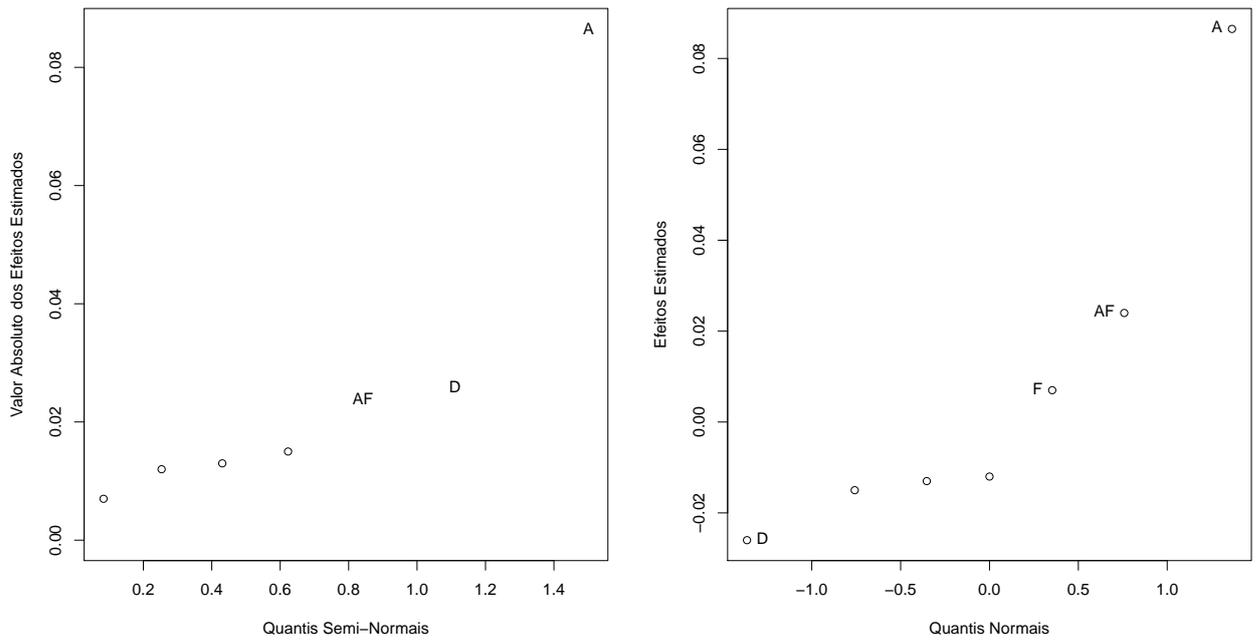


Figura 5.22: Gráficos de Probabilidade para o Exemplo 5.22.

Exemplo 5.23 (LAWSON, 2008 - exemplo 3): Trata-se de um fatorial fracionado 2^{5-1} . Embora o autor tenha detectado um valor discrepante nos dados pelo uso de seu método; tal conclusão não é óbvia pela análise dos gráficos da Figura 5.23.

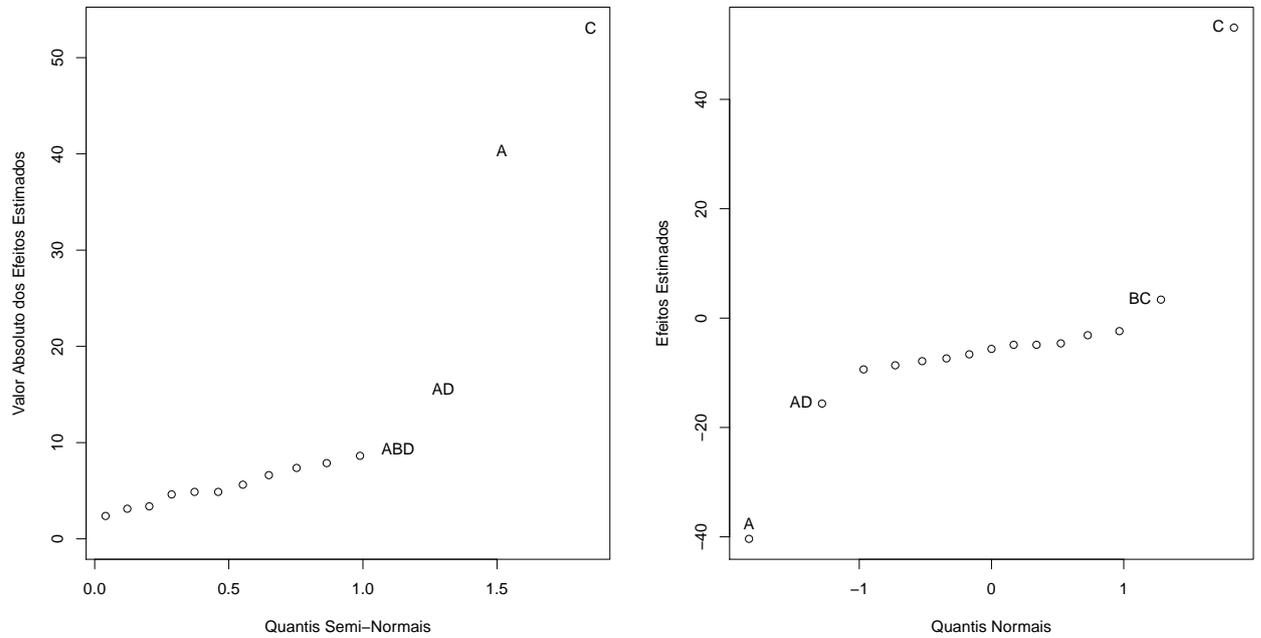


Figura 5.23: Gráficos de Probabilidade para o Exemplo 5.23.

Análise dos Resultados

Sobre a habilidade dos gráficos em detectar a presença de valores discrepantes, de acordo com os exemplos aqui apresentados bem como aqueles observados nos artigos selecionados, nota-se que o gráfico normal apresenta um padrão visual mais claramente identificável do que o gráfico semi-normal. Nas situações em que a magnitude do valor discrepante era considerada grande, tais como $10\sigma_{obs}$, ambos os gráficos mostraram-se igualmente úteis. Já naquelas em que o valor discrepante não era tão grande assim, a identificação de sua presença através do gráfico normal foi mais evidente. Parece ser mais perceptível visualizar a abertura (*gap*) em torno da origem, que é o padrão do gráfico normal quando um valor discrepante está presente; do que perceber um pequeno deslocamento da origem, característico do gráfico semi-normal nesta situação.

Vale ressaltar ainda que a presença de um valor discrepante nem sempre será identificada ou percebida pelo uso das técnicas gráficas. Desta forma, o fato de um gráfico de probabilidade normal ou semi-normal não apresentar a forma esperada produzida pela presença de um valor discrepante nas observações, não significa necessariamente que ele não esteja lá, influenciando os dados (HUND; MASSART; SMEYERS-VERBEKE, 2002). Nesta situação, se o pesquisador tiver suspeitas acerca da presença de observações aberrantes, é indicado o uso de testes mais formais, tais como os propostos por Daniel (1960), Goupy (2006) e Lawson (2008).

Alguns comentários são feitos acerca da presença de valores discrepantes nos planos fatoriais não replicados. Por exemplo, Daniel (1959) explica que, quando dois valores discrepantes têm a mesma magnitude, eles agem conjuntamente para “inflacionar” metade dos efeitos; se os dois tiverem magnitudes diferentes, então, todos os efeitos serão afetados, metade deles pela diferença da magnitude dos valores discrepantes, e a outra metade pela soma. Uma vantagem do gráfico semi-normal é que sua habilidade de detectar valores discrepantes nos dados não depende de quais efeitos são julgados significativos (DANIEL, 1960), uma vez que o tamanho do *gap* depende do tamanho do valor discrepante presente. Ao usarem o gráfico normal para detectar tais valores, Hund, Massart e Smeyers-Verbeke (2002) apresentam algumas importantes conclusões: o gráfico normal é útil para detectar apenas grandes valores discrepantes nos dados (de magnitude $10\sigma_{obs}$); um valor discrepante que afeta todos os efeitos do mesmo modo (o

que ocorre nos casos de experimentos não balanceados) não será detectado no gráfico normal; o gráfico normal não mostrou-se um método satisfatório para identificar a presença de dois valores discrepantes nos dados, a sugestão é utilizá-lo para a identificação de um único valor discrepante, e que este tenha grande magnitude. Segundo Goupy (2006), os métodos gráficos são testes de detecção visual e, portanto, subjetivos, geralmente não são muito sensíveis e não indicam qual a resposta aberrante. Lawson (2008) ainda comenta que se o número de provas num fatorial não replicado for maior que 16 (tais como 32 e 64), tal experimento é bastante robusto a valores discrepantes. Desta forma, os efeitos estimados não serão grandemente viciados, e poderão ser identificados usando o gráfico normal ou o gráfico semi-normal. Entretanto, se $n = 8$ ou 16 , todos os efeitos calculados serão viciados, tornando difícil identificar aqueles que são significativos. Tais comentários serão levados em consideração na construção dos cenários das simulações.

5.2.3 Identificação de Restrições na Aleatorização (*Inadvertent Plot-Splitting*)

É comum em planejamento de experimentos a presença de fatores que são mais caros, consomem mais tempo ou são mais difíceis de mudar que outros. Isto torna inviável a execução do experimento sob a forma de um completamente aleatorizado, fazendo com que o investigador economize seus esforços por minimizar os custos dos fatores difíceis de mudar, criando então uma restrição na aleatorização (KULAHCI; BISGAARD, 2005). Esta situação dá origem aos experimentos *split-plot*, comumente conhecidos como experimentos em parcelas subdivididas.

As primeiras referências aos experimentos em parcelas subdivididas foram aplicadas no contexto da agricultura, e apareceram na literatura por volta da década de 30, com os trabalhos de Fisher, que conduziu experimentos agrícolas com fertilizantes e culturas (ARVIDSSON *et all*, 2001). O lote inteiro (*whole plot*) de terra era subdividido ou parcelado (*split*) em sub-lotes (*sub-plots*) ou sub-unidades, para que os diferentes fertilizantes fossem aplicados. Em seguida, os diferentes tipos de culturas poderiam então ser cultivados nestas diferentes parcelas. Daí a origem do nome *split-plot* ou parcelas subdivididas. Nestes planejamentos, os fatores difíceis de mudar são chamados de *whole-plot factors* ou fatores da parcela, e os mais fáceis de mudar de *sub-plot factors*

ou fatores da subparcela.

Os planos em parcelas subdivididas apresentam restrições na aleatorização do experimento, sendo uma opção prática de planejamento nos casos em que não é possível ou razoável realizar o experimento numa ordem completamente aleatória e irrestrita. A idéia de restrição na aleatorização é facilmente entendida pela explicação de Bingham e Sitter (1999, p.62): “*Suppose we wish to run an experiment with k factors, each at two levels. A natural choice is a 2^k factorial experiment in which it is required that the experimental runs be performed in a completely random order. Suppose, however, that it is very expensive or difficult to change the levels for some of the factors, say k_1 of them. To save costs or time, we could instead randomly choose one of the factor-level settings of these k_1 hard-to-change factors and then run all of the level combinations of the remaining k_2 factors in a random order while holding the k_1 factors fixed. This would then be repeated for each level combination of the k_1 factors. The design matrix for this experimental setup is identical to a 2^k factorial design. It is only the randomization structure that is different. This restriction in randomization results in what is commonly known as a split-plot structure in which the k_1 and k_2 factors are called whole-plot (WP) and sub-plot (SP) factors, respectively.*”

Quando um experimento fatorial não replicado tem restrições na aleatorização, Daniel (1959) sugeriu que fossem construídos dois gráficos semi-normais, um para os efeitos da parcela (*whole plot*) e outro para os efeitos da subparcela (*sub-plot*). A razão para isso é comentada no artigo: quando o gráfico é construído com todos os efeitos juntos, os dois tipos de erro (o erro associado à parcela e o erro associado à subparcela) estão misturados, o que acaba afetando a forma do gráfico e, conseqüentemente, o julgamento dos efeitos. Segundo Daniel, a forma curvada com que os pontos são dispostos no gráfico semi-normal é um indicativo de que há restrições na aleatorização do experimento, o que serve de alerta para o investigador sobre a presença de um plano em parcelas subdivididas. É um erro comum apresentar um experimento como completamente aleatorizado quando, na verdade, se trata de um planejamento em parcelas subdivididas. A proposta de Daniel é que o gráfico semi-normal identifique esta situação e oriente o investigador a corrigi-la. A esta situação ele se referiu como *inadvertent plot-splitting* ou parcelas subdivididas inadvertidas, isto é, embora tenha realizado o experimento em parcelas subdivididas, o pesquisador inadvertidamente o analisa de

outra forma, por não reconhecer a finalidade da aleatorização.

Søren Bisgaard, Howard Fuller e Ernesto Barrios (1995) aplicam para o gráfico normal o mesmo uso que Daniel apontou para o gráfico semi-normal no que diz respeito a situação de parcelas subdivididas inadvertidas comentada no parágrafo anterior. A sugestão dos autores é calcular grupos de efeitos que têm variâncias diferentes, e então construir dois gráficos normais, um para cada grupo, a fim de avaliar a significância dos efeitos. Isto é útil para não esconder efeitos significativos, nem tampouco para julgar como ativo um efeito que na verdade é inativo. Quando é feito apenas um único gráfico para todos os efeitos do experimento em parcelas subdivididas, a distribuição misturada tem caudas mais pesadas fazendo com que a curva traçada no gráfico normal tenha uma forma se S^1 .

Por serem tão comumente usados na pesquisa industrial, por serem os mais convincentes (DANIEL, 1976) e em muitos casos, os únicos possíveis (DANIEL, 1976); os planejamentos em parcelas subdivididas não poderiam deixar de ser considerados neste trabalho. Como há indicações das duas técnicas gráficas para o uso em tais experimentos, esta seção aborda a comparação entre o gráfico normal e o semi-normal em cenários que apresentam experimentos com restrições na aleatorização; a fim de verificar qual deles apresenta a melhor visualização que ajude a identificar mais prontamente o caso das parcelas subdivididas inadvertidas: se a forma curvada do gráfico semi-normal (DANIEL, 1959) ou a forma de S do gráfico normal (BISGAARD; FULLER; BARRIOS, 1995).

¹Esta forma de S é notada quando o gráfico é construído com a seguinte configuração dos pontos: $(y_{(j)}, \Phi^{-1}(\hat{p}_j))$. Na versão R 2.10.0, usada nesta dissertação, os pontos marcados são da forma $(\Phi^{-1}(\hat{p}_j), y_{(j)})$. Com esta configuração, no caso das parcelas subdivididas, a forma S não é observada.

Exemplo 5.24 (BISGAARD; FULLER; BARRIOS, 1995): Trata-se de um experimento fatorial 2^5 em parcelas subdivididas. Neste artigo os autores discutem o uso do gráfico normal para detectar estas restrições na aleatorização. Por meio deste exemplo, eles comentam a forma S padrão do gráfico normal na presença dos planos em parcelas subdivididas. Discutem também como separar os efeitos da parcela e os da subparcela para construir os gráficos normais separados. A Figura 5.24 mostra os gráficos com os pontos na forma $(y_{(j)}, \Phi^{-1}(\hat{p}_j))$. É esta disposição dos pontos que forma o padrão S do gráfico normal e o padrão curvo descrito por Daniel (1959) no gráfico semi-normal, para o caso considerado.

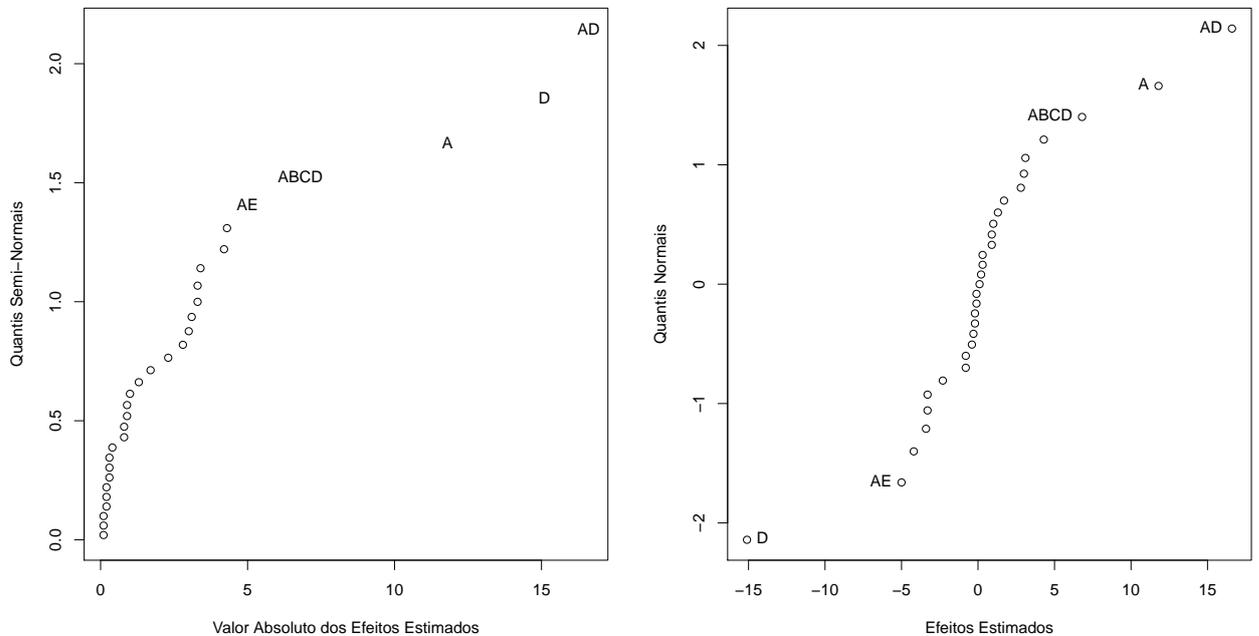


Figura 5.24: Gráficos de Probabilidade para o Exemplo 5.24.

Exemplo 5.25 (BISGAARD; FULLER; BARRIOS, 1995): Neste exemplo, utiliza-se os mesmos dados do Exemplo 5.24, diferindo apenas na construção dos gráficos: os pontos marcados são da forma $(\Phi^{-1}(\hat{p}_j), y_{(j)})$, que é o padrão dos gráficos produzidos pela versão R 2.10.0. *A*, *B*, *C* e *D* são os efeitos da parcela, e *E* é o efeito da subparcela.

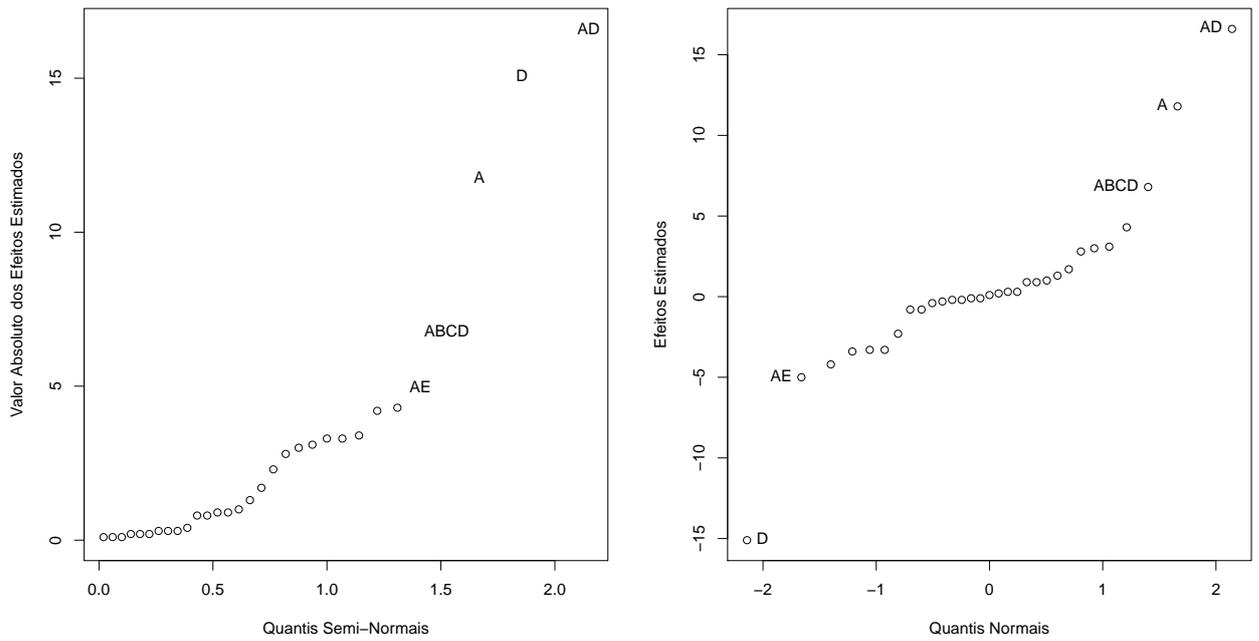


Figura 5.25: Gráficos de Probabilidade para o Exemplo 5.25.

Exemplo 5.26 (BISGAARD; FULLER; BARRIOS, 1995): Esta é apenas uma ilustração da sugestão dada pelos autores bem como por Daniel (1959) no caso da análise gráfica dos planos com parcelas subdivididas: o ideal, nesta situação, é construir dois gráficos separadamente, um para os efeitos da parcela e outro para os efeitos da subparcela. Tal procedimento permite ao investigador realizar um julgamento mais claro sobre quais dos efeitos são realmente significativos. Quando se constrói os gráficos com todos os efeitos misturados, apenas os efeitos A , D e AD são julgados significativos, como se pode ver pela Figura 5.24 e pela Figura 5.25. Mas quando se contrói o gráfico dos efeitos separados, pode-se perceber o experimento mais claramente: A , D e AD são os efeitos significativos devidos à parcela, e os efeitos E e AE são significativos devidos à subparcela. Sendo assim, o experimento em questão tem cinco efeitos significativos, e não apenas três. Outro aspecto importante que pode ser notado pelos gráficos é que o erro associado à parcela é maior que o erro associado à subparcela: basta olhar a inclinação das retas nos dois casos. A Figura 5.26 e a Figura 5.27 ilustram estas conclusões.

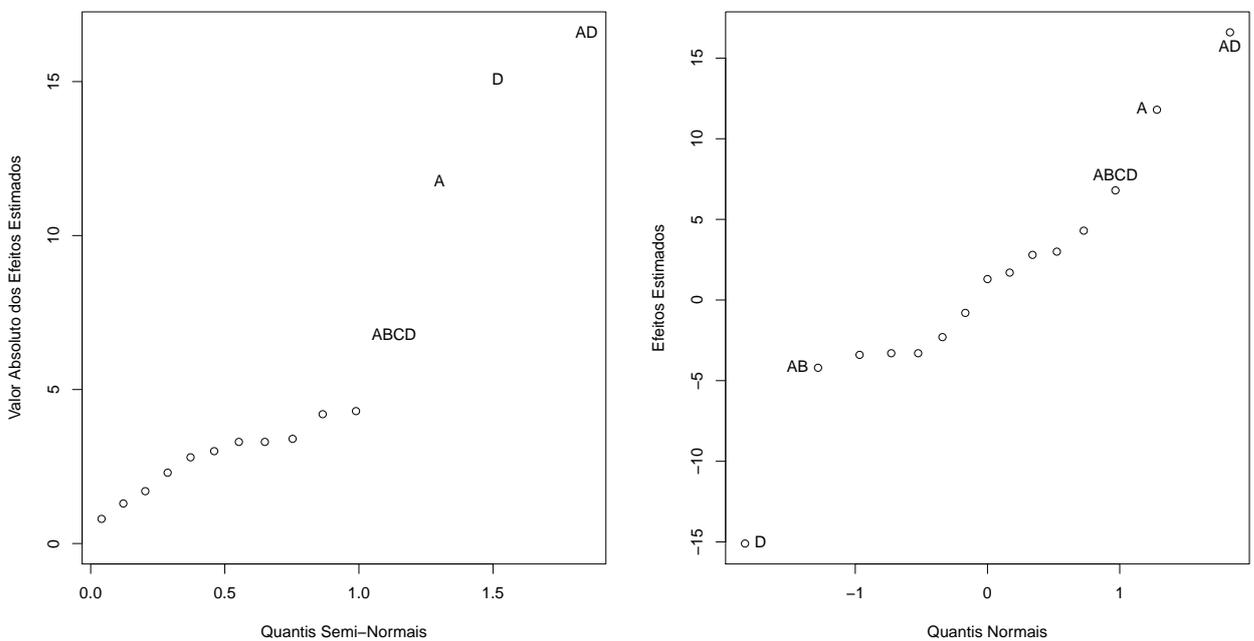


Figura 5.26: Gráficos de Probabilidade para o Exemplo 5.26. Gráficos com os efeitos da parcela.

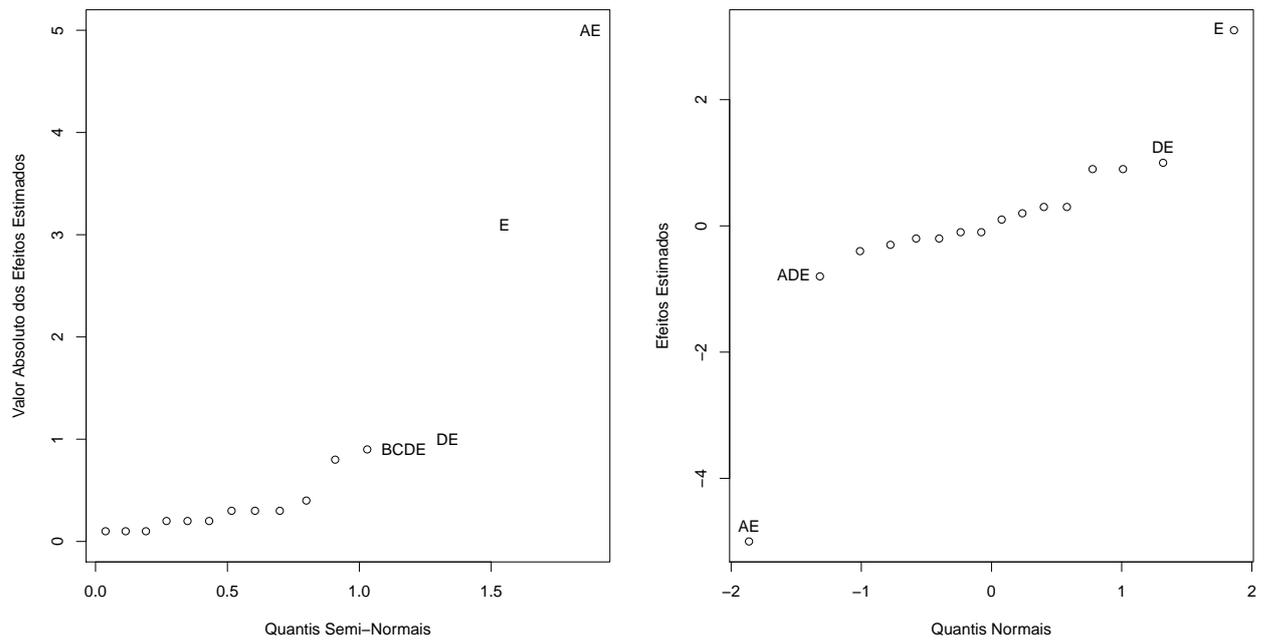


Figura 5.27: Gráficos de Probabilidade para o Exemplo 5.26. Gráficos com os efeitos da subparcela.

Exemplo 5.27 (MONTGOMERY; KULAHCI; ALMIMI, 2007): Trata-se de um fatorial 2^5 . A Figura 5.28 mostra os gráficos com os pontos na forma $(y_{(j)}, \Phi^{-1}(\hat{p}_j))$. A nítida forma de S no gráfico normal é um indicativo da ocorrência de parcelas subdivididas. Neste exemplo, A , B e C são os efeitos da parcela, e D e E os efeitos da subparcela.

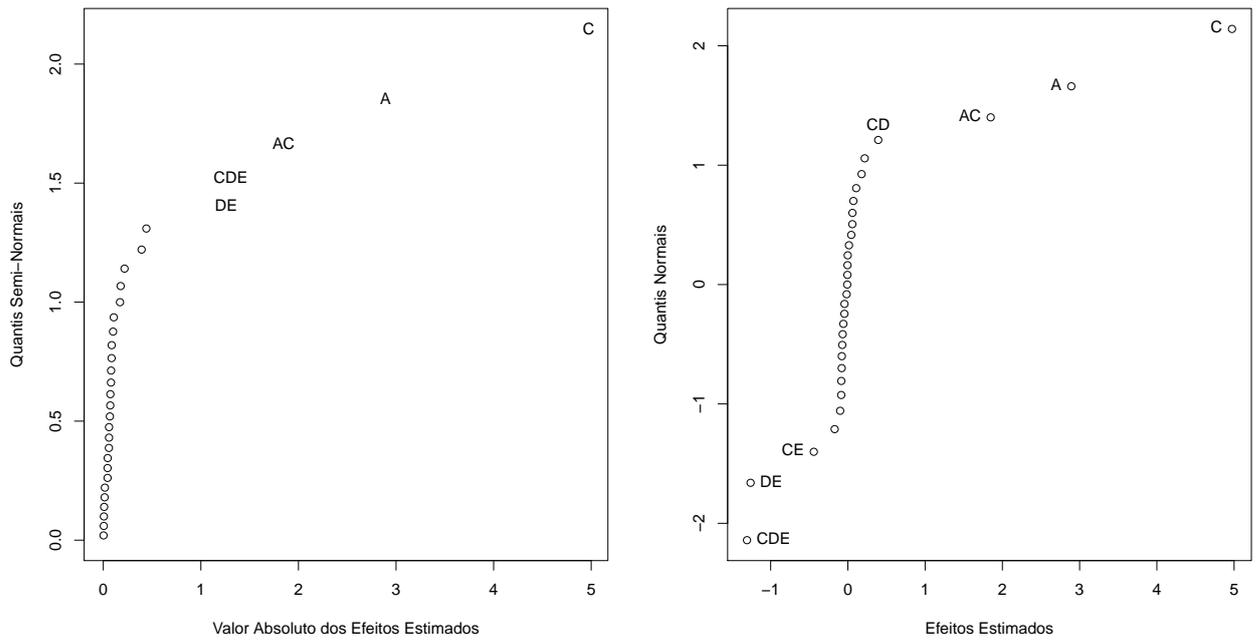


Figura 5.28: Gráficos de Probabilidade para o Exemplo 5.27.

Exemplo 5.28 (MONTGOMERY; KULAHCI; ALMIMI, 2007): Utiliza-se os mesmos dados do Exemplo 5.27 para ilustrar os gráficos de probabilidade construídos com o formato do R 2.10.0.

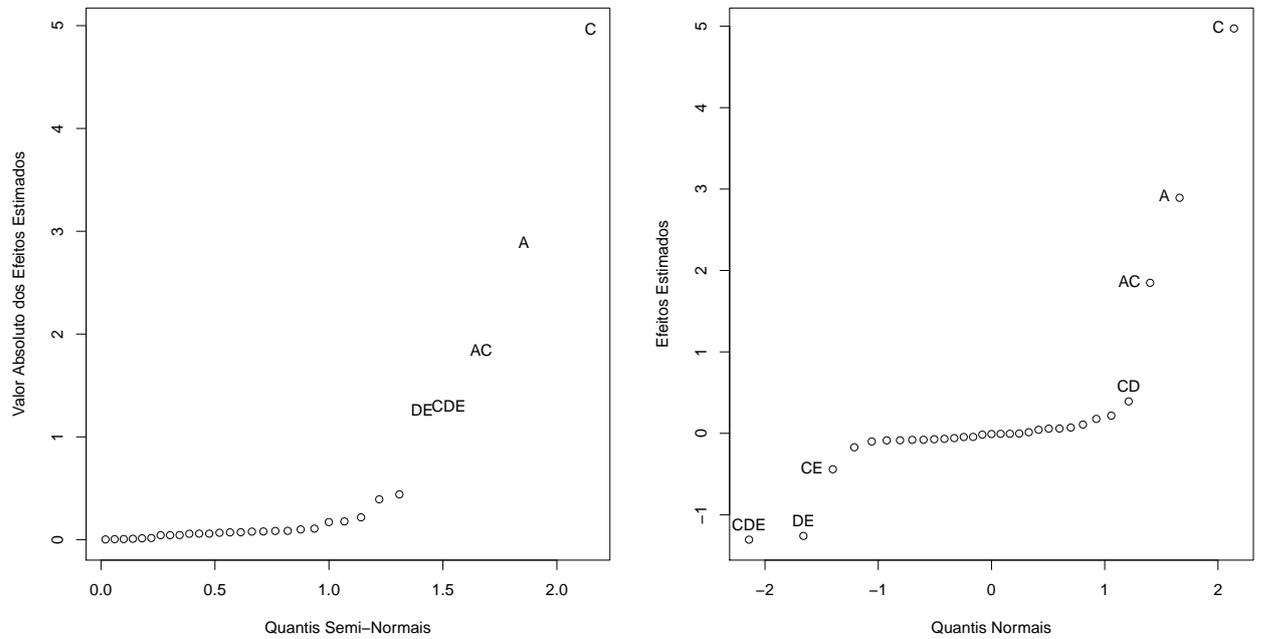


Figura 5.29: Gráficos de Probabilidade para o Exemplo 5.28.

Análise dos Resultados

Sobre o aspecto da habilidade dos gráficos de probabilidade em identificar restrições na aleatorização, pela análise dos exemplos aqui apresentados bem como daqueles observados nos poucos artigos encontrados que tratam do assunto, o gráfico normal mostra-se a ferramenta mais útil e de melhor identificação visual em relação ao gráfico semi-normal. A sua característica forma de S , padrão para os casos dos experimentos com parcelas subdivididas, mostrou-se evidente nos cenários apresentados, tornando mais clara a detecção desta restrição na aleatorização. No caso do gráfico semi-normal, não foi percebida uma forma padrão para esta situação que ajudasse a identificar o plano em parcelas subdivididas conforme Daniel (1959) afirmou. Embora se note, é verdade, uma certa curvatura no gráfico semi-normal nestes casos, não ficou muito claro que tipo de forma este gráfico assume nestas circunstâncias, ao contrário do que já se espera do gráfico normal.

Vale destacar ainda a situação em que os gráficos são construídos com a disposição dos pontos padrão do R 2.10.0. Tendo por base apenas os exemplos acima apresentados, nota-se que nesta configuração é mais difícil encontrar uma forma, em ambos os gráficos, que possa auxiliar na identificação visual da ocorrência de restrições na aleatorização. Isso talvez seja uma indicação de que, no caso das parcelas subdivididas, construir os gráficos na forma $(y_{(j)}, \Phi^{-1}(\hat{p}_j))$ seja uma escolha mais apropriada para a identificação gráfica desta situação.

5.3 Cenários para as Simulações

Esta seção trata da parte principal acerca da metodologia adotada para a consecução do objetivo central deste trabalho: a construção de simulações de cenários experimentais para realizar um estudo comparativo dos gráficos normal e semi-normal no contexto da análise dos planos fatoriais não replicados. O levantamento bibliográfico forneceu indicações para a metodologia adotada na construção destes cenários que, semelhante ao realizado na Seção 5.2, encontram-se divididos em três grupos, em que cada grupo representa o aspecto sob o qual o desempenho dos gráficos é comparado:

1. Julgamento dos Efeitos. Para comparar a habilidade dos gráficos em avaliar a significância dos efeitos, os seguintes fatores são levados em consideração na construção dos cenários das simulações:

- **Número de provas.** Concentra-se nos fatoriais cujas provas são $2^3 = 8$, $2^4 = 16$, $2^5 = 32$ e $2^6 = 64$ (DANIEL, 1976; DONG, 1993; LAWSON, 2008).
- **Porcentagem de efeitos ativos.** Com relação à quantidade de efeitos ativos presentes nos cenários, usa-se os seguintes níveis: 0%, 10%, 20% e 40%. Tal escolha é feita pois, de acordo com Daniel (1967), o gráfico semi-normal é útil em detectar significância quando até 20% dos efeitos são ativos. Além desse valor, sua habilidade é comprometida. A razão disso é que os gráficos de probabilidade são construídos sob a suposição do princípio da esparsidade dos efeitos (BOX; MEYER, 1986; DONG, 1993).
- **Tamanho dos efeitos.** Os efeitos ativos usados nas simulações têm os seguintes tamanhos: $\pm 3\sigma_{ef}$, $\pm 5\sigma_{ef}$ e $\pm 10\sigma_{ef}$, em que σ_{ef} é o desvio padrão dos efeitos não ativos. A escolha destes tamanhos tem como intuito comparar a habilidade dos gráficos em avaliar pequenos, médios e grandes efeitos, respectivamente.
- Nesta fase inicial das simulações (sem a presença de valores discrepantes), o **método objetivo** utilizado para a comparação dos gráficos é a distância vertical dos efeitos significativos aos pontos marcados sobre a reta $y = \sigma_{ef} \cdot G^{-1}(p_j)$. Tal expressão representa a reta traçada nos gráficos de probabilidade no caso em que nenhum efeito é significativo. Por tal critério de comparação, o “melhor” gráfico é aquele que apresenta os maiores valores para esta distância, denotando que a identificação visual dos efeitos ativos, em tal gráfico, é mais clara, uma vez que estes se apresentam mais nitidamente distantes da reta.

2. Detecção de valores discrepantes (*outliers*) nos dados. A partir das respostas simuladas obtidas sob as condições da fase anterior, o passo seguinte é criar cenários na presença de valores discrepantes, considerando os seguintes fatores:

- **Número de valores discrepantes.** Com relação ao número de valores discrepantes, consideram-se dois casos (DANIEL, 1959; LAWSON, 2008): utilizam-se

cenários com um valor discrepante presente, bem como com dois valores discrepantes nos dados.

- **Tamanho dos valores discrepantes.** Os valores discrepantes usados nas simulações têm os seguintes tamanhos: $\pm 3\sigma_{obs}$, $\pm 5\sigma_{obs}$ e $\pm 10\sigma_{obs}$, em que σ_{obs} é o desvio padrão das observações (HUND; MASSART; SMEYERS-VERBEKE, 2002; LAWSON, 2008).
- **Posição do valor discrepante.** A influência da posição do valor discrepante também é verificada, sob as seguintes características: no tratamento em que todos os fatores estão no nível “+”; no tratamento em que todos os fatores estão no nível “-”; num outro tratamento qualquer.
- Nos cenários em que dois valores discrepantes estão presentes, são consideradas duas situações: (1) os dois de mesma magnitude; (2) os dois de magnitudes diferentes.
- O **método objetivo** utilizado para a comparação gráfica é o tamanho dos *gaps*. A técnica gráfica terá o melhor desempenho em detectar a presença de valores discrepantes nos dados quanto maior for o tamanho do *gap* formado em tal gráfico, vindo assim a facilitar a identificação visual de tal situação.

3. Identificação de Restrições na Aleatorização (*Inadvertent Plot-Splitting*).

Embora esta seja uma importante situação em que os gráficos de probabilidade são utilizados para a análise dos experimentos fatoriais não replicados, nenhum método objetivo para a comparação dos gráficos em tais situações é usado nas simulações aqui apresentadas. Motivo: embora algumas técnicas e métodos tenham sido elaborados e testados durante o andamento desta dissertação, ainda não se obteve êxito em encontrar algum que pudesse ser útil para este estudo experimental comparativo.

Inicialmente é avaliada a influência do número N de simulações usadas para cada cenário. Para tal avaliação, usou-se como valores $N = 5000$ e $N = 10000$. Nas figuras que se seguem foram escolhidos, apenas como ilustração para os gráficos apresentados, o efeito A e o tratamento ab , visando resumir as informações dos resultados, análogos para todos os outros efeitos. Nos cenários em que o efeito A é ativo, o tamanho $10\sigma_{ef}$ usado para A varia de acordo com o número de provas do experimento, da seguinte

forma²: para os experimentos com 8 provas temos que $ef_A = 7,1$; com 16 provas, $ef_A = 5$; para um experimento com 32 provas, $ef_A = 3,5$; e para 64 provas $ef_A = 2,5$. A função criada para a realização das simulações é a função `teste()`, que pode ser vista no Apêndice desta dissertação. Tal função realiza os dois métodos objetivos propostos neste estudo experimental para a comparação dos gráficos de probabilidade normal e semi-normal: o cálculo das distâncias dos efeitos à reta, bem como o cálculo do tamanho dos *gaps*.

Apesar do objetivo principal desta fase inicial ser verificar se há ou não grandes diferenças no número de simulações utilizadas, algumas características podem ser percebidas. Nos casos em que o efeito A não é ativo ($ef_A = 0$), em grande parte das simulações as estimativas do efeito A estão próximas de zero. Nota-se também que, quanto maior o número de provas (com destaque para os cenários com 32 e 64 provas), mais próximo de zero os resultados se encontram. Não se percebem grandes diferenças para $N = 5000$ e $N = 10000$. As Figuras 5.30 e 5.31 ilustram estes fatos.

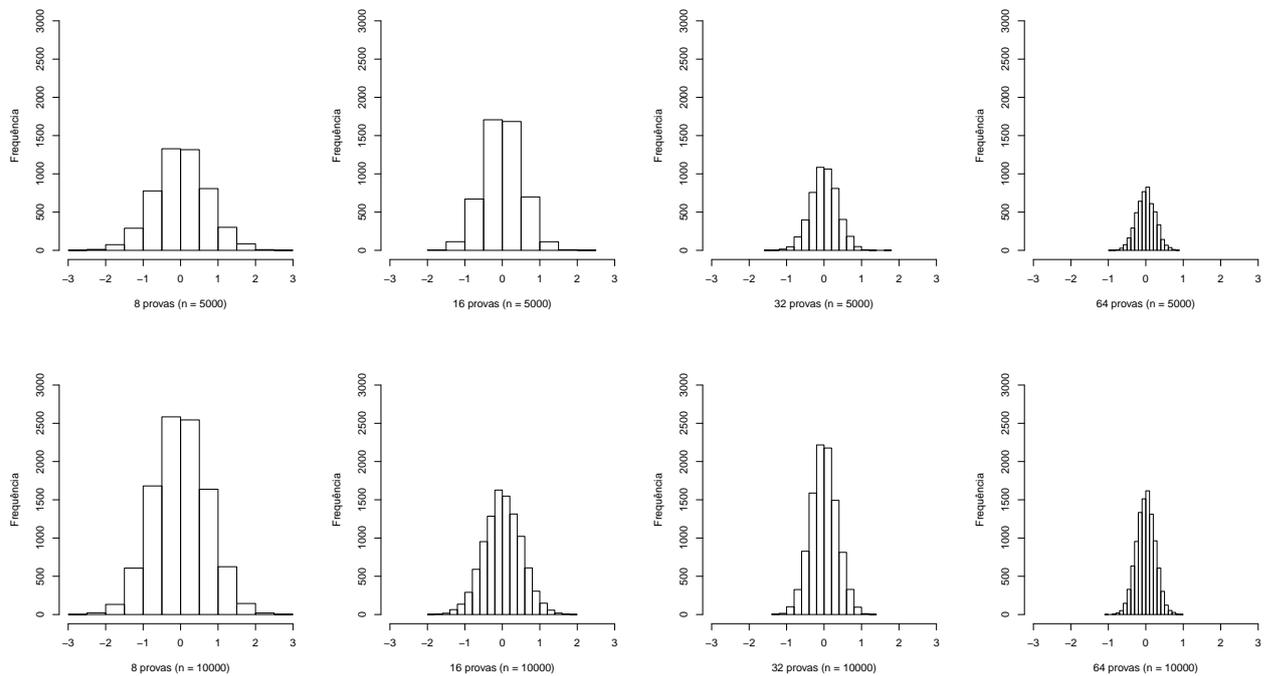


Figura 5.30: **Histogramas das Estimativas do Efeito A .** Para estes cenários, o efeito A não é significativo (seu valor é igual a zero), e não há valores discrepantes nos dados.

²Ver Seção 5.3.1.

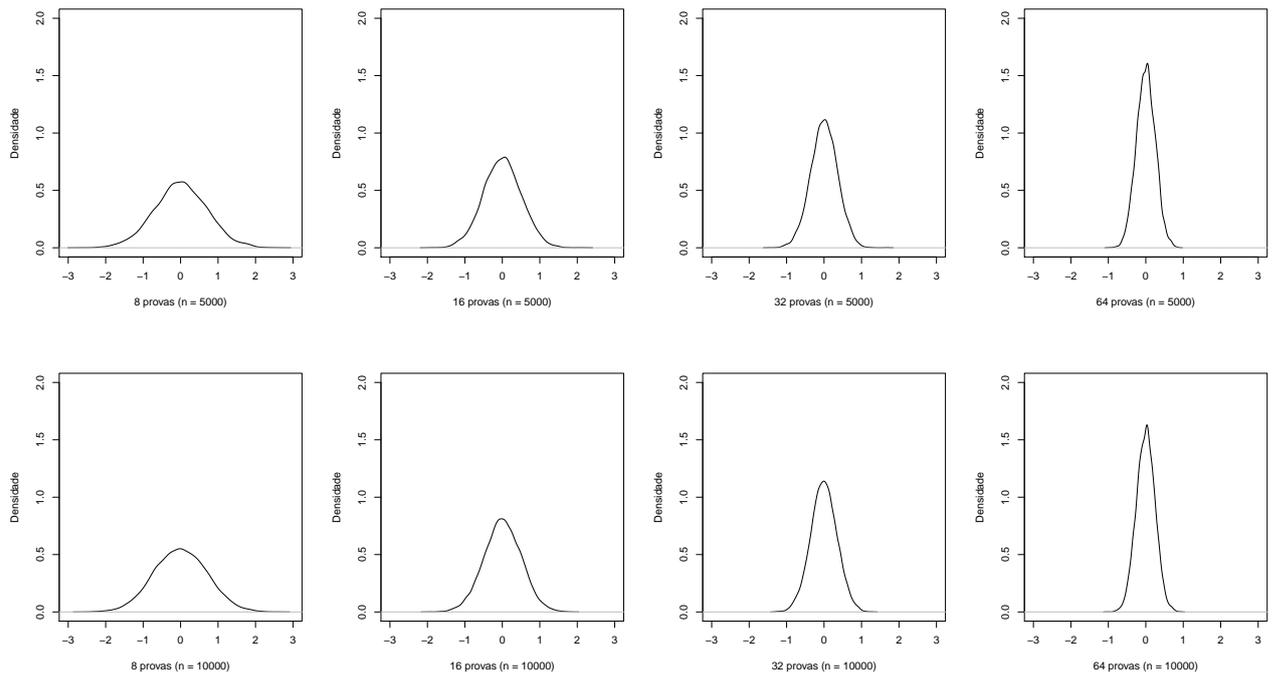


Figura 5.31: **Densidade das Estimativas do Efeito A .** Para estes cenários, o efeito A não é significativo (seu valor é igual a zero), e não há valores discrepantes nos dados.

Nos cenários em que o efeito A é ativo, nota-se, para ambos os casos ($N = 5000$ e $N = 10000$), que os resultados das simulações se aproximam do verdadeiro valor de A correspondente ao número de provas do experimento. Relembrando: $ef_A = 7,1$ para oito provas; $ef_A = 5$ para dezesseis provas; $ef_A = 3,5$ para trinta e duas provas; $ef_A = 2,5$ para sessenta e quatro provas. As Figuras 5.32 e 5.33 exemplificam estas afirmações.

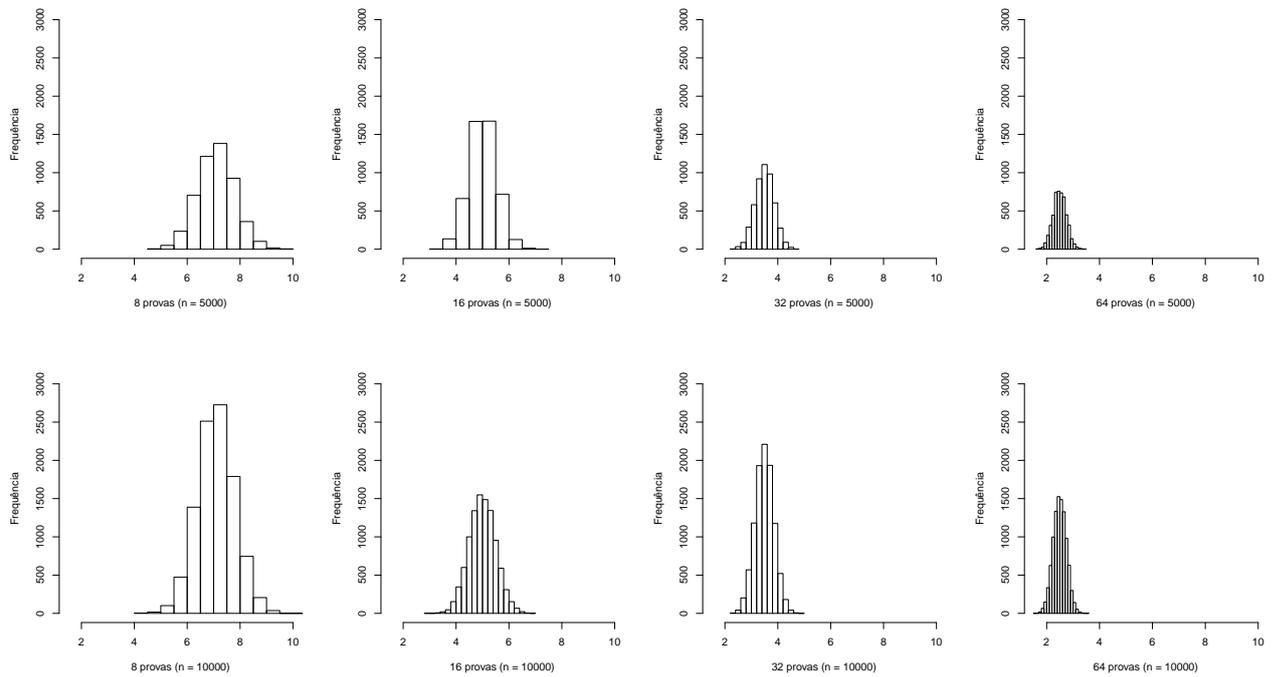


Figura 5.32: **Histogramas das Estimativas do Efeito A.** Para estes cenários, o efeito A é significativo, cujo tamanho é $10\sigma_{ef}$. Não há valores discrepantes nos dados.

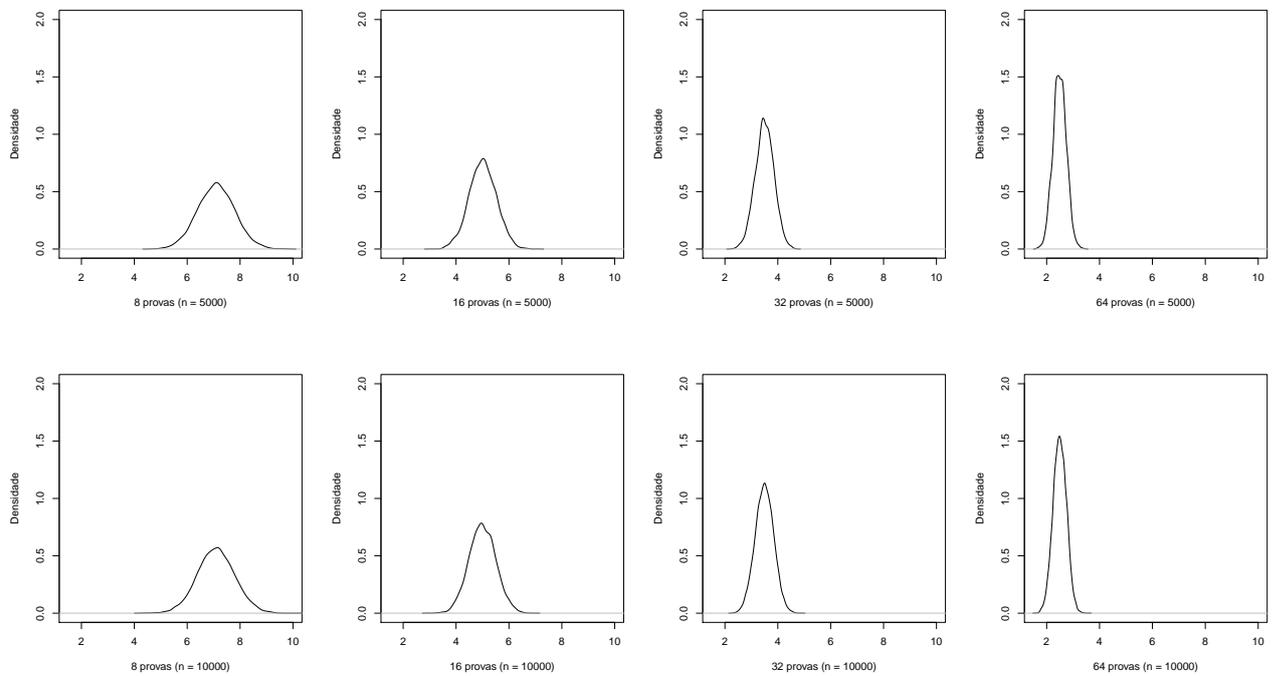


Figura 5.33: **Densidade das Estimativas do Efeito A.** Para estes cenários, o efeito A é significativo, cujo tamanho é $10\sigma_{ef}$. Não há valores discrepantes nos dados.

Nos cenários em que um valor discrepante de tamanho 10 está presente no tratamento ab , nota-se que as estimativas do efeito A aumentam em relação aos valores usuais, correspondentes ao número de provas considerado. Esta característica não é tão evidente nos experimentos com maior número de provas, em especial nos casos com 64 provas. Em tais casos, as estimativas do efeito A parecem não ser tão fortemente influenciadas pela presença do valor discrepante. Tanto para $N = 5000$ quanto para $N = 10000$ nota-se estas características semelhantes, conforme se pode ver nas Figuras 5.34 e 5.35.

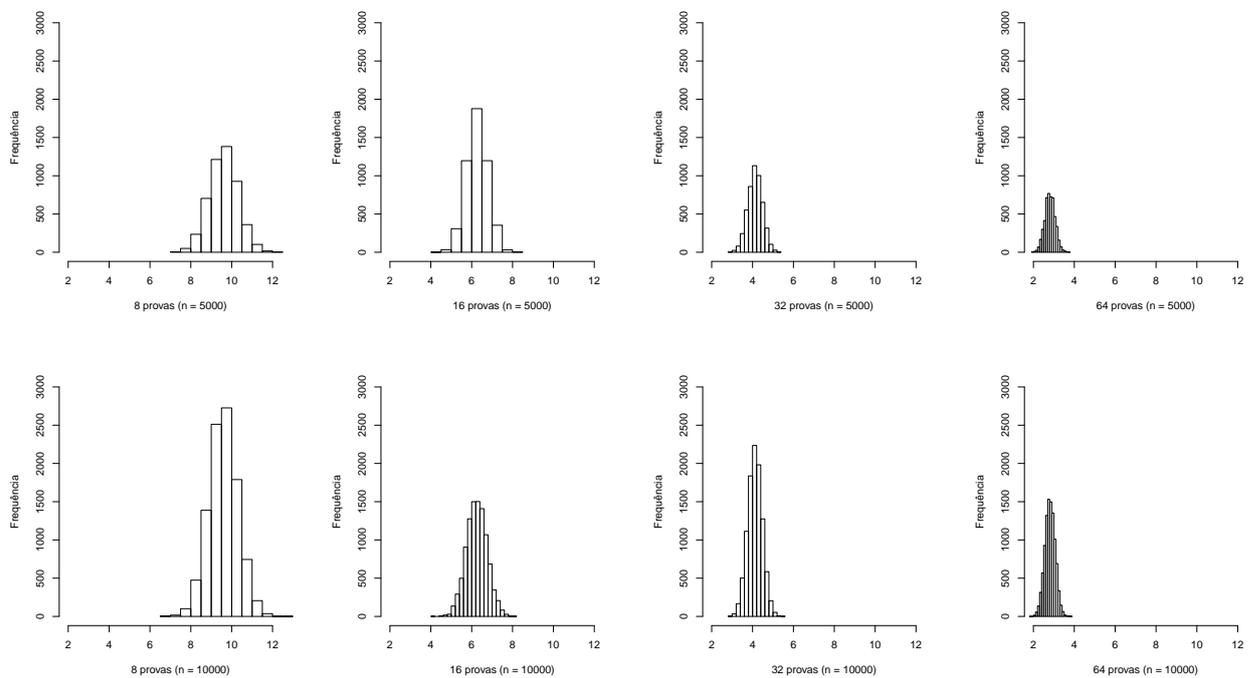


Figura 5.34: **Histogramas das Estimativas do Efeito A .** Para estes cenários, o efeito A é significativo, cujo tamanho é $10\sigma_{ef}$. Há um valor discrepante no tratamento ab , cujo tamanho é 10.

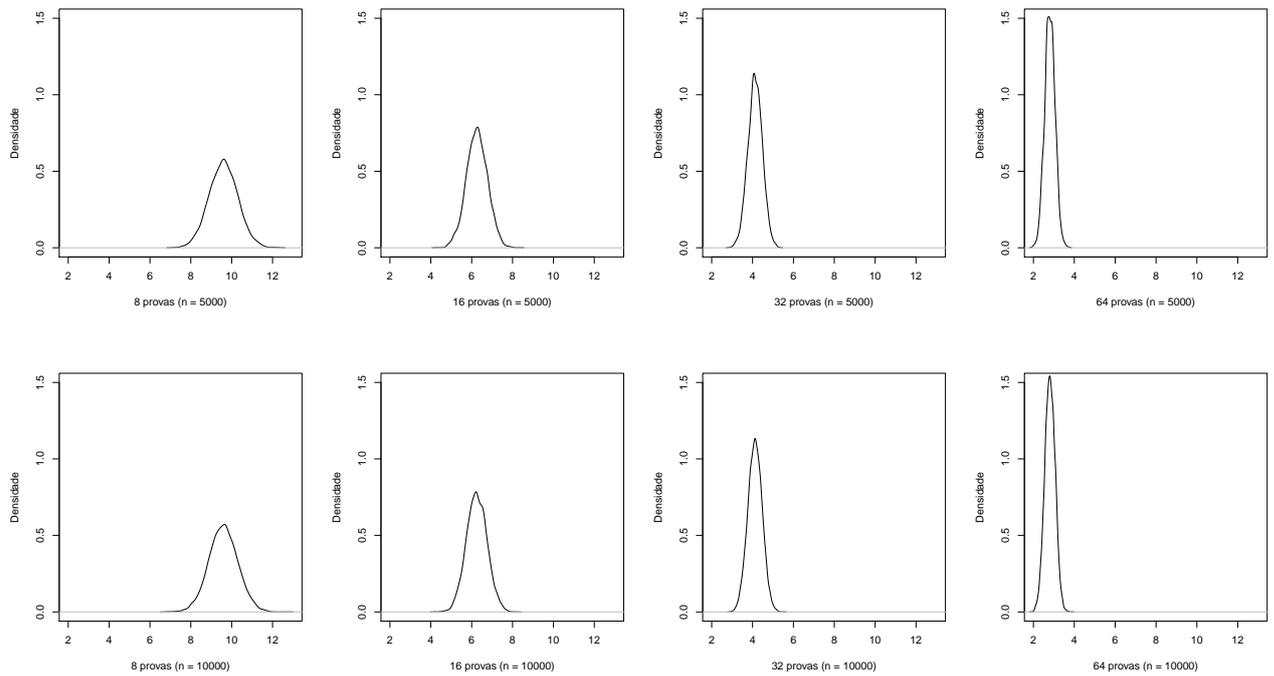


Figura 5.35: **Densidade das Estimativas do Efeito A.** Para estes cenários, o efeito A é significativo, cujo tamanho é $10\sigma_{ef}$. Há um valor discrepante no tratamento ab , cujo tamanho é 10.

Quando se analisa a habilidade dos gráficos em julgar a significância dos efeitos, nota-se, conforme se pode ver pelas Figuras 5.36, 5.37, 5.38 e 5.39, que as distâncias simuladas, calculadas para o efeito A , são consideravelmente maiores no gráfico semi-normal do que no gráfico normal. Tanto para os casos em que $N = 5000$ quanto para $N = 10000$, os resultados são análogos.

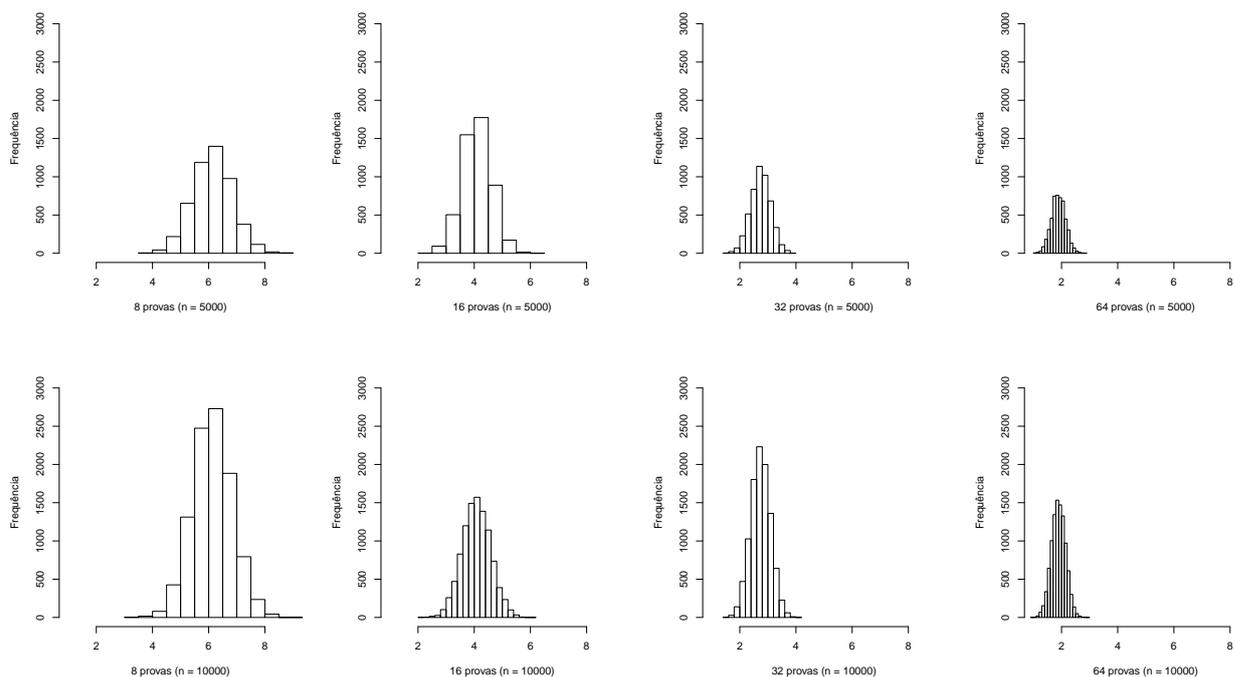


Figura 5.36: **Histogramas das Distâncias Calculadas para o Efeito A no Gráfico Normal.** Para estes cenários, o efeito A é significativo, cujo tamanho é $10\sigma_{ef}$. Os histogramas medem a distância do efeito A à reta $y = \sigma_{ef} \cdot G^{-1}(p_j)$.

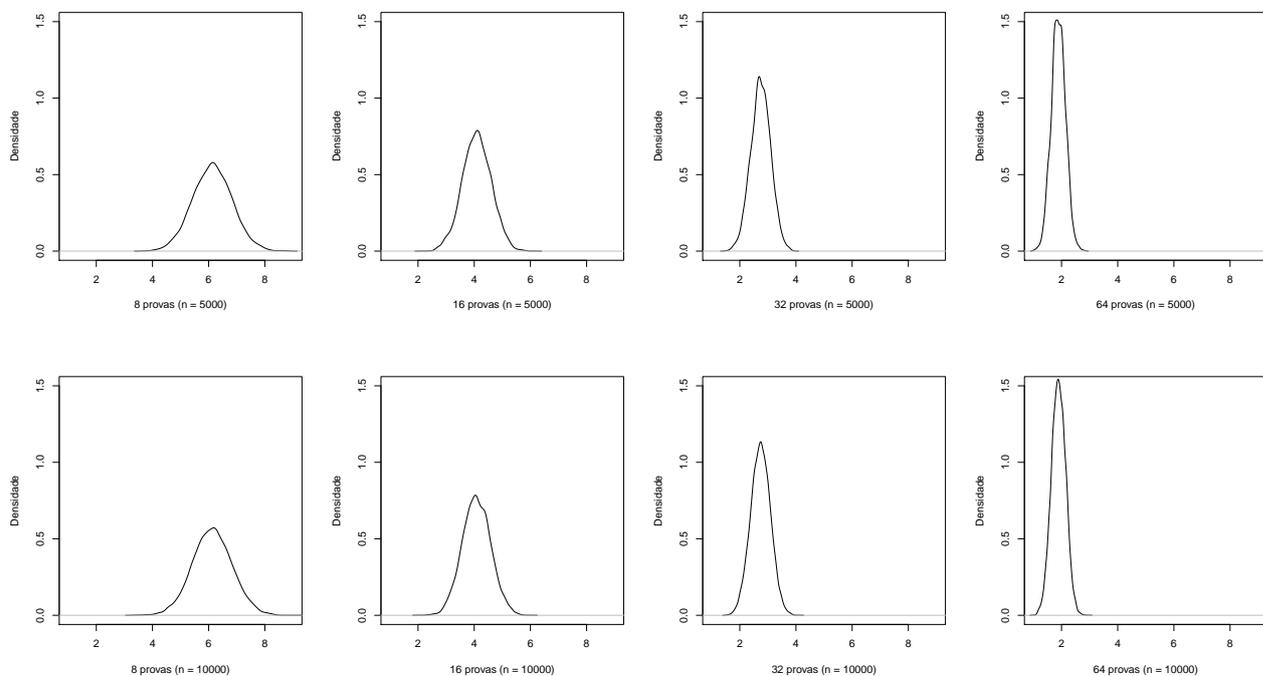


Figura 5.37: **Densidade das Estimativas das Distâncias Calculadas para o Efeito A no Gráfico Normal.** Para estes cenários, o efeito A é significativo, cujo tamanho é $10\sigma_{ef}$. Os histogramas medem a distância do efeito A à reta $y = \sigma_{ef} \cdot G^{-1}(p_j)$.

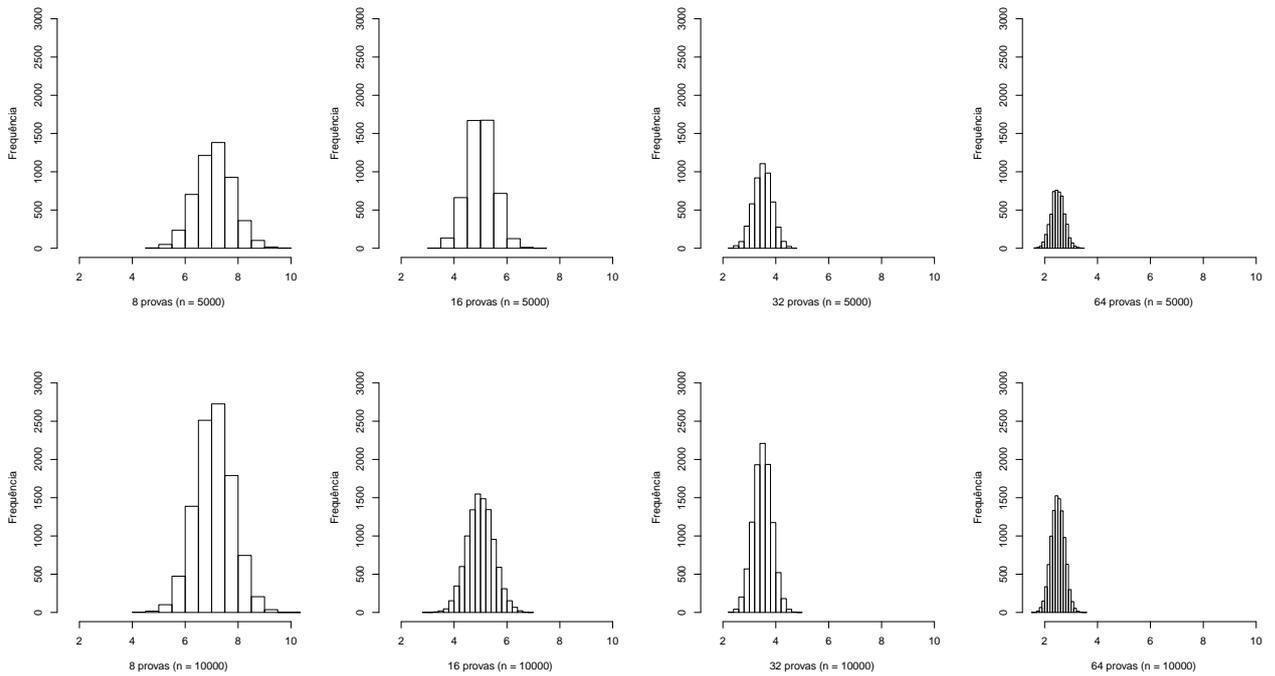


Figura 5.38: **Histogramas das Distâncias Calculadas para o Efeito A no Gráfico Semi-Normal.** Para estes cenários, o efeito A é significativo, cujo tamanho é $10\sigma_{ef}$. Os histogramas medem a distância do efeito A à reta $y = \sigma_{ef} \cdot G^{-1}(p_j)$.

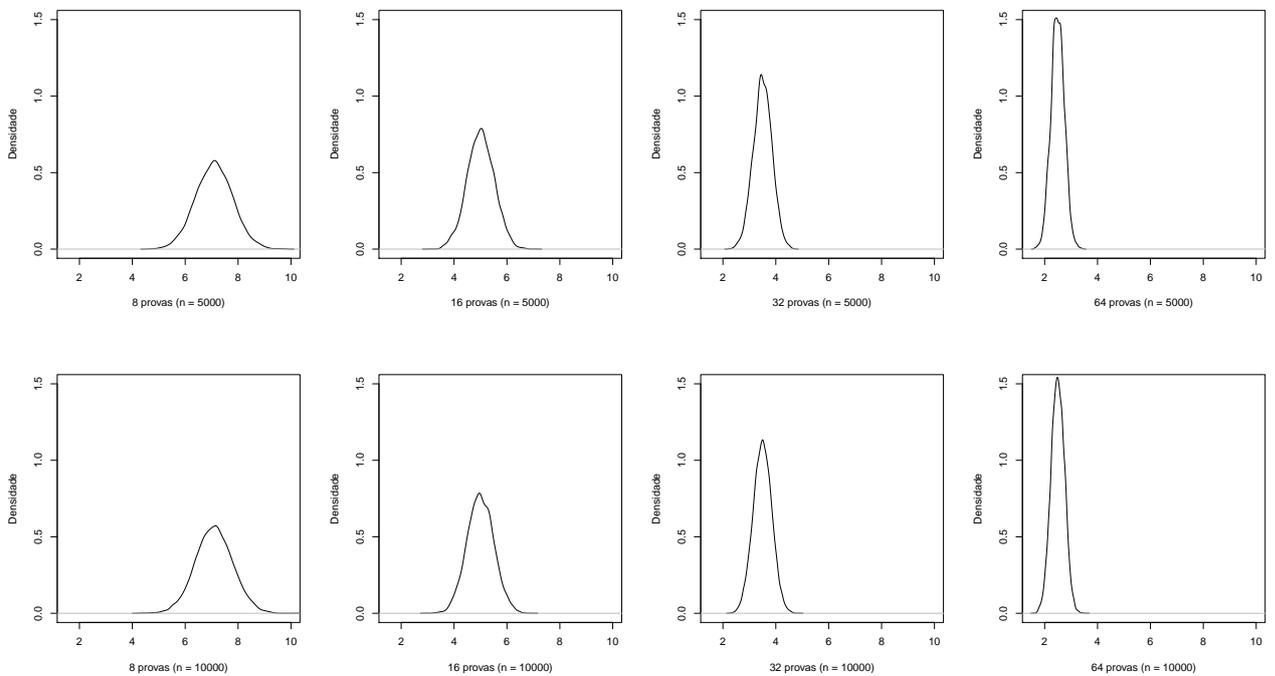


Figura 5.39: **Densidade das Estimativas das Distâncias Calculadas para o Efeito A no Gráfico Semi-Normal.** Para estes cenários, o efeito A é significativo, cujo tamanho é $10\sigma_{ef}$. Os histogramas medem a distância do efeito A à reta $y = \sigma_{ef} \cdot G^{-1}(p_j)$.

Na comparação da habilidade dos gráficos em detectar a presença de valores discrepantes nos dados, percebe-se que o *gap* formado no gráfico normal é destacadamente maior que aquele formado no gráfico semi-normal. Outra característica observada, em ambos os gráficos, é o fato de que os *gaps* formados nos experimentos com 32 e 64 provas são consideravelmente pequenos, muito próximos de zero; o que explica a forma assimétrica apresentada em tais gráficos, evidenciando que tais experimentos são mais robustos à presença de valores discrepantes. As figuras 5.40, 5.41, 5.42 e 5.43 ilustram estas considerações.

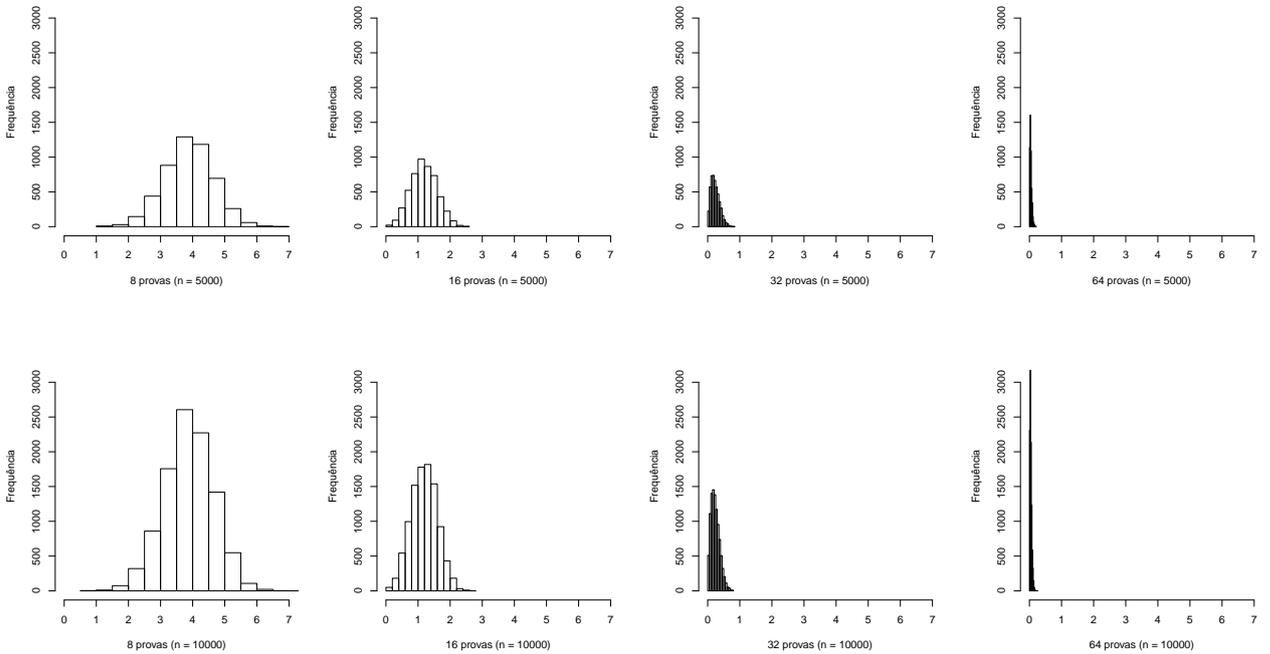


Figura 5.40: **Histogramas do Tamanho dos Gaps no Gráfico Normal.** Para estes cenários, o efeito A é significativo, cujo tamanho é $10\sigma_{ef}$, e há um valor discrepante de tamanho 10 no tratamento ab .

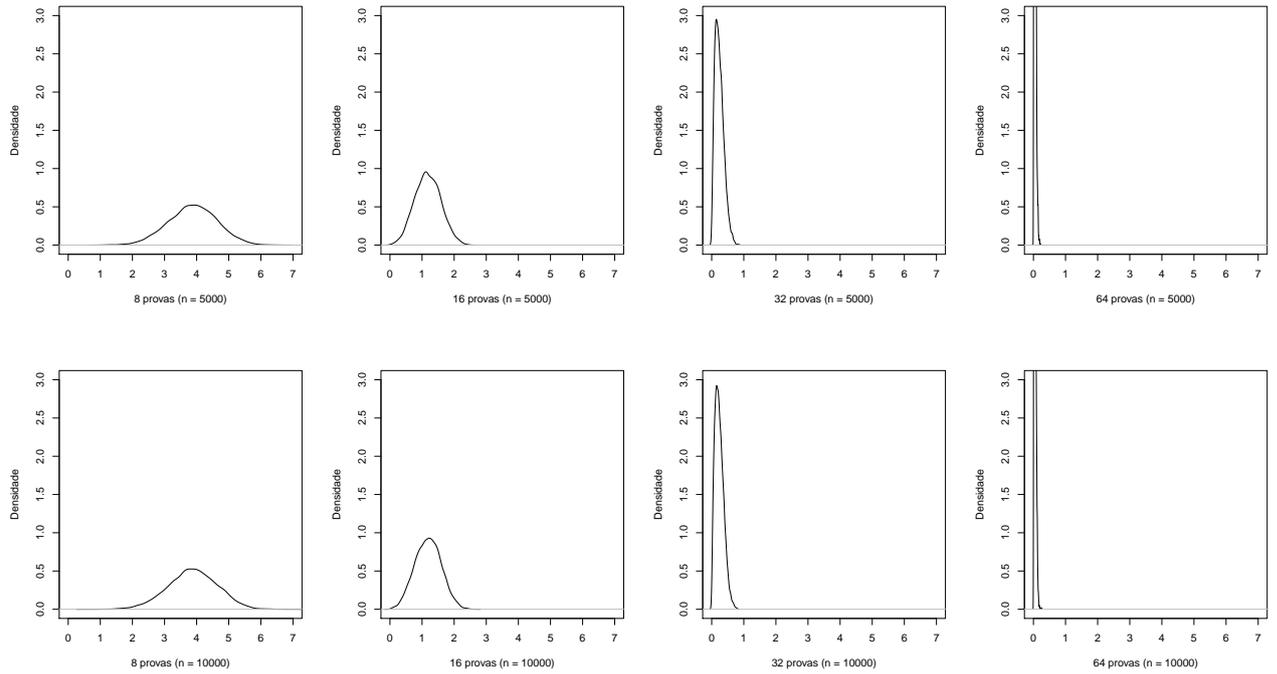


Figura 5.41: Densidade das Estimativas do Tamanho dos *Gaps* no Gráfico Normal. Para estes cenários, o efeito *A* é significativo, cujo tamanho é $10\sigma_{ef}$, e há um valor discrepante de tamanho 10 no tratamento *ab*.

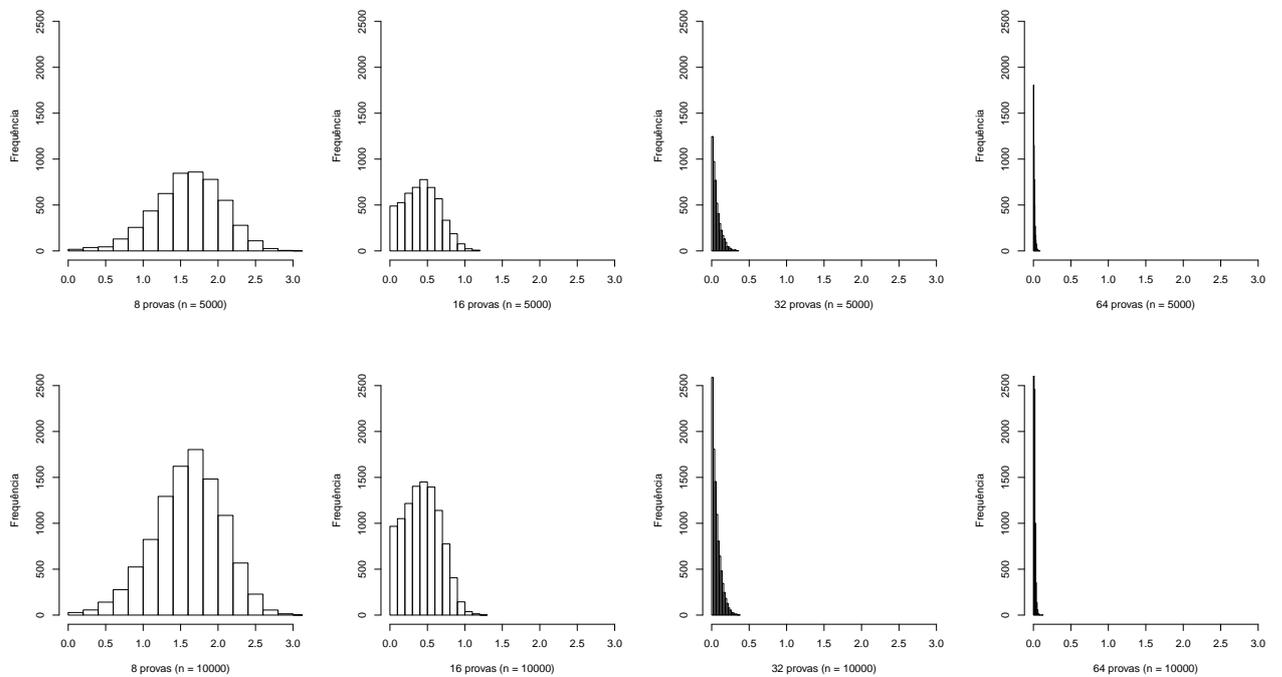


Figura 5.42: Histogramas do Tamanho dos *Gaps* no Gráfico Semi-Normal. Para estes cenários, o efeito *A* é significativo, cujo tamanho é $10\sigma_{ef}$, e há um valor discrepante de tamanho 10 no tratamento *ab*.

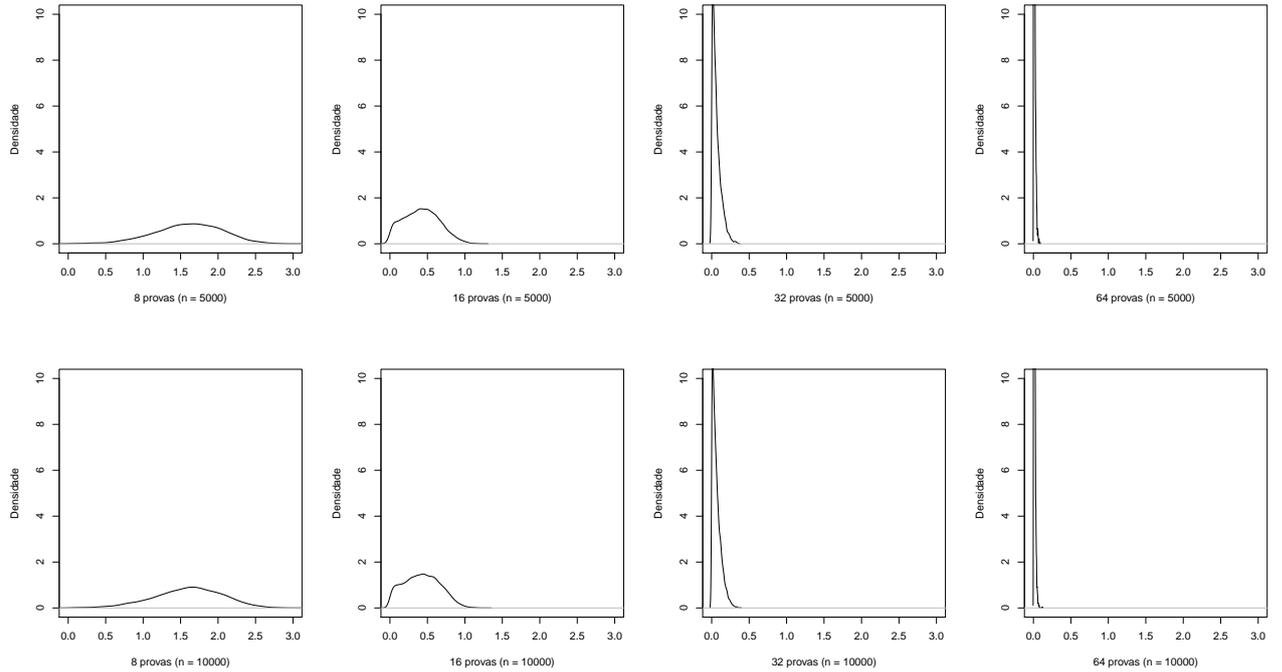


Figura 5.43: **Densidade das Estimativas do Tamanho dos *Gaps* no Gráfico Semi-Normal.** Para estes cenários, o efeito A é significativo, cujo tamanho é $10\sigma_{ef}$, e há um valor discrepante de tamanho 10 no tratamento ab .

Como se pode ver pela análise das figuras e dos comentários acima, não se percebe diferenças acentuadas entre $N = 5000$ e $N = 10000$. Portanto, visando a redução de tempo e de esforços computacionais, é utilizado $N = 5000$ como o número de simulações para cada cenário apresentado e discutido nesta seção. As seções que se seguem apresentam um exame mais detalhado dos resultados do experimento.

5.3.1 Julgamento da Significância dos Efeitos

Esta seção trata da análise dos resultados das simulações sobre a comparação do desempenho dos gráficos normal e semi-normal em avaliar a significância dos efeitos. De acordo com a metodologia utilizada, os efeitos ativos usados nas simulações têm os tamanhos: $\pm 3\sigma_{ef}$, $\pm 5\sigma_{ef}$ e $\pm 10\sigma_{ef}$, em que σ_{ef} , para os experimentos não replicados, é dado por (MONTGOMERY, 2001):

$$\sigma_{ef} = \frac{2 \cdot \sigma_{obs}}{\sqrt{2^k}} \quad (5.1)$$

no qual σ_{obs} é o desvio padrão das observações que, por uma questão de simplicidade computacional e facilidade nas interpretações, nos cenários aqui apresentados, $\sigma_{obs} = 1$. Desta forma, os tamanhos dos efeitos podem ser calculados como mostrado na Tabela 5.2. Analogamente, a quantidade de efeitos ativos nos cenários, de acordo com a metodologia adotada, pode ser vista na Tabela 5.3.

Tabela 5.2: Tamanho dos efeitos ativos usados nas simulações.

Número de provas	Tamanho dos efeitos		
	$\pm 3\sigma_{ef}$	$\pm 5\sigma_{ef}$	$\pm 10\sigma_{ef}$
$2^3 = 8$	$\pm 2,13$	$\pm 3,55$	$\pm 7,07$
$2^4 = 16$	$\pm 1,50$	$\pm 2,50$	$\pm 5,00$
$2^5 = 32$	$\pm 1,05$	$\pm 1,75$	$\pm 3,53$
$2^6 = 64$	$\pm 0,75$	$\pm 1,25$	$\pm 2,50$

Tabela 5.3: Número de efeitos ativos usados nas simulações.

Número de provas	Efeitos significativos			
	0%	10%	20%	40%
$2^3 = 8$	0	1	2	3
$2^4 = 16$	0	2	3	6
$2^5 = 32$	0	3	6	13
$2^6 = 64$	0	6	13	26

Como já dito, o critério de comparação usado para verificar a habilidade dos gráficos no julgamento dos efeitos é a distância destes à reta gerada pela maioria dos pontos. A Tabela 5.4 resume os dados das médias das diferenças das distâncias do gráfico semi-normal para as do gráfico normal, nesta ordem, para cada um dos efeitos de um fatorial 2^3 . Como se pode perceber, o gráfico semi-normal apresenta, em média, distâncias maiores do que o gráfico normal. Esta mesma característica também é notada nos cenários com outros números de provas, tal como o exemplificado na Tabela 5.5.

Tabela 5.4: Diferença média entre as distâncias calculadas nos gráficos num fatorial 2^3 .

Efeitos	0%	10%	20%	40%
<i>A</i>	0,271	0,921	0,921	0,927
<i>B</i>	0,264	0,211	0,964	0,536
<i>C</i>	0,273	0,229	0,201	0,965
<i>AB</i>	0,274	0,227	0,201	0,149
<i>AC</i>	0,266	0,231	0,208	0,150
<i>AD</i>	0,271	0,222	0,209	0,140
<i>ABC</i>	0,258	0,228	0,203	0,150

Tabela 5.5: Diferença média entre as distâncias calculadas nos gráficos num fatorial 2^4 .

Efeitos	0%	10%	20%	40%
<i>A</i>	0,242	0,857	0,612	0,479
<i>B</i>	0,247	0,916	0,916	0,629
<i>C</i>	0,247	0,236	0,917	0,917
<i>D</i>	0,249	0,233	0,210	0,917
<i>AB</i>	0,245	0,228	0,207	0,160
<i>AC</i>	0,247	0,229	0,216	0,477
<i>AD</i>	0,251	0,229	0,210	0,159
<i>BC</i>	0,244	0,232	0,211	0,157
<i>BD</i>	0,240	0,228	0,212	0,628
<i>CD</i>	0,251	0,233	0,208	0,154
<i>ABC</i>	0,247	0,231	0,209	0,159
<i>ABD</i>	0,244	0,231	0,206	0,162
<i>ACD</i>	0,239	0,236	0,208	0,159
<i>BCD</i>	0,244	0,229	0,210	0,160
<i>ABCD</i>	0,255	0,235	0,208	0,156

As Figuras 5.44, 5.45, 5.46, 5.47, 5.48, 5.49, 5.50 e 5.51 fornecem uma visão comparativa do comportamento das distâncias calculadas em ambos os gráficos. É nítido que as distâncias do gráfico semi-normal são maiores que as do gráfico normal. Sendo assim, conforme já comentado na Seção 5.2.1, o gráfico semi-normal é a técnica mais útil em auxiliar o investigador no seu julgamento sobre a significância dos efeitos dos fatores. Uma outra característica observada é que, à medida que aumenta a quantidade de efeitos significativos, os valores das distâncias calculadas, para ambos os gráficos, também aumentam, assim como aumenta também o número de efeitos com distâncias semelhantes; o que torna difícil a interpretação visual do julgamento destes efeitos: uma vez que muitos efeitos estão, aproximadamente, a uma mesma distância da reta, fica difícil ao investigador discernir, apenas pela avaliação visual, qual deles é realmente ativo. Por conta disso, percebe-se que nas situações em que 40% dos efeitos são significativos, a habilidade das técnicas gráficas em julgar tais efeitos é comprometida. A situação “ideal” para o uso dos gráficos de probabilidade é aquela em que menos de 20% dos efeitos são significantes (DANIEL, 1967; DONG, 1993).

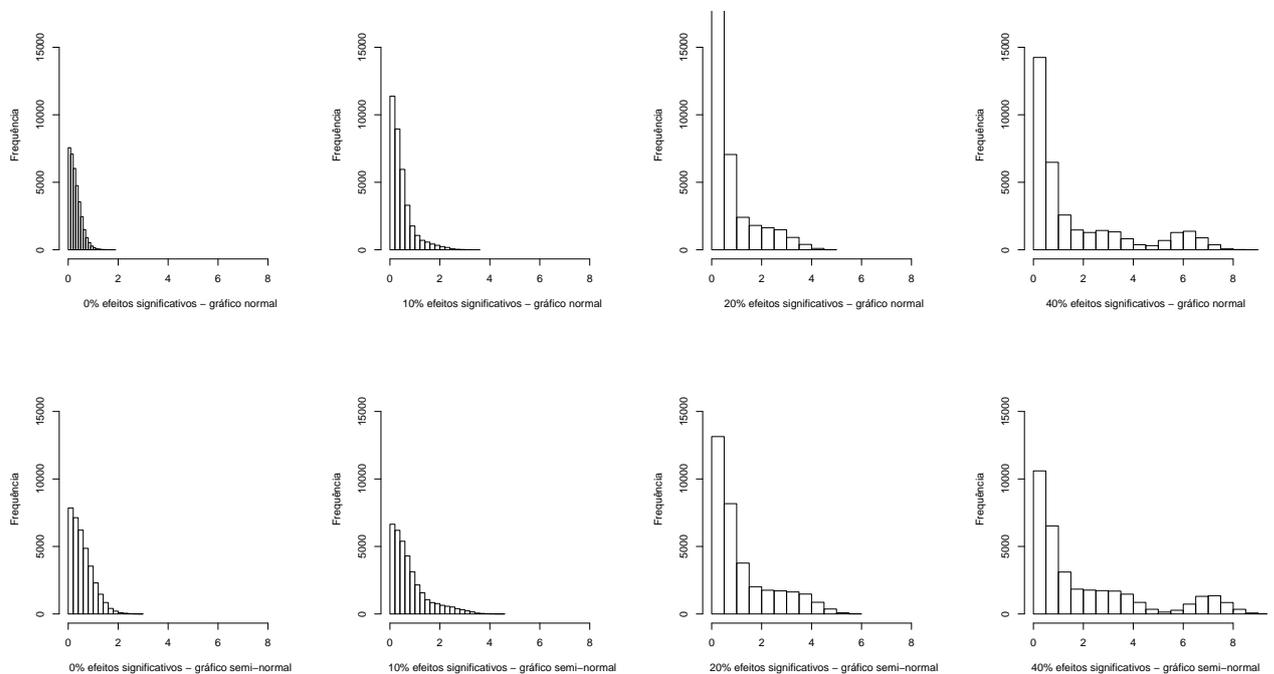


Figura 5.44: Comparação das distâncias nos gráficos de probabilidade. Trata-se de simulações de planos fatoriais 2^3 .

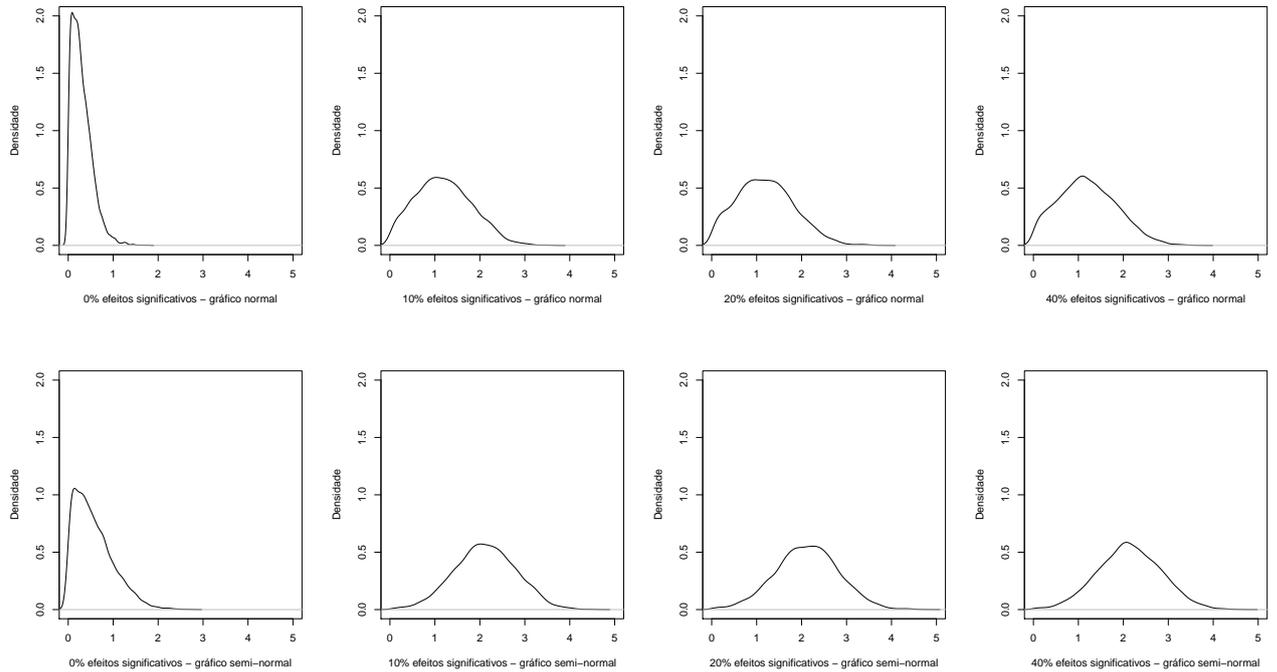


Figura 5.45: Estimação de Densidade das distâncias nos gráficos de probabilidade. Trata-se de simulações de planos fatoriais 2^3 .

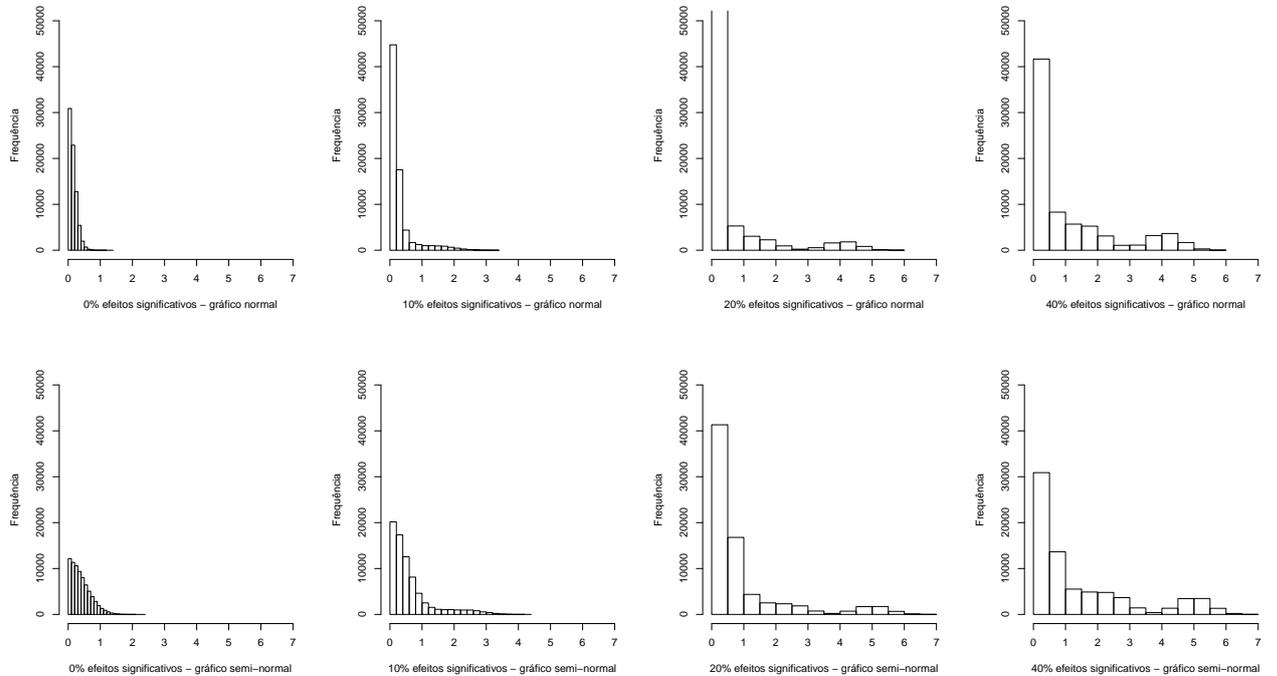


Figura 5.46: Comparação das distâncias nos gráficos de probabilidade. Trata-se de simulações de planos fatoriais 2^4 .

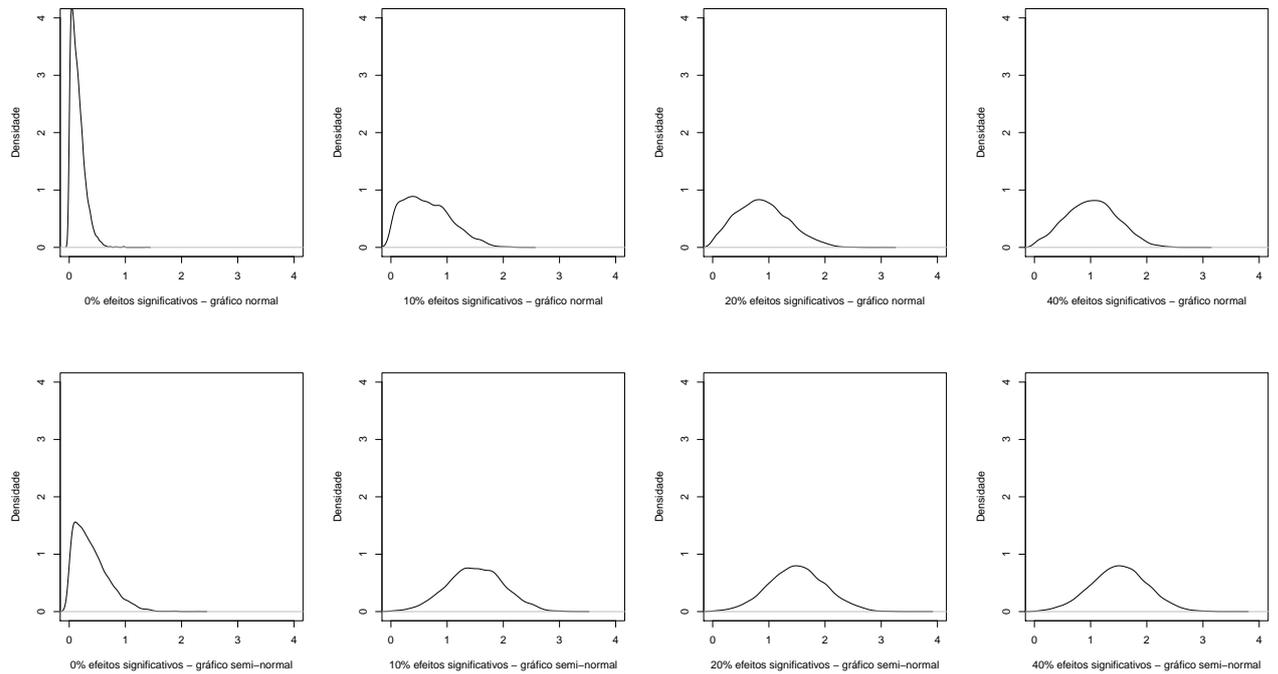


Figura 5.47: **Estimação de Densidade das distâncias nos gráficos de probabilidade.** Trata-se de simulações de planos fatoriais 2^4 .

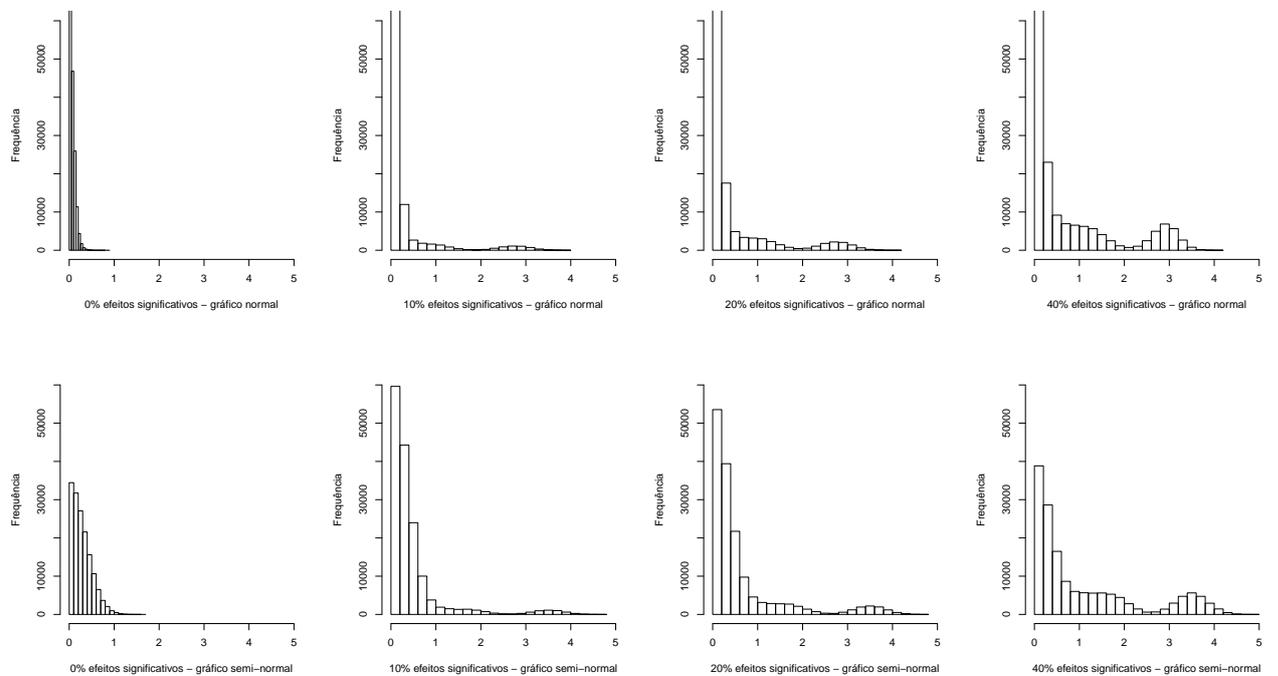


Figura 5.48: **Comparação das distâncias nos gráficos de probabilidade.** Trata-se de simulações de planos fatoriais 2^5 .

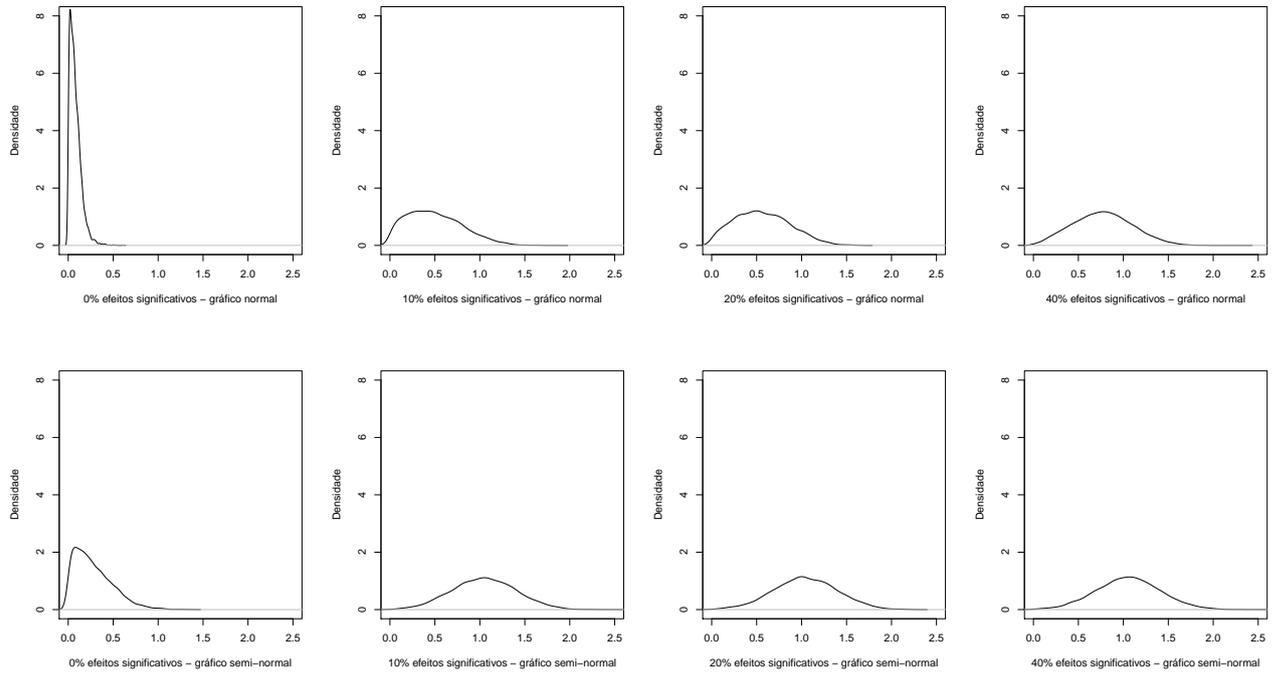


Figura 5.49: Estimação de Densidade das distâncias nos gráficos de probabilidade. Trata-se de simulações de planos fatoriais 2^5 .

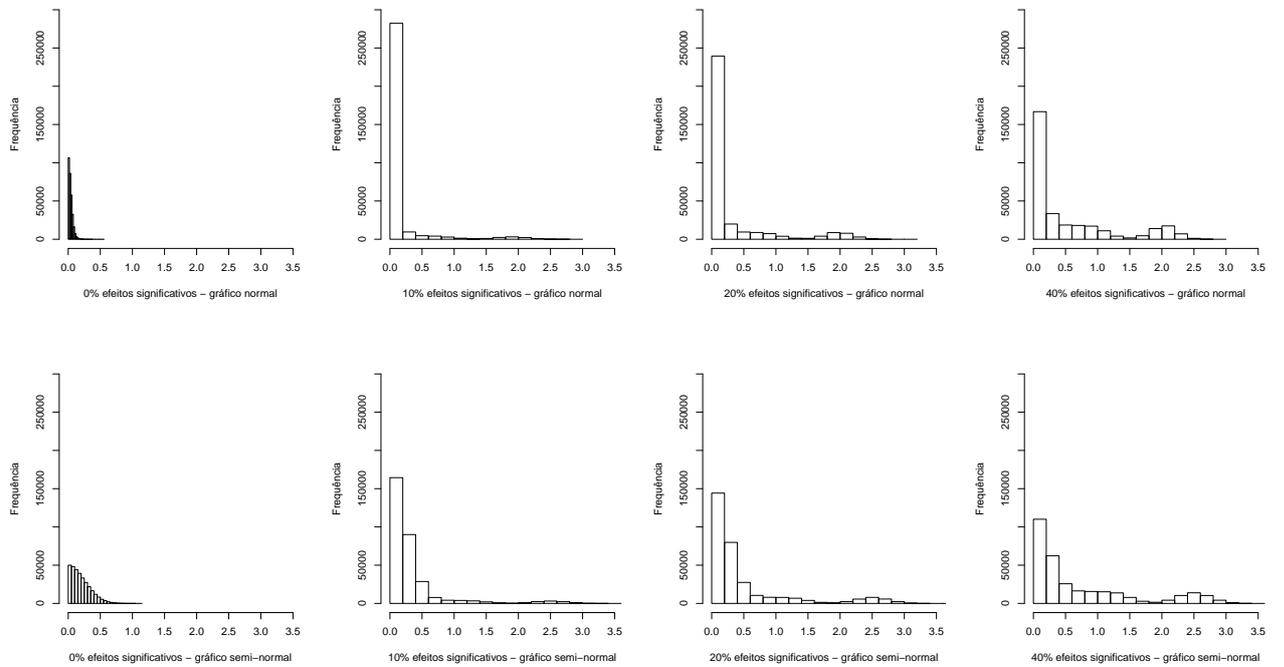


Figura 5.50: Comparação das distâncias nos gráficos de probabilidade. Trata-se de simulações de planos fatoriais 2^6 .

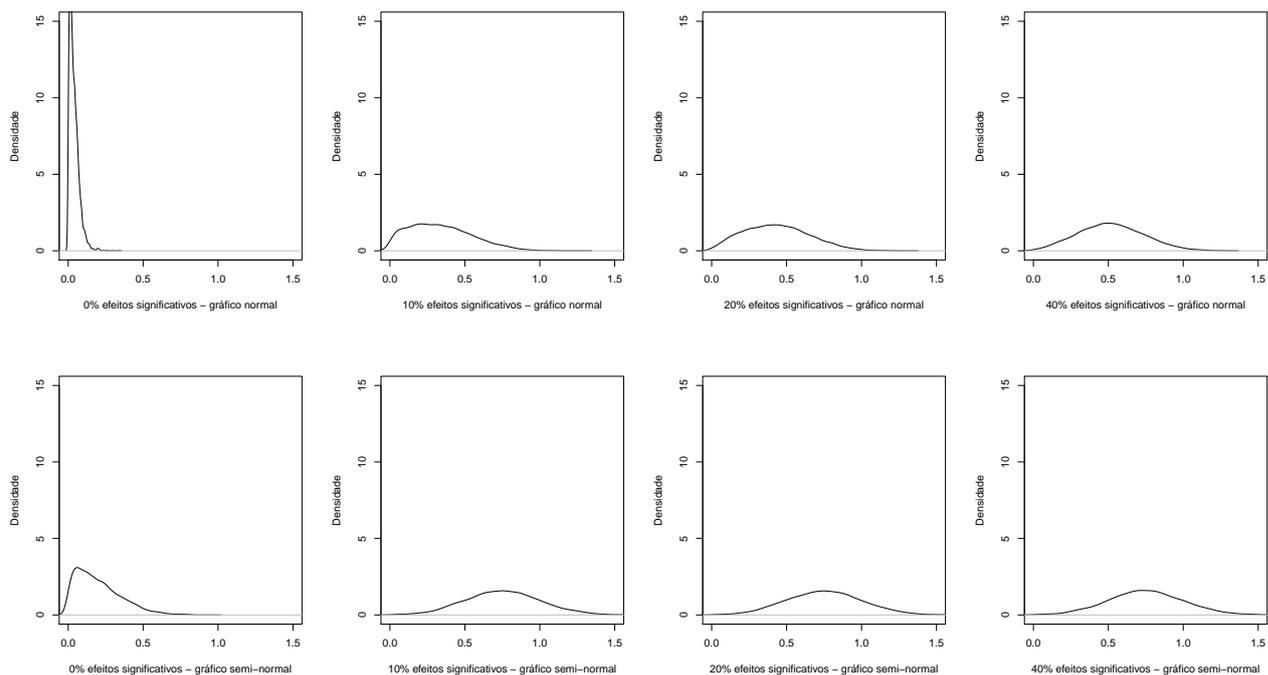


Figura 5.51: **Estimação de Densidade das distâncias nos gráficos de probabilidade.** Trata-se de simulações de planos fatoriais 2^6 .

Com relação ao tamanho dos efeitos, nota-se que os efeitos $\pm 3\sigma_{ef}$ são os mais difíceis de julgar, pois suas distâncias à reta são as menores, e quanto maior o número de provas (especialmente 2^5 e 2^6), mais difícil é essa identificação. Portanto, pode-se concluir que os gráficos de probabilidade são úteis em identificar efeitos significativos apenas se estes forem poucos (menor ou igual a 20%) e grandes (preferencialmente de tamanhos a partir de $\pm 10\sigma_{ef}$).

5.3.2 Detecção de Valores Discrepantes nos Dados

Os resultados da comparação da habilidade dos gráficos normal e semi-normal em detectar valores discrepantes nos dados são discutidos nesta seção. De acordo com a metodologia utilizada, os valores discrepantes usados nas simulações têm tamanhos: $\pm 3\sigma_{obs}$, $\pm 5\sigma_{obs}$ e $\pm 10\sigma_{obs}$. Como $\sigma_{obs} = 1$ para os cenários utilizados aqui, os valores discrepantes são ± 3 , ± 5 e ± 10 . As análises são feitas considerando dois casos separadamente: um valor discrepante está presente nos dados; dois valores discrepantes estão presentes. Por uma questão didática e de objetividade no texto, os cenários e os exemplos apresentados nesta seção referem-se às simulações em que 0% dos efeitos é

significativo. O motivo de tal escolha é que, sendo todos os efeitos iguais a zero, fica mais fácil a interpretação e análise dos resultados, como também deixa mais evidente a influência dos valores discrepantes sobre a estimativa dos efeitos, que também é uma característica a ser investigada.

Um Valor Discrepante Presente

As simulações realizadas apontam para os seguintes resultados:

- Os sinais de “+” e “-” em ± 3 , ± 5 e ± 10 nos valores discrepantes não influenciam o tamanho dos *gaps*. Estes são afetados apenas pelo valor absoluto do tamanho do valor discrepante. Os valores negativos produzem resultados semelhantes aos valores positivos.
- O *gap* formado no gráfico normal é consideravelmente maior que aquele formado no gráfico semi-normal. Uma análise cuidadosa dos gráficos da Figura 5.54 até a Figura 5.65, atentando para os limites expressos nas abscissas de tais gráficos, revela que há fortes evidências para se crer que a identificação de valores discrepantes nos dados será mais clara no gráfico normal, uma vez que a percepção visual do *gap* é mais evidente.
- Acerca da posição em que se encontra o valor discrepante, os dois gráficos apresentam características distintas. No caso do gráfico semi-normal, o tamanho do *gap* permanece aproximadamente o mesmo, independentemente da posição do valor discrepante; só mudando de tamanho quando o valor discrepante muda seu tamanho também. No gráfico normal, a única posição que altera o tamanho do *gap* é quando o valor discrepante encontra-se na observação associada ao tratamento em que todos os fatores estão em seu nível “+”. Quando isso ocorre, há uma diminuição no tamanho do *gap*. Nota-se, então, que o gráfico semi-normal é sensível ao tamanho do valor discrepante; já o gráfico normal é sensível tanto ao tamanho quanto à posição do valor discrepante.
- Quanto maior for o valor discrepante, maior será o tamanho do *gap*. A situação gráfica que permite uma identificação mais clara da presença de valores discrepantes é quando este tem o tamanho 10. Como se pode ver pelas Figuras 5.62, 5.63,

5.64 e 5.65, é neste caso que a distribuição dos *gaps* tem uma forma diferente dos outros casos. Isso evidencia que os gráficos de probabilidade são úteis em detectar grandes valores discrepantes (DANIEL, 1959, 1960; LAWSON, 2008). Pequenos valores discrepantes tornam a identificação gráfica difícil e ainda mais subjetiva.

- Outra importante característica observada nas simulações é o fato de que os experimentos com 32 e 64 provas são mais robustos à presença de valores discrepantes (LAWSON, 2008). De acordo com as Figuras 5.54 à 5.65, tais experimentos, diferentemente daqueles com 8 e 16 provas, não são tão grandemente influenciados pelos valores discrepantes. A forma gráfica da distribuição de seus *gaps* é bastante parecida à forma assumida quando nenhum valor discrepante está presente (Figuras 5.52 e 5.53). Isso indica que a identificação de valores discrepantes nos gráficos de probabilidade construídos para tais planos experimentais pode ser mais difícil, uma vez que os *gaps* de tais gráficos têm tamanhos relativamente menores.
- As considerações apresentadas nos itens acima mostram que, o fato dum gráfico de probabilidade não evidenciar o característico *gap* que indica a presença de valores discrepantes, não significa que eles estejam ausentes. Pelo menos duas possibilidades podem ser consideradas: ou o valor discrepante é consideravelmente pequeno, ou o experimento é robusto à sua presença.

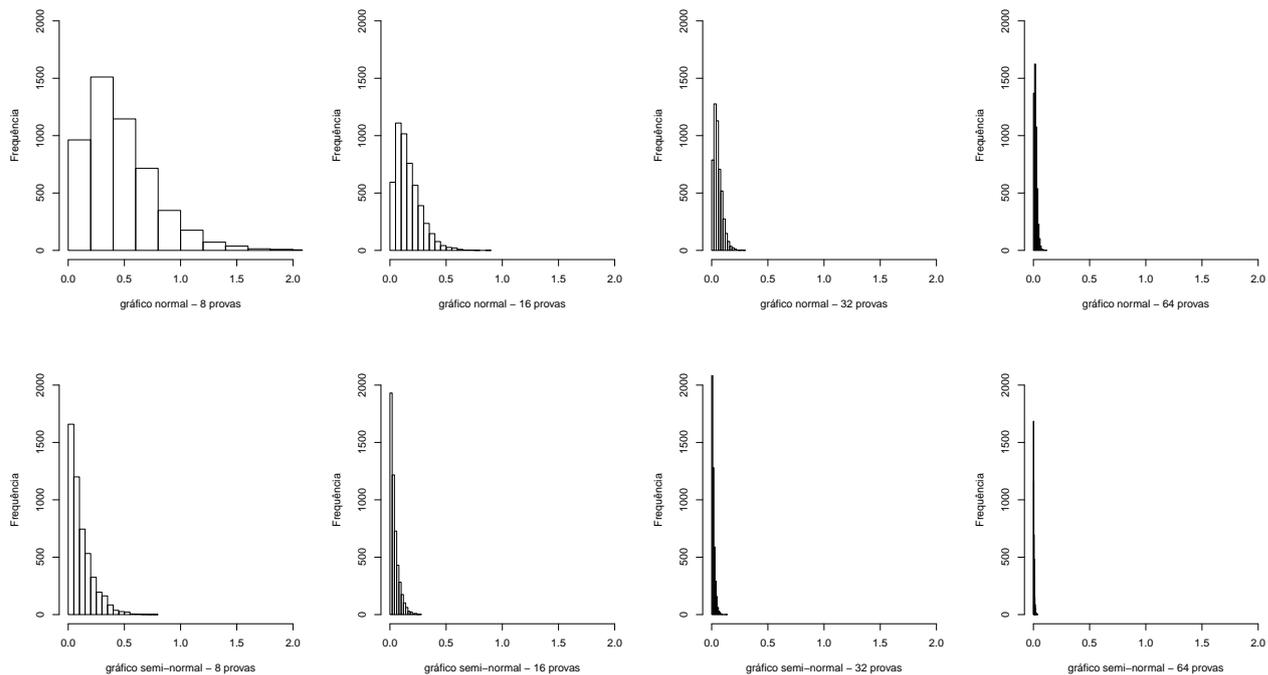


Figura 5.52: Comparação dos *gaps* nos gráficos de probabilidade sem a presença de valores discrepantes.

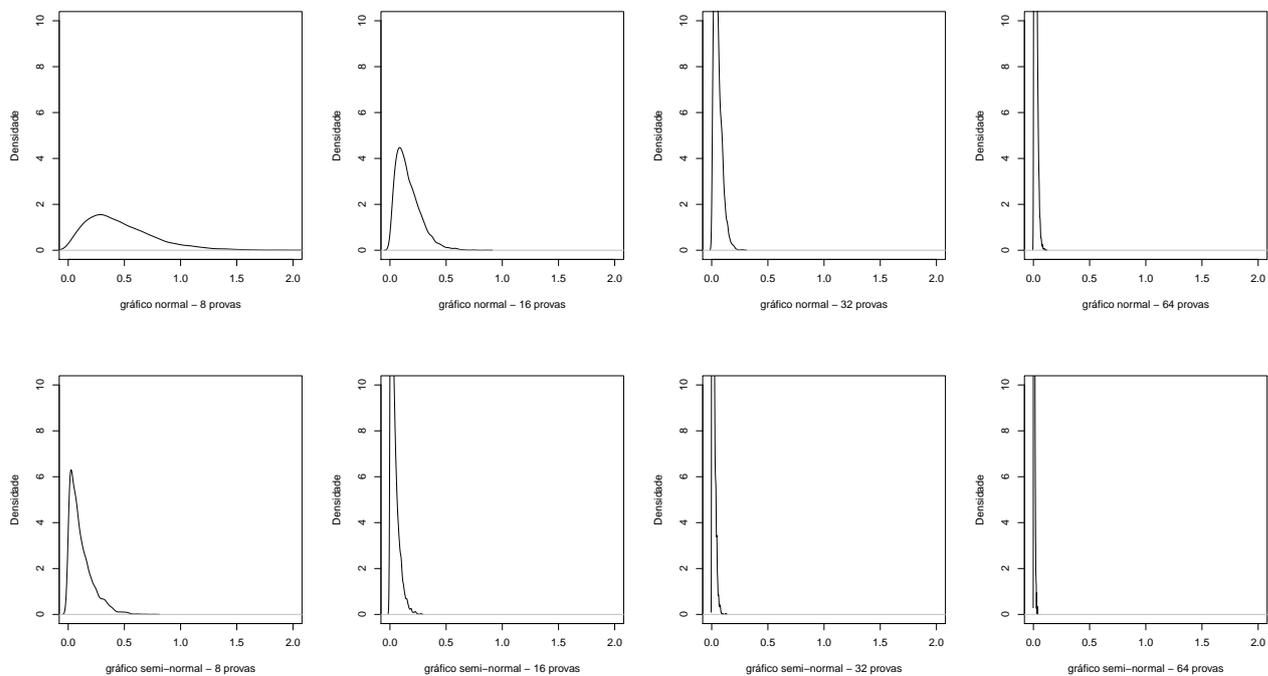


Figura 5.53: Estimação de Densidade dos *gaps* nos gráficos de probabilidade sem a presença de valores discrepantes.

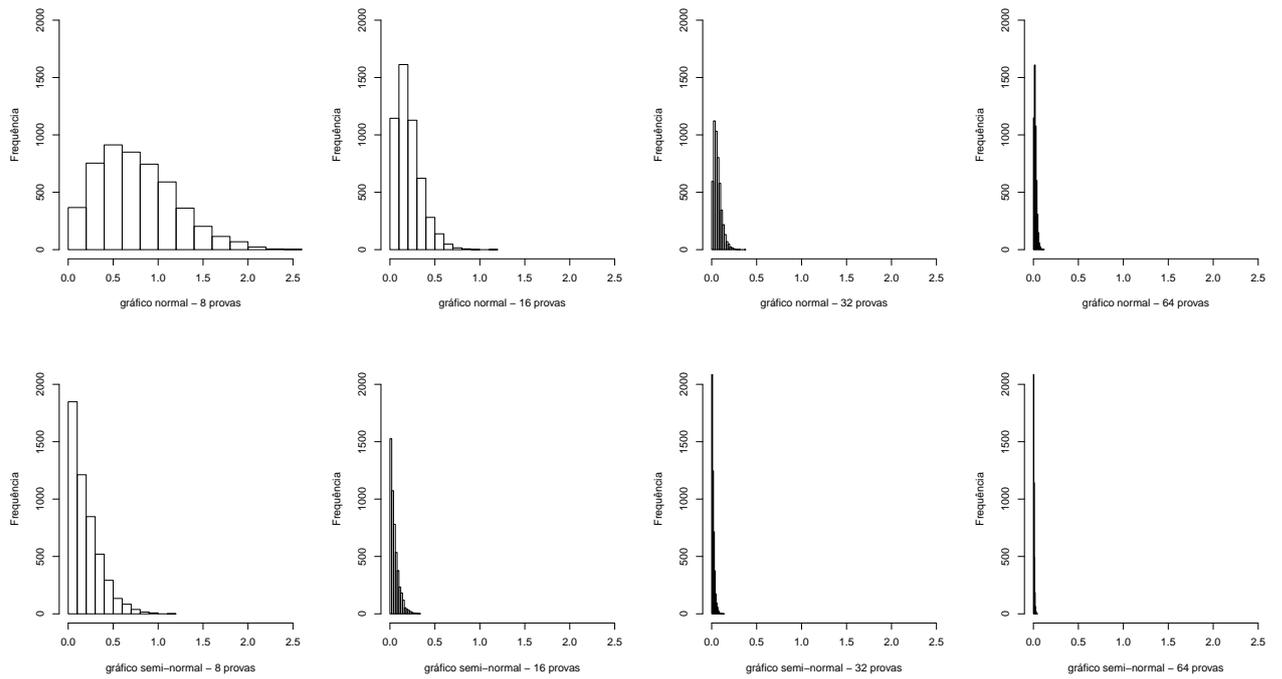


Figura 5.54: Comparação dos *gaps* nos gráficos de probabilidade quando o valor discrepante é igual a 3 e encontra-se no tratamento (1).

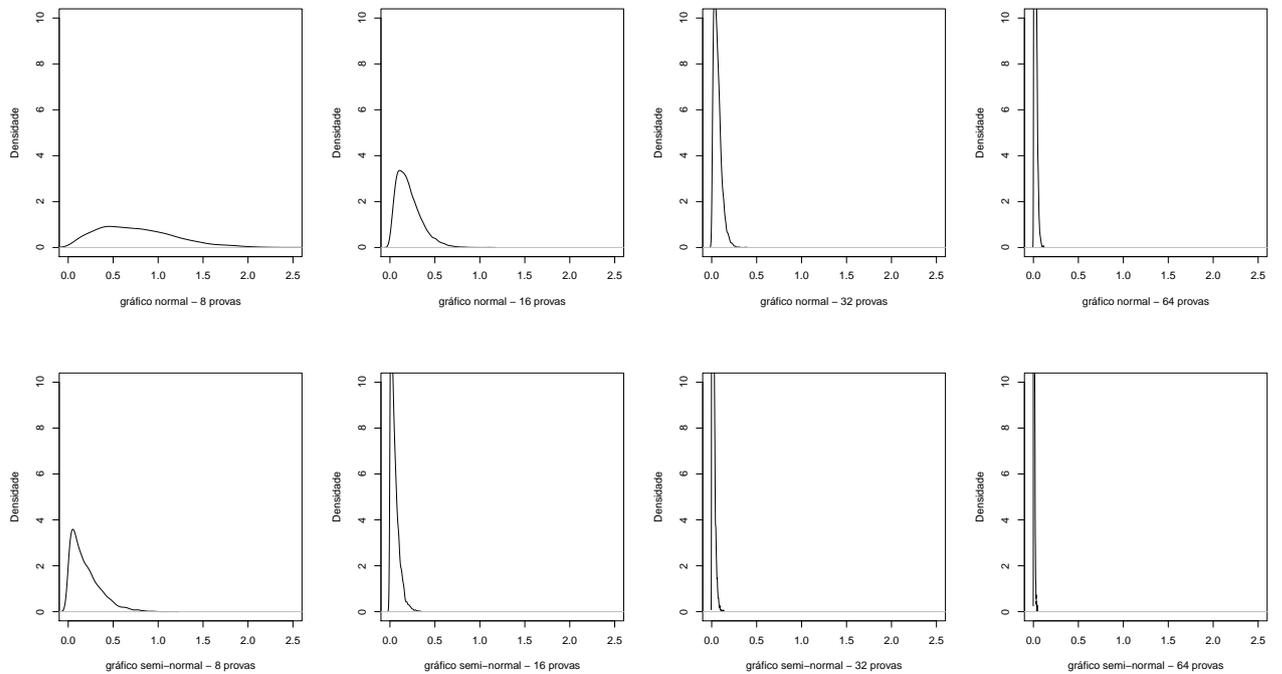


Figura 5.55: Estimação de Densidade dos *gaps* nos gráficos de probabilidade quando o valor discrepante é igual a 3 e encontra-se no tratamento (1).

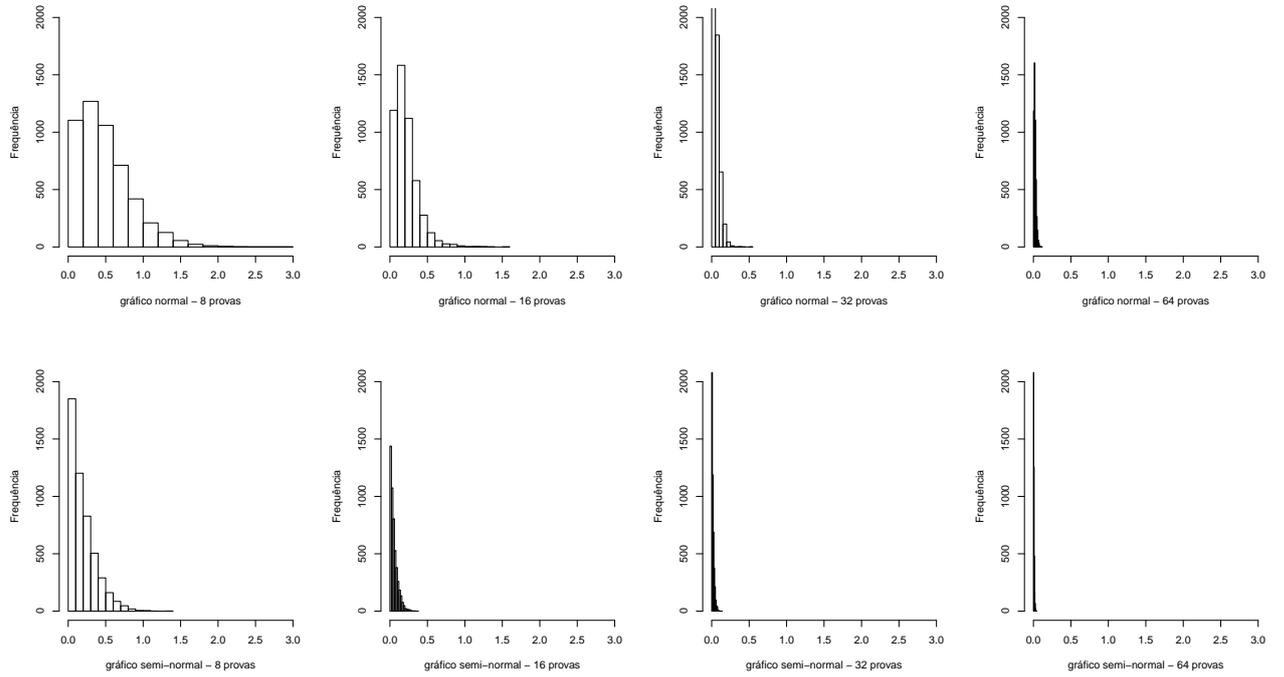


Figura 5.56: Comparação dos *gaps* nos gráficos de probabilidade quando o valor discrepante é igual a 3 e encontra-se no tratamento em que todos os fatores estão em seu nível “+”.

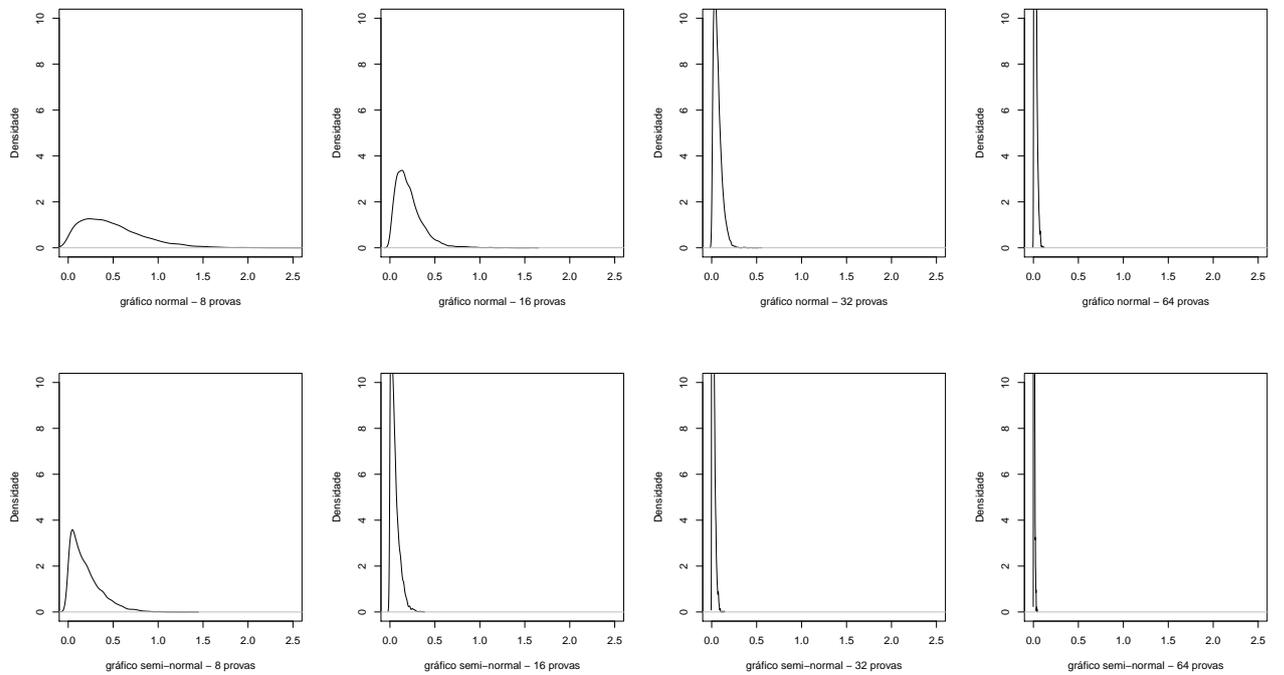


Figura 5.57: Estimação de Densidade dos *gaps* nos gráficos de probabilidade quando o valor discrepante é igual a 3 e encontra-se no tratamento em que todos os fatores estão em seu nível “+”.

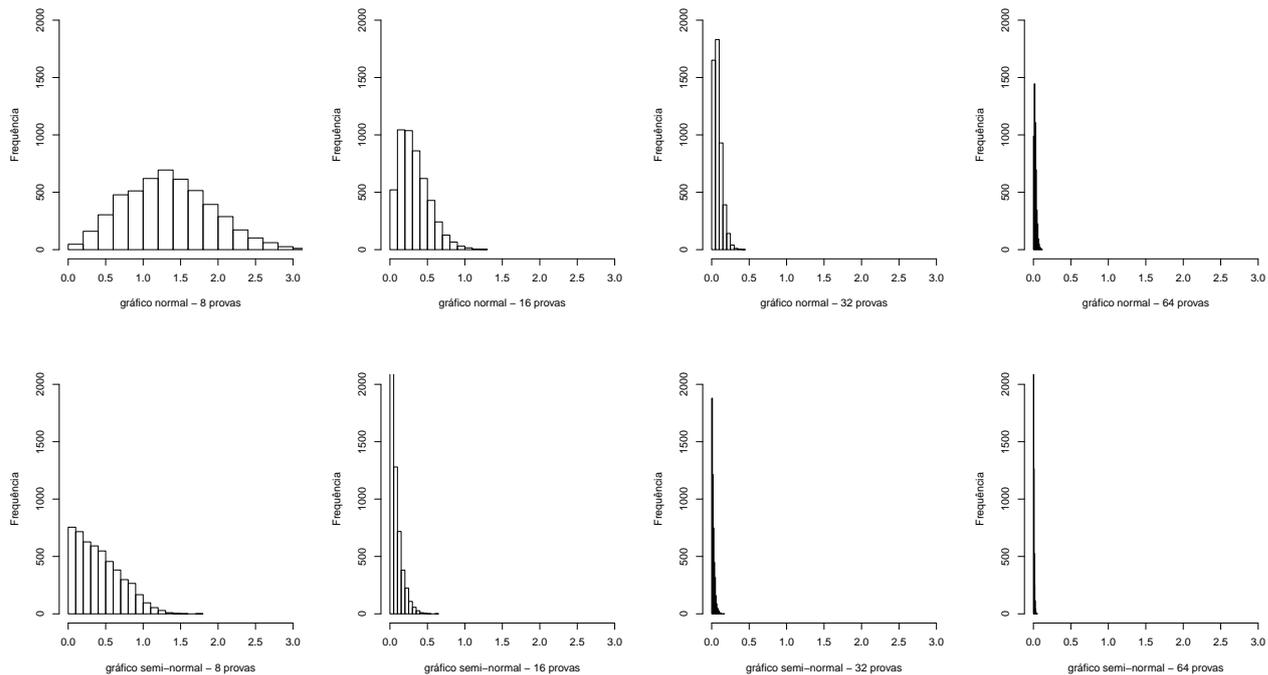


Figura 5.58: Comparação dos *gaps* nos gráficos de probabilidade quando o valor discrepante é igual a 5 e encontra-se no tratamento (1).

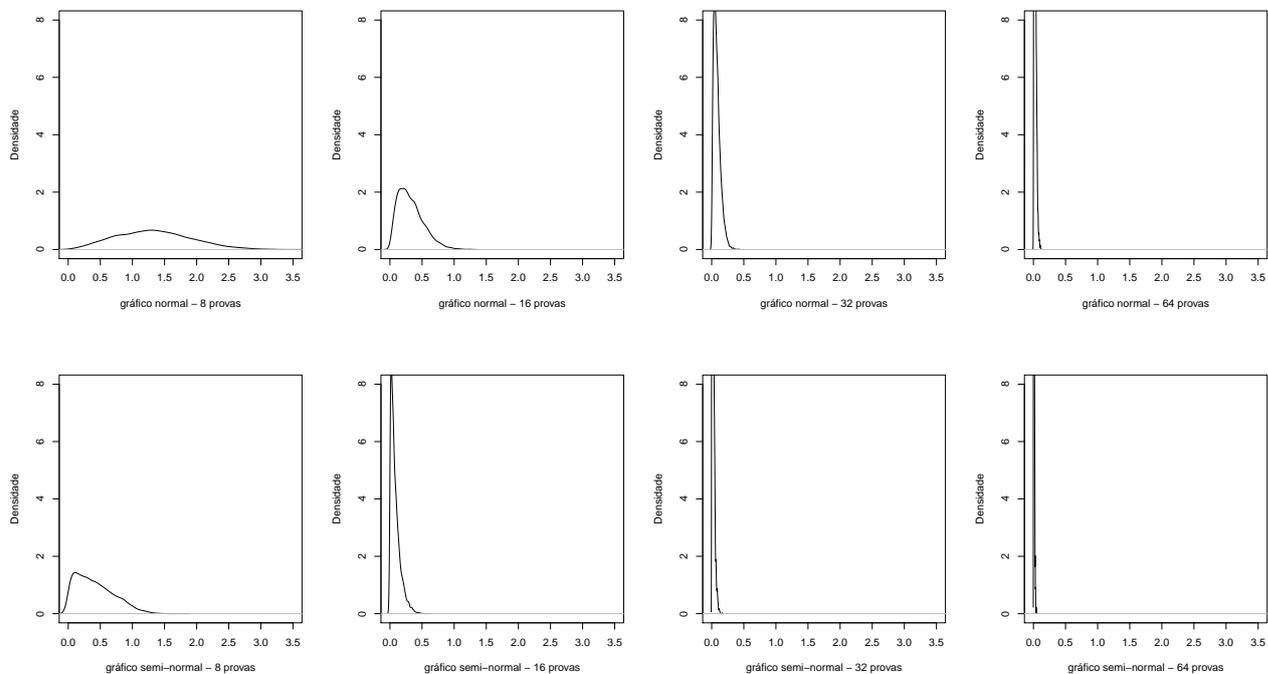


Figura 5.59: Estimação de Densidade dos *gaps* nos gráficos de probabilidade quando o valor discrepante é igual a 5 e encontra-se no tratamento (1).

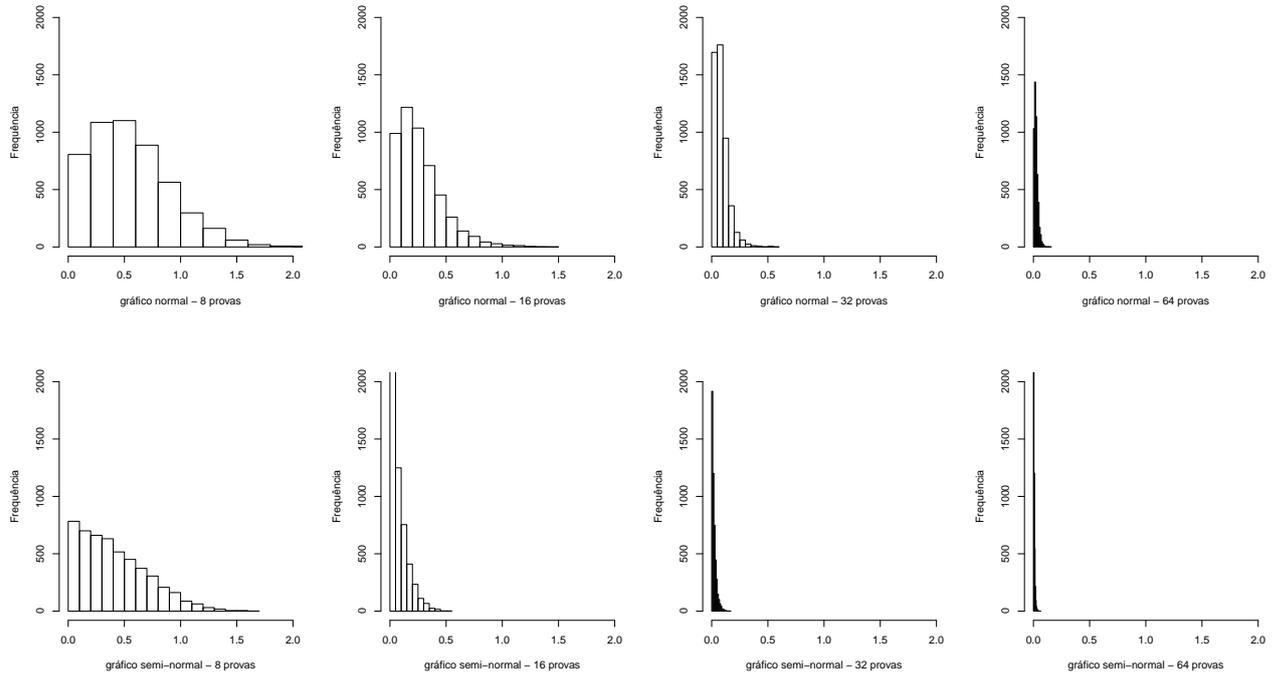


Figura 5.60: Comparação dos *gaps* nos gráficos de probabilidade quando o valor discrepante é igual a 5 e encontra-se no tratamento em que todos os fatores estão em seu nível “+”.

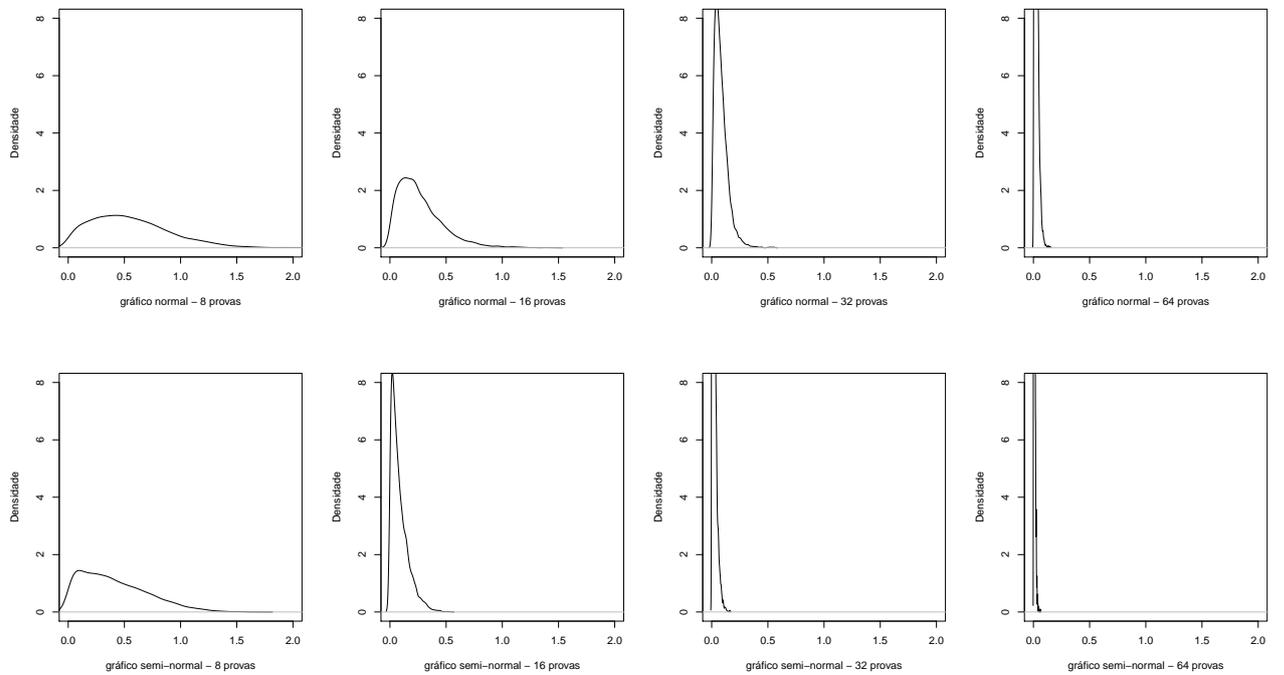


Figura 5.61: Estimação de Densidade dos *gaps* nos gráficos de probabilidade quando o valor discrepante é igual a 5 e encontra-se no tratamento em que todos os fatores estão em seu nível “+”.

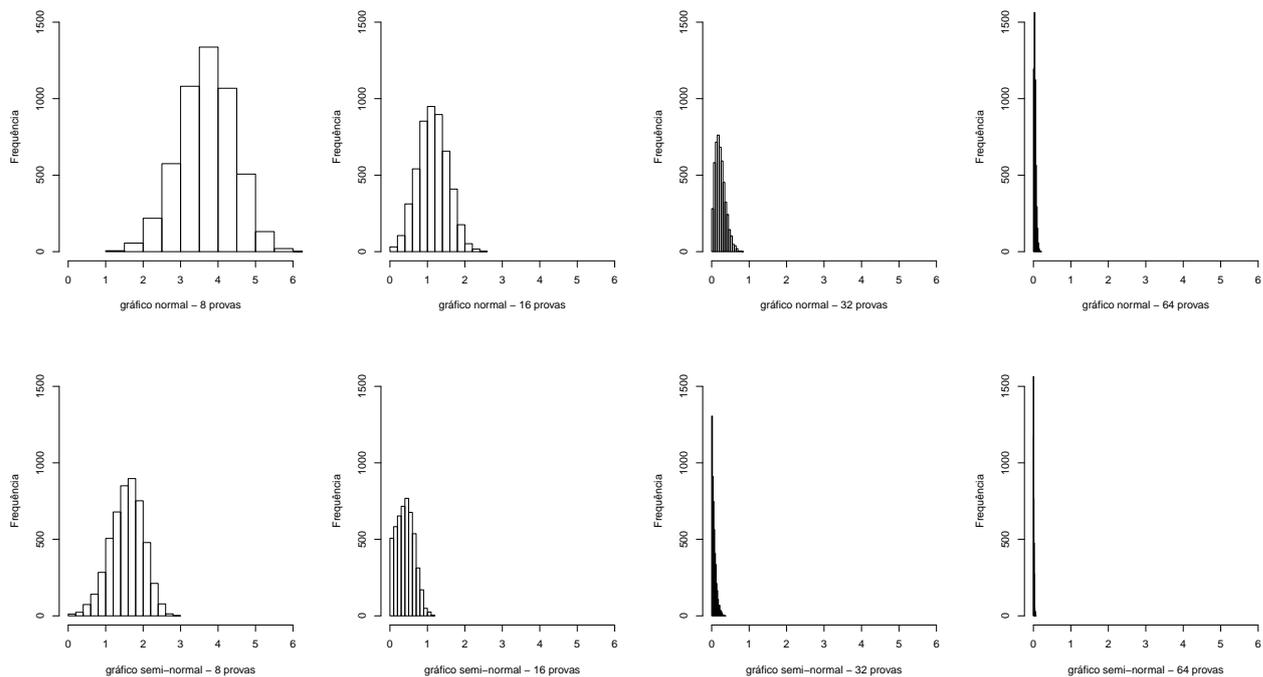


Figura 5.62: Comparação dos *gaps* nos gráficos de probabilidade quando o valor discrepante é igual a 10 e encontra-se no tratamento (1).

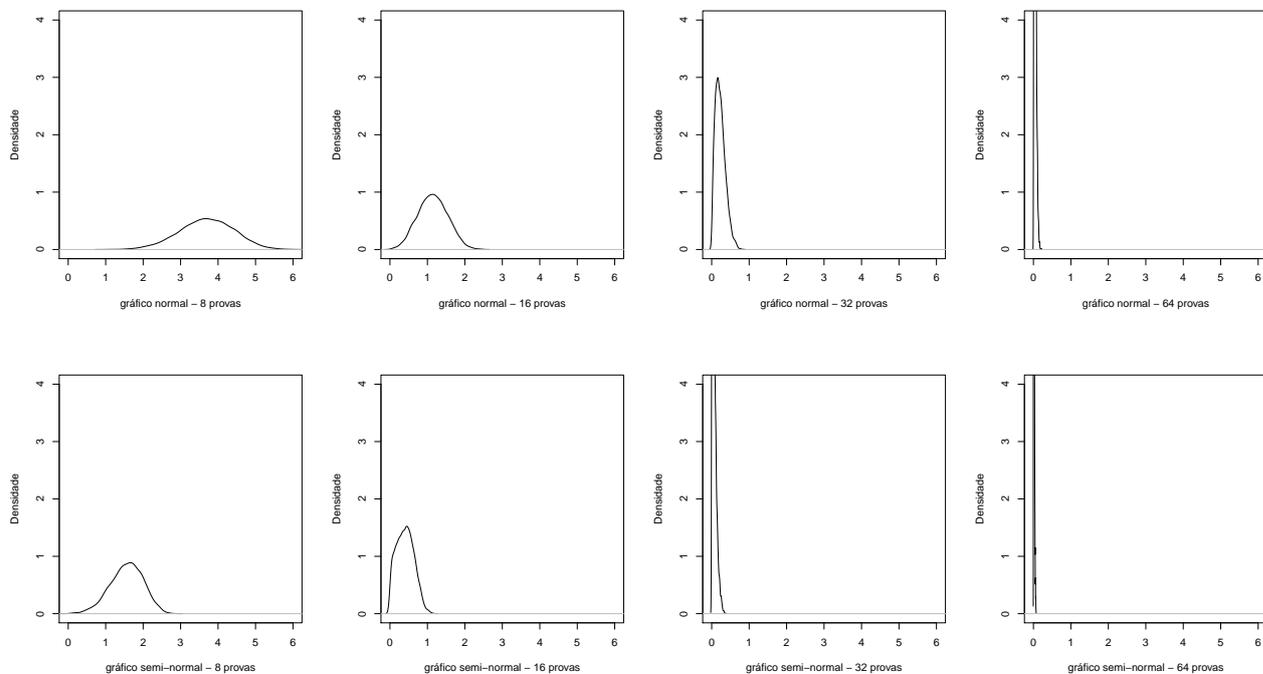


Figura 5.63: Estimação de Densidade dos *gaps* nos gráficos de probabilidade quando o valor discrepante é igual a 10 e encontra-se no tratamento (1).

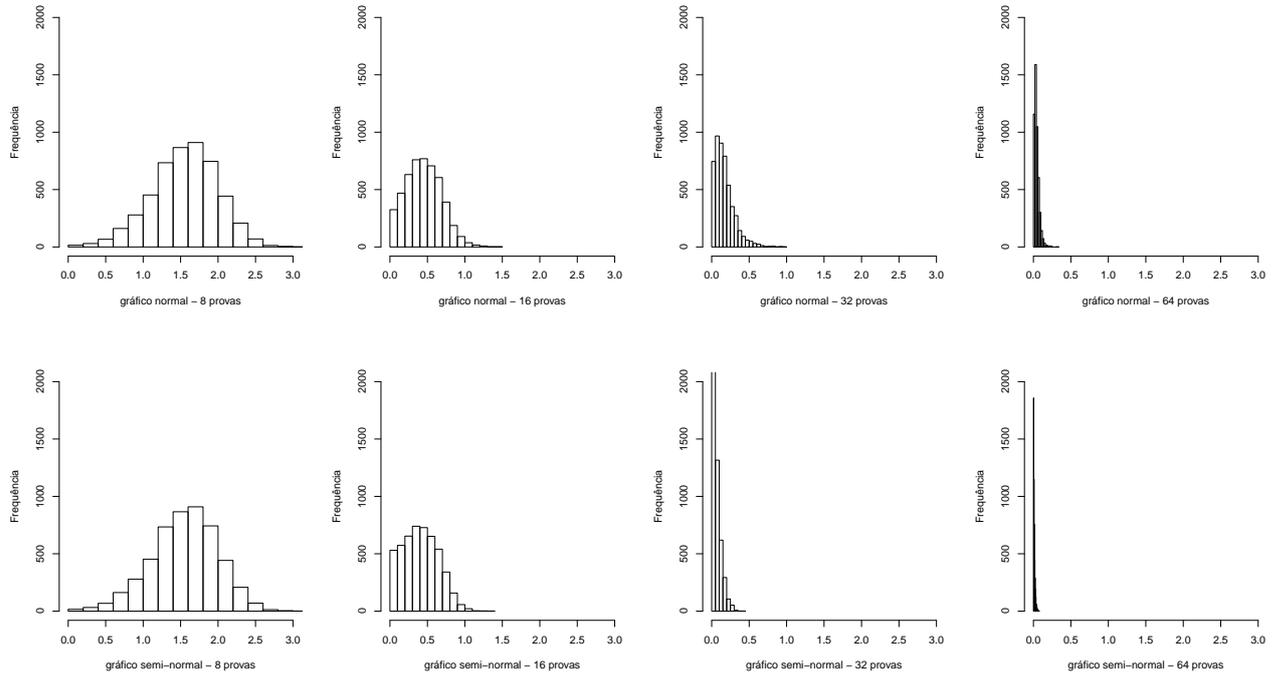


Figura 5.64: Comparação dos *gaps* nos gráficos de probabilidade quando o valor discrepante é igual a 10 e encontra-se no tratamento em que todos os fatores estão em seu nível “+”.

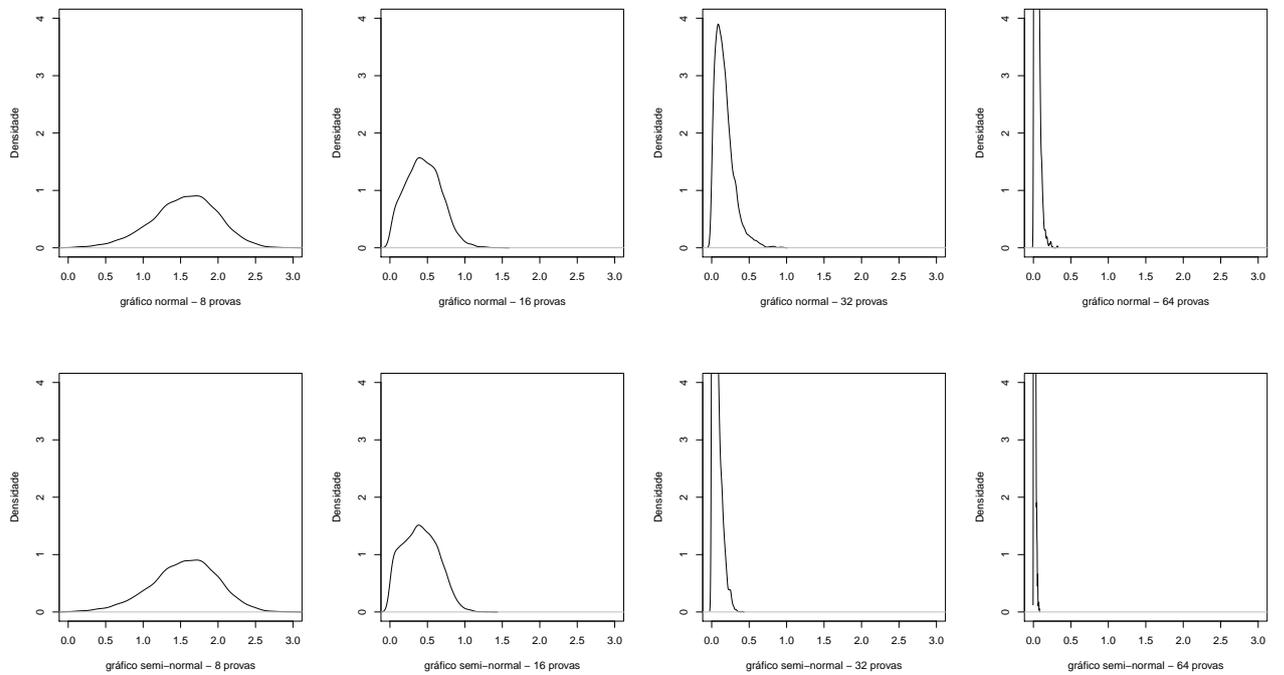


Figura 5.65: Estimação de Densidade dos *gaps* nos gráficos de probabilidade quando o valor discrepante é igual a 10 e encontra-se no tratamento em que todos os fatores estão em seu nível “+”.

As Tabelas 5.6, 5.7, 5.8 e 5.9 mostram as médias dos tamanhos dos *gaps* nos gráficos de probabilidade dos cenários simulados. Estes valores fornecem evidência adicional aos resultados discutidos no início desta seção. Por exemplo, nota-se que o tamanho dos *gaps* aumenta à medida que o valor discrepante aumenta. Percebe-se também que, enquanto o gráfico semi-normal mantém o tamanho de seu *gap* aproximadamente constante, para cada caso específico, o gráfico normal é sensível à posição do valor discrepante no tratamento + + +: quanto maior o valor discrepante nessa posição maior será a redução do *gap*, podendo chegar a menos da metade do tamanho nos experimentos com 8 e 16 provas, os quais mostraram ser os mais afetados pela presença do valor discrepante. É nítido que os experimentos com 32 e 64 provas são mais robustos à presença de tais valores.

Tabela 5.6: Tamanho dos *gaps* no gráfico normal para o valor discrepante no tratamento (1).

Número de provas	Tamanho do valor discrepante			
	0	3	5	10
2^3	0,471	0,761	1,333	3,691
2^4	0,161	0,210	0,324	1,144
2^5	0,055	0,064	0,082	0,231
2^6	0,020	0,021	0,024	0,042

Tabela 5.7: Tamanho dos *gaps* no gráfico normal para o valor discrepante no tratamento + + +.

Número de provas	Tamanho do valor discrepante			
	0	3	5	10
2^3	0,471	0,490	0,555	1,551
2^4	0,161	0,214	0,268	0,449
2^5	0,055	0,065	0,083	0,170
2^6	0,020	0,022	0,024	0,044

Tabela 5.8: Tamanho dos *gaps* no gráfico semi-normal para o valor discrepante no tratamento (1).

Número de provas	Tamanho do valor discrepante			
	0	3	5	10
2^3	0,116	0,193	0,415	1,558
2^4	0,039	0,052	0,083	0,409
2^5	0,014	0,016	0,020	0,063
2^6	0,005	0,005	0,006	0,010

Tabela 5.9: Tamanho dos *gaps* no gráfico semi-normal para o valor discrepante no tratamento + + +.

	Número de provas	Tamanho do valor discrepante			
		0	3	5	10
	2^3	0,116	0,195	0,410	1,549
	2^4	0,039	0,053	0,085	0,408
	2^5	0,014	0,016	0,020	0,064
	2^6	0,005	0,005	0,006	0,011

Considere agora a razão entre os tamanhos dos *gaps* nos gráficos normal e semi-normal, apresentada nas Tabelas 5.10 e 5.11. Percebe-se que, em média, o tamanho do *gap* do gráfico normal é aproximadamente quatro vezes o tamanho do *gap* do gráfico semi-normal nos casos em que o valor discrepante é pequeno ou o número de provas é superior a 16. Essa razão tende a diminuir à medida que o tamanho do valor discrepante aumenta, mas permanece aproximadamente constante na presença de pequenos valores discrepantes ou nos experimentos robustos.

Tabela 5.10: Razão entre os tamanhos dos *gaps* nos gráficos normal e semi-normal com o valor discrepante no tratamento (1).

	Número de provas	Tamanho do valor discrepante		
		3	5	10
	2^3	3,947	3,209	2,369
	2^4	4,069	3,922	2,798
	2^5	3,995	4,037	3,685
	2^6	4,022	3,923	4,077

Tabela 5.11: Razão entre os tamanhos dos *gaps* nos gráficos normal e semi-normal com o valor discrepante no tratamento + + +.

	Número de provas	Tamanho do valor discrepante		
		3	5	10
	2^3	2,503	1,354	1,001
	2^4	4,000	3,143	1,101
	2^5	4,149	4,111	2,659
	2^6	3,956	4,035	4,116

No que diz respeito à influência do valor discrepante sobre a estimativa dos efeitos, uma característica observada é que, embora os sinais “+” e “-” dos valores discrepantes não apresentem diferenças acentuadas na influência do tamanho dos *gaps*, percebe-se que eles interferem na estimativa dos efeitos. Seja Δ o tamanho do valor discrepante (no caso das simulações, $\Delta = \pm 3, \pm 5, \pm 10$). Considere, inicialmente, o caso em que o

valor discrepante está presente na observação associada ao tratamento em que todos os fatores estão em seu nível “+”. Se Δ for positivo, então todos os efeitos estimados aumentam $\frac{\Delta}{2^{(k-1)}}$. Caso Δ seja negativo, então todos os efeitos diminuem $\frac{\Delta}{2^{(k-1)}}$. Agora, se o valor discrepante estiver em qualquer outra posição, então metade dos efeitos é aumentada $\frac{\Delta}{2^{(k-1)}}$ e a outra metade diminui $\frac{\Delta}{2^{(k-1)}}$. A influência dos sinais neste caso é a seguinte: os efeitos que aumentam quando o valor discrepante é positivo, diminuem quando este é negativo, e vice-versa.

Dois Valores Discrepantes Presentes

No caso de dois valores discrepantes presentes nos dados, são considerados dois casos nas simulações: (1) os dois valores discrepantes têm a mesma magnitude; (2) os dois têm magnitudes diferentes. As magnitudes consideradas para estes cenários são $\Delta = 3, 5, 10$. Os valores negativos não são escolhidos pois, conforme se notou no caso anterior, eles produzem os mesmos resultados para os tamanhos dos *gaps*. Deseja-se também verificar a influência da posição dos valores discrepantes sob a seguinte metodologia: um valor discrepante será mantido fixo no tratamento (1), ao passo que o outro “percorrerá” o tratamento *a* bem como o tratamento em que todos os fatores estão em seu nível “+”. Como há três opções de tamanhos para cada um dos dois valores discrepantes, e dois casos considerados (mesma magnitude e magnitudes diferentes), há um total de 18 combinações diferentes a serem analisadas para cada um dos quatro números de prova considerados (8, 16, 32, 64). A escolha pelos tratamentos (1) e *a* é apenas ilustrativa, uma vez que, conforme discutido no caso anterior, a única posição que altera o tamanho dos *gaps*, e isso apenas para o gráfico normal, é quando o valor discrepante posiciona-se no tratamento “+++”. Para as outras posições, o tamanho dos *gaps* permanece aproximadamente o mesmo.

As Tabelas 5.12, 5.13 e 5.14 mostram os tamanhos dos *gaps* nos gráficos normal e semi-normal nas configurações acima mencionadas. A análise dessas tabelas permite listar algumas conclusões:

- Nos casos em que dois valores discrepantes estão presentes nos dados, nota-se que, de um modo geral, o tamanho dos *gaps* é consideravelmente menor quando comparado com o caso de um único valor discrepante. A redução no tamanho

dos *gaps* torna mais difícil a identificação visual da presença destes valores discrepantes nos dados por meio dos gráficos de probabilidade. Isso é uma indicação de que tais gráficos podem não ser úteis na identificação de mais de um valor discrepante nos dados (BOX, 1990; LAWSON, 2008). Apenas nas situações em que o tamanho de tais valores é grande, principalmente no caso em que o valor 10 está presente, a identificação pode ser percebida mais claramente. Apesar disso, usando apenas os gráficos normal ou semi-normal, não é possível identificar se há mais de um valor discrepante, uma vez que apenas um único *gap* continua aparecendo, com tamanhos semelhantes ao caso de apenas um valor discrepante presente.

- Assim como aconteceu no caso anterior, para todas as simulações, o *gap* formado no gráfico normal mostra-se destacadamente maior que o formado no gráfico semi-normal. Isso reforça a indicação de que, no gráfico normal, a percepção da presença de valores atípicos pode ser mais clara.
- No caso do **gráfico semi-normal** algumas particularidades podem ser destacadas:

Mantém a característica, já observada, de que seu *gap* não é influenciado ou modificado se o valor discrepante estiver na observação associada ao tratamento “+ + +”;

Quando os dois valores discrepantes têm a mesma magnitude, independente se 3, 5 ou 10, o tamanho do *gap* é aproximadamente o mesmo;

Quando os valores discrepantes são 3 e 5, independente da posição em que estejam, o *gap* tem tamanhos bem parecidos. O mesmo acontece com as combinações 3 e 10; 10 e 5.

- Considerando o **gráfico normal** nota-se que:

Mesmo quando os valores discrepantes têm a mesma magnitude, o *gap* não preserva o mesmo valor. Isso é explicado pelo fato, já observado anteriormente, de o tamanho do *gap* ser reduzido quando o valor discrepante encontra-se associado ao tratamento “+ + +”;

Nos casos em que o valor discrepante tem tamanho 10, o gráfico apresenta os maiores valores para o *gap*. Mais uma vez a indicação de que, quanto maior o valor discrepante, melhor a habilidade gráfica em identificá-lo.

Portanto, pode-se concluir que os gráficos de probabilidade são úteis em detectar a presença de um único valor discrepante, e se este for de grande magnitude (principalmente de tamanhos a partir de $10\sigma_{obs}$).

Tabela 5.12: Tamanhos dos *gaps* nos gráficos normal e semi-normal nas simulações de fatoriais 2^3 na presença de dois valores discrepantes.

Posição dos valores discrepantes	<i>gap</i> - gráfico normal	<i>gap</i> - gráfico semi-normal
(1) = 3, $a = 3$	0,784	0,179
(1) = 3, $abc = 3$	0,683	0,175
(1) = 5, $a = 5$	0,964	0,183
(1) = 5, $abc = 5$	0,774	0,186
(1) = 10, $a = 10$	1,272	0,183
(1) = 10, $abc = 10$	0,974	0,187
(1) = 3, $a = 5$	0,998	0,234
(1) = 3, $abc = 5$	0,603	0,228
(1) = 3, $a = 10$	2,717	1,031
(1) = 3, $abc = 10$	1,048	1,018
(1) = 5, $a = 3$	0,992	0,231
(1) = 5, $abc = 3$	1,071	0,229
(1) = 5, $a = 10$	1,832	0,603
(1) = 5, $abc = 10$	0,711	0,598
(1) = 10, $a = 3$	2,706	1,031
(1) = 10, $abc = 3$	3,652	1,035
(1) = 10, $a = 5$	1,825	0,594
(1) = 10, $abc = 5$	3,427	0,604

Tabela 5.13: Tamanhos dos *gaps* nos gráficos normal e semi-normal nas simulações de fatoriais 2^4 na presença de dois valores discrepantes.

Posição dos valores discrepantes	<i>gap</i> - gráfico normal	<i>gap</i> - gráfico semi-normal
(1) = 3, $a = 3$	0,235	0,058
(1) = 3, $abcd = 3$	0,235	0,056
(1) = 5, $a = 5$	0,283	0,068
(1) = 5, $abcd = 5$	0,283	0,069
(1) = 10, $a = 10$	0,303	0,069
(1) = 10, $abcd = 10$	0,298	0,071
(1) = 3, $a = 5$	0,288	0,071
(1) = 3, $abcd = 5$	0,280	0,071
(1) = 3, $a = 10$	0,827	0,257
(1) = 3, $abcd = 10$	0,365	0,255
(1) = 5, $a = 3$	0,288	0,071
(1) = 5, $abcd = 3$	0,290	0,072
(1) = 5, $a = 10$	0,559	0,152
(1) = 5, $abcd = 10$	0,326	0,149
(1) = 10, $a = 3$	0,821	0,256
(1) = 10, $abcd = 3$	1,051	0,252
(1) = 10, $a = 5$	0,557	0,149
(1) = 10, $abcd = 5$	0,816	0,146

Tabela 5.14: Tamanhos dos *gaps* nos gráficos normal e semi-normal nas simulações de fatoriais 2^5 na presença de dois valores discrepantes.

Posição dos valores discrepantes	<i>gap</i> - gráfico normal	<i>gap</i> - gráfico semi-normal
(1) = 3, $a = 3$	0,070	0,017
(1) = 3, $abcde = 3$	0,070	0,017
(1) = 5, $a = 5$	0,090	0,022
(1) = 5, $abcde = 5$	0,090	0,023
(1) = 10, $a = 10$	0,108	0,026
(1) = 10, $abcde = 10$	0,106	0,026
(1) = 3, $a = 5$	0,085	0,021
(1) = 3, $abcde = 5$	0,086	0,021
(1) = 3, $a = 10$	0,190	0,049
(1) = 3, $abcde = 10$	0,165	0,050
(1) = 5, $a = 3$	0,083	0,021
(1) = 5, $abcde = 3$	0,084	0,021
(1) = 5, $a = 10$	0,148	0,037
(1) = 5, $abcde = 10$	0,149	0,038
(1) = 10, $a = 3$	0,186	0,048
(1) = 10, $abcde = 3$	0,202	0,050
(1) = 10, $a = 5$	0,147	0,037
(1) = 10, $abcde = 5$	0,158	0,038

Com relação à influência de dois valores discrepantes, digamos Δ_1 e Δ_2 , sobre a estimativa dos efeitos, de acordo com as Tabelas 5.15, 5.16, 5.17 e 5.18, que ilustram os resultados para as simulações de fatoriais 2^3 , pode-se notar que:

- **Quando os dois têm a mesma magnitude:** (1) Se nenhum deles estiver na posição “+++”, tem-se: metade dos efeitos permanece inalterada e, dos restantes, metade aumenta $\frac{\Delta_1+\Delta_2}{2^{k-1}}$ e a outra metade diminui $\frac{\Delta_1+\Delta_2}{2^{k-1}}$; (2) Se um deles estiver na posição “+++”, então metade dos efeitos estimados permanecem inalterados, e a outra metade aumenta $\frac{\Delta_1+\Delta_2}{2^{k-1}}$.
- **Quando os dois têm magnitudes diferentes:** (1) Se nenhum deles estiver na posição “+++”, tem-se: 25% aumentam $\frac{\Delta_1+\Delta_2}{2^{k-1}}$, 25% diminuem $\frac{\Delta_1+\Delta_2}{2^{k-1}}$, 25% aumentam $\frac{\Delta_1-\Delta_2}{2^{k-1}}$ e 25% diminuem $\frac{\Delta_1-\Delta_2}{2^{k-1}}$; (2) Se um deles estiver na posição “+++”, então, metade dos efeitos aumenta $\frac{\Delta_1+\Delta_2}{2^{k-1}}$ e a outra metade aumenta $\frac{\Delta_1-\Delta_2}{2^{k-1}}$.

Tabela 5.15: Influência de dois valores discrepantes de mesmo tamanho sobre os efeitos de fatoriais 2^3 . Tais valores encontram-se nas posições (1) e a ; (1) e abc .

Efeitos	valor: 3	valor: 3	valor: 5	valor: 5	valor: 10	valor: 10
<i>A</i>	0,011	-0,003	0,007	0,015	-0,002	-0,002
<i>B</i>	-1,513	0,013	-2,509	-0,007	-5,011	-0,006
<i>AB</i>	-1,511	0,007	-2,491	0,001	-5,015	0,004
<i>C</i>	-0,006	1,498	0,002	2,482	0,012	5,009
<i>AC</i>	0,014	1,500	-0,003	2,493	0,003	5,010
<i>BC</i>	1,497	1,511	2,500	2,500	4,988	5,005
<i>ABC</i>	0,005	0,004	0,009	0,010	-0,007	0,005

Tabela 5.16: Influência de dois valores discrepantes de tamanhos diferentes sobre os efeitos de fatoriais 2^3 .

Efeitos	(1) = 3, $a = 5$	(1) = 3, $abc = 5$	(1) = 3, $a = 10$	(1) = 3, $abc = 10$
<i>A</i>	0,496	0,505	1,751	1,733
<i>B</i>	-1,999	0,495	-3,256	1,760
<i>AB</i>	-2,001	0,501	-3,253	1,736
<i>C</i>	-0,493	1,994	-1,747	3,247
<i>AC</i>	-0,509	1,994	-1,760	3,267
<i>BC</i>	2,009	1,996	3,282	3,270
<i>ABC</i>	0,504	0,504	1,747	1,733

Tabela 5.17: Influência de dois valores discrepantes de tamanhos diferentes sobre os efeitos de fatoriais 2^3 .

Efeitos	(1) = 5, $a = 3$	(1) = 5, $abc = 3$	(1) = 5, $a = 10$	(1) = 5, $abc = 10$
<i>A</i>	-0,483	-0,500	1,246	1,257
<i>B</i>	-2,018	-0,500	-3,753	1,256
<i>AB</i>	-2,004	-0,479	-3,736	1,258
<i>C</i>	0,486	1,988	-1,234	3,746
<i>AC</i>	0,491	1,984	-1,258	3,743
<i>BC</i>	1,999	1,996	3,748	3,749
<i>ABC</i>	-0,502	-0,498	1,255	1,247

Tabela 5.18: Influência de dois valores discrepantes de tamanhos diferentes sobre os efeitos de fatoriais 2^3 .

Efeitos	(1) = 10, $a = 3$	(1) = 10, $abc = 3$	(1) = 10, $a = 5$	(1) = 10, $abc = 5$
<i>A</i>	-1,749	-1,770	-1,245	-1,258
<i>B</i>	-3,244	-1,750	-3,754	-1,273
<i>AB</i>	-3,253	-1,753	-3,760	-1,235
<i>C</i>	1,742	3,248	1,237	3,762
<i>AC</i>	1,746	3,248	1,237	3,746
<i>BC</i>	3,252	3,261	3,746	3,755
<i>ABC</i>	-1,750	-1,750	-1,261	-1,264

Capítulo 6

Considerações Finais

6.1 Conclusão

Em 1959, Cuthbert Daniel desenvolveu um método inédito para a análise de experimentos fatoriais não replicados: o gráfico semi-normal dos efeitos. Com o passar do tempo, o gráfico normal também começa a ser usado com os mesmos objetivos do gráfico semi-normal (DANIEL, 1976; BOX; MEYER, 1986). Isso gerou uma controvérsia entre os pesquisadores de tais experimentos, e a escolha de qual técnica usar tornou-se uma questão subjetiva, na maioria dos casos investigados. Tendo em vista que os planos fatoriais não replicados são frequentes em experimentação, e que os gráficos de probabilidade, normal e semi-normal, são as ferramentas mais utilizadas para a análise destes planos, a necessidade de um estudo que pudesse indicar, objetivamente, as reais habilidades e utilidades destes gráficos na análise de tais experimentos era premente. Foi exatamente isso que propôs esta dissertação. O estudo experimental comparativo discutido no Capítulo 5 fornece algumas respostas às perguntas de investigação elaboradas na Introdução desta dissertação.

Inicialmente, não há um procedimento absolutamente melhor que outro. As duas técnicas gráficas apresentaram vantagens e limitações, dependendo do caso considerado. Sendo assim, a decisão de qual gráfico usar dependerá da situação que se deseja analisar no experimento. Por exemplo, no caso em que se deseja julgar a significância dos efeitos, o gráfico semi-normal mostrou-se melhor, uma vez que a distância do efeito à reta é consideravelmente maior neste gráfico do que no gráfico normal, facilitando assim o seu julgamento. Por conta disso, a análise feita através do gráfico semi-normal talvez

possa deixar mais claro o julgamento de determinado efeito de tamanho pequeno. Na situação em que se deseja detectar a presença de valores discrepantes nos dados, o recomendado é o gráfico normal, pois, o tamanho do *gap* característico para esses casos é consideravelmente maior em tal gráfico do que no gráfico semi-normal. Apesar disso, uma vantagem que o gráfico semi-normal mostra nessa situação é o fato do *gap*, em tal gráfico, não depender da posição em que os valores discrepantes se encontram. Embora não se tenha apresentado neste trabalho nenhum procedimento formal para comparar os dois gráficos no caso da identificação das parcelas subdividas inadvertidas, pelos poucos artigos encontrados que tratam do assunto, a indicação é usar o gráfico normal, uma vez que a forma de S característica do gráfico normal para essa situação é mais claramente perceptível.

No tocante às restrições no uso de cada gráfico, percebe-se que, não apenas o gráfico semi-normal, mas também para o gráfico normal, foram identificadas algumas limitações: estes são úteis apenas se a quantidade de efeitos significativos for pequena; além de poucos, os efeitos têm que ter tamanhos consideráveis, afim de que possam ser visivelmente identificados pelos gráficos; sua capacidade de detectar valores discrepantes é comprometida quando há mais de um presente, ou quando sua magnitude é pequena. Além disso, tendo apenas como base alguns poucos artigos analisados, ainda não ficou muito evidente qual a configuração do gráfico semi-normal na presença de restrições na aleatorização.

Portanto, com o conhecimento adquirido através desta pesquisa, tanto por meio dos exemplos discutidos dos artigos, quanto pelos resultados das simulações realizadas, a recomendação feita é que, ao analisar experimentos fatoriais 2^k e 2^{k-p} não replicados com o uso das técnicas gráficas que estão sendo consideradas aqui, o pesquisador utilize o gráfico normal e o gráfico semi-normal simultaneamente, em vez de optar pela escolha de uma ou de outra técnica, pois, conforme discutido, cada uma delas possui vantagens e limitações que, juntas, podem contribuir para uma análise subjetiva mais completa dos dados.

6.2 Pesquisas Futuras

O campo do uso dos gráficos de probabilidade normal nos experimentos fatoriais não replicados ainda é bem vasto. Muito mais pode ser feito. A seguir, segue algumas indicações para pesquisas futuras, muitas das quais surgiram durante o estudo dos artigos selecionados, bem como dos resultados que se observava nas simulações; mas que não puderam ser feitas, quer por fugirem do objetivo central deste trabalho, quer devido às limitações de tempo:

- Estender o levantamento bibliográfico para outros importantes periódicos, tais como o *CIS - Current Index of Statistics*.
- Investigar as configurações do centro e dos extremos dos gráficos de probabilidade.
- Desenvolver um teste de comparação dos gráficos usando envelopes.
- Ajustar uma curva suave para a distribuição empírica como um ajuste alternativo para o gráfico normal.
- Construir um teste de comparação das habilidades dos gráficos no caso de restrições na aleatorização (*inadvertent plot-splitting*).
- Verificar se blocos negligenciados mostrariam algum padrão nos gráficos de probabilidade e, em caso afirmativo, comparar esse padrão.
- Estudar a distribuição dos *gaps*.
- Estudar o caso da mistura crítica (*critical mix*).
- Usar o método das distâncias para comparar o desempenho dos gráficos no julgamento dos efeitos na presença de valores discrepantes.
- Verificar se a escolha do \hat{p}_j tem alguma influência nos gráficos.
- Comparar as estimativas de σ^2 feitas pelo gráfico normal e pelo gráfico semi-normal.

Referências Bibliográficas

- [1] ADDELMAN, Sidney. Irregular fractions of the 2^n factorial experiments. *Technometrics*, v. 3, n. 4, p. 479-496, 1961.
- [2] ADDELMAN, Sidney. Orthogonal main-effect plans for asymmetrical factorial experiments. *Technometrics*, v. 4, n. 1, p. 21-46, 1962.
- [3] ADDELMAN, Sidney. Designs for the sequential application of factors. *Technometrics*, v. 6, n. 4, p. 365-370, 1964.
- [4] AGUIRRE-TORRES, Víctor; PÉREZ-TREJO, M. Esther. Outliers and the use of the rank transformation to detect active effects in unreplicated 2^f experiments. *Communications in Statistics - Simulation*, v. 30, n. 3, p. 637-663, 2001.
- [5] AL-SHIHA, A. A.; YANG, Shie-Shien. A multistage procedure for analyzing unreplicated factorial experiments. *Biometrical Journal*, v. 41, n. 6, p. 659-670, 1999.
- [6] ANSCOMBE, F. J.; TUKEY, J. W. The examination and analysis of residuals. *Technometrics*, v. 5, n. 2, p. 141-160, 1963.
- [7] ANTONY, J. Improving the wire bonding process quality using statistically designed experiments. *Microelectronics Journal*, v. 30, p. 161-168, 1999.
- [8] ARVIDSSON, M.; KAMMERLIND, P.; HYNÉN, A.; BERGMAN, B. Identification of factors influencing dispersion in split-plot experiments. *Journal of Applied Statistics*, v.28. n. 3-4, 269-283, 2001.
- [9] ATKINSON, A. C. Two graphical displays for outlying and influential observations in regression. *Biometrika*, v. 68, n. 1, p. 13-20, 1981.

- [10] ATKINSON, A. C. Regression diagnostics, transformations and constructed variables. *Journal of the Royal Statistical Society - Series B (Methodological)*, v. 44, n. 1, p. 1-36, 1982.
- [11] ATKINSON, A. C. Diagnostic regression analysis and shifted power transformations. *Technometrics*, v. 25, n. 1, p. 23-33, 1983.
- [12] AUER, C., KUNERT, J. On a heuristic analysis of highly fractionated 2^n factorial experiments. *Metrika*, v. 63, p. 43-54, 2005.
- [13] BACON-SHONE, J.; FUNG, W. K. A new graphical method for detecting single and multiple outliers in univariate and multivariate data. *Applied Statistics*, v. 36, n. 2, p. 153-162, 1987.
- [14] BECKMAN, R. J.; COOK, R. D. Outliers. *Technometrics*, v. 25, n. 2, p. 119-149, 1983.
- [15] BERGMAN, Bo; HYNÉN, Anders. Dispersion effects from unreplicated designs in the 2^{k-p} series. *Technometrics*, v. 39, n. 2, p. 191-198, 1997.
- [16] BINGHAM, Derek; SITTER, Randy R. Minimum-aberration two-level fractional factorial split-plot designs. *Technometrics*, v. 41, n. 1, p. 62-70, 1999.
- [17] BIRNBAUM, A. On the analysis of factorial experiments without replication. *Technometrics*, v. 1, n. 4, p. 343-357, 1959.
- [18] BISGAARD, Søren; FULLER, Howard T., BARRIOS, Ernesto. Two-level factorials run as split-plot experiments. *Report Series in Quality and Productivity - University of Wisconsin-Madison*, n. 136, 1995.
- [19] BISGAARD, Søren; VIVACQUA, Carla Almeida; PINHO, André L. S. de. Quality Quandaries: Not All Models Are Polynomials! *Quality Engineering*, v. 17, n. 1, p. 181-186, 2004.
- [20] BLOM, Gunnar. *Probability and statistics: theory and applications*. Springer texts in statistics. New York: Springer-Verlag, 1989.
- [21] BOX, G. E. P. The R. A. Fisher memorial lecture, 1988: quality improvement: an expanding domain for the application of scientific method. *Philosophical Transactions of the Royal Society. Series A - Mathematical, Physical & Engineering Sciences*, n. 327, p. 617-630, 1989.

- [22] BOX, G. E. P. Finding bad values in factorial designs. *Report Series in Quality and Productivity - University of Wisconsin-Madison*, n. 57, 1990.
- [23] BOX, G. E. P.; HUNTER, J. S.; HUNTER, W. G. *Statistics for experimenters: design, innovation, and discovery*. 2. ed. Hoboken, N.J.: Wiley-Interscience, 2005.
- [24] BOX, G. E. P.; MEYER, R. D. An analysis for unreplicated fractional factorials. *Technometrics*, v. 28, n. 1, p. 11-18, 1986.
- [25] BOX, G. E. P.; MEYER, R. D.; STEINBERG, David M. Follow-Up designs to resolve confounding in multifactor experiments. *Technometrics*, v. 38, n. 4, p. 303-313, 1996.
- [26] BRADU, D; HAWKINS, D. M. Location of multiple outliers in two-way tables, using tetrads. *Technometrics*, v. 24, n. 2, p. 103-108, 1982.
- [27] CLEVELAND, William S. Research in Statistical Graphics. *Journal of the American Statistical Association*, v. 82, n. 398, p. 419-423, 1987.
- [28] COOK, R. D.; WANG, P. C. Transformations and influential cases in regression. *Technometrics*, v. 25, n. 4, p. 337-343, 1983.
- [29] COX, D. R. Interaction. *International Statistical Review*, v. 52, n. 1, p. 1-24, 1984.
- [30] COX, D. R.; LAUH, Elizabeth. A note on the graphical analysis of multidimensional contingency tables. *Technometrics*, v. 9, n. 3, p. 481-488, 1967.
- [31] DANIEL, C. Fractional replication in industrial research. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, v. 5, p. 87-98. Berkeley and Los Angeles: University of California Press, 1956.
- [32] DANIEL, C. Use of half-normal plots in interpreting factorial two-level experiments. *Technometrics*, v. 1, n. 4, p. 311-341, 1959.
- [33] DANIEL, C. Locating Outliers in Factorial Experiments. *Technometrics*, v. 2, n. 2, p. 149-156, 1960.

- [34] DANIEL, C. *Half-Normal Plots*. In: Encyclopedia of Statistical Sciences, edição de Samuel Kotz & Norman L. Johnson. v. 3. New York: J. Wiley & Sons, Inc., 1967.
- [35] DANIEL, C. *Applications of statistics to industrial experimentation*. New York: J. Wiley & Sons, Inc., 1976.
- [36] DEAN, Angela; VOSS, Daniel. *Design and Analysis of Experiments*. New York: Springer, 1999.
- [37] DODGSON, Jeffrey H. A graphical method for assessing mean squares in saturated fractional designs. *Journal of Quality Technology*, v. 35, n. 2, p. 206-212, 2003.
- [38] DONG, Fang. On the identification of active contrasts in unreplicated fractional factorials. *Statistica Sinica*, v. 3, p. 209-217, 1993.
- [39] DYKSTRA, O. Partial Duplication of Factorial Experiments. *Technometrics*, v. 1, n. 1, p. 63-75, 1959.
- [40] ELANDT, Regina C. The folded normal distribution: two methods of estimating parameters from moments. *Technometrics*, v. 3, n. 4, p. 551-562, 1961.
- [41] EVANS, D. A. The influence of computers on the teaching of statistics. *Journal of the Royal Statistical Society. Series A - General*, v. 136, n. 2, p. 153-190, 1973.
- [42] FEDER, Paul I. Graphical techniques in statistical data analysis-tools for extracting information from data. *Technometrics*, v. 16, n. 2, p. 287-299, 1974.
- [43] FLACK, Virginia F.; FLORES, Rafael A. Using simulated envelopes in the evaluation of normal probability plots of regression residuals. *Technometrics*, v. 31, n. 2, p. 219-225, 1989.
- [44] FREENY, Anne E.; LANDWEHR, James M. Graphical analysis for a large designed experiment. *Technometrics*, v. 37, n. 1, p. 1-14, 1995.
- [45] GAN, F. F.; KOEHLER, K. J. Goodness-of-fit tests based on P-P Probability Plots. *Technometrics*, v. 32, n. 3, p. 289-303, 1990.

- [46] GENTLEMAN, J. F.; WILK, M. B. Detecting outliers in a two-way table: I. statistical behavior of residuals. *Technometrics*, v. 17, n. 1, p. 1-14, 1975.
- [47] GERSON, Marion. The techniques and uses of probability plotting. *The Statistician*, v. 24, n. 4, p. 235-257, 1975.
- [48] GORMAN, J. W.; TOMAN, R. J. Selection of variables for fitting equations to data. *Technometrics*, v. 8, n. 1, p. 27-51, 1966.
- [49] GOUPY, Jacques. Factorial experimental design: Detecting an outlier with the dynamic variable and the Daniel's diagram. *Chemometrics and Intelligent Laboratory Systems*, v. 80, p. 156 – 166, 2006.
- [50] HAMADA, M.; BALAKRISHNAN, N. Analyzing unreplicated factorial experiments: a review with some new proposals. *Statistica Sinica*, v. 8, p. 1-41, 1998.
- [51] HARTEER, H. LEON. Expected values of normal order statistics. *Biometrika*, v. 48, n. 1/2, p. 151-165, 1961.
- [52] HEALY, M. J. R. Multivariate normal plotting. *Applied Statistics*, v. 17, n. 2, p. 157-161, 1968.
- [53] HOLM, Sture; MARK, Sigyn; ADOLFSSON, Tobias. A step-down test for effects in unreplicated factorial designs. *Communications in Statistics — Theory and Methods*, v. 34, p. 405-416, 2005.
- [54] HOLMGREN, Eric B. The P-P Plot as a method for comparing treatment effects. *Journal of the American Statistical Association*, v. 90, n. 429, p. 360-365, 1995.
- [55] HOLMS, A. G.; BERRETTONY, J. N. Chain-Pooling ANOVA for two-level factorial replication-free experiments. *Technometrics*, v. 11, n. 4, p. 725-746, 1969.
- [56] HUND, Edelgard; MASSART, D. Luc; SMEYERS-VERBEKE, Johanna. Robust regression and outlier detection in the evaluation of robustness tests with different experimental designs. *Analytica Chimica Acta*, n. 463, p. 53-73, 2002.
- [57] HUNTER, J. S. Sequential factorial estimation. *Technometrics*, v. 6, n. 1, p. 41-55, 1964.

- [58] ILZARBE, Laura; ÁLVAREZ, María Jesús; VILES, Elisabeth; TANCO, Martín. Practical applications of Design of Experiments in the field of engineering: A Bibliographical Review. *Quality and Reliability Engineering International*, v. 24, p. 417–428, 2008.
- [59] JOHNSON, Richard A.; BHATTACHARYYA, Gouri K. *Statistics: principles e methods*. 2. ed. New York: John Wiley & Sons, Inc., 1992.
- [60] KEMPTHORNE, Oscar; HINKELMANN, Klaus. *Design and Analysis of Experiments - Volume 1: Introduction to Experimental Design*. 2. ed. New York : John Wiley & Sons, Inc., 2008.
- [61] KUEHL, Robert O. *Design of experiments: statistical principles of research design and analysis*. 2. ed. New York : Thomson Learning, 2000.
- [62] KULAHCI, Murat; BISGAARD, Søren. The use of Plackett–Burman designs to construct split-plot designs. *Technometrics*, v. 47, n. 4, p. 495-501, 2005.
- [63] LABRECQUE, John. Goodness-of-fit tests based on nonlinearity in probability plots. *Technometrics*, v. 19, n. 3, p. 293-306, 1977.
- [64] LAWSON, J. SAS macros for analysis of unreplicated 2^k and 2^{k-p} designs with a possible outlier. *Journal of Statistical Software*, v. 25, n. 13, p. 1-17, 2008.
- [65] LAWSON, J.; GRIMSHAW, Scott; BURT, Jason. A quantitative method for identifying active contrasts in unreplicated factorial designs based on the half-normal plot. *Computational Statistics & Data Analysis*, v. 26, p. 425-436, 1998.
- [66] LENTH, R. V. Quick and easy analysis of unreplicated factorials. *Technometrics*, v. 31, n. 4, p. 469-473, 1989.
- [67] LEONE, F. C.; NELSON, L. S.; NOTTINGHAM, R. B. The folded normal distribution. *Technometrics*, v. 3, n. 4, p. 543-550, 1961.
- [68] LOEPPKY, J. L.; SITTER, R. R. Analyzing unreplicated blocked or split-plot fractional factorial designs. *Journal of Quality Technology*, v. 34, n. 3, p. 229-243, 2002.
- [69] LOUGHIN, T. M. Calibration of the Lenth test for unreplicated factorial designs. *Journal of Quality Technology*, v. 30, n. 2, p. 171-175, 1998.

- [70] MARGOLIN, B. H. Systematic methods for analyzing $2^n 3^m$ factorial experiments with applications. *Technometrics*, v. 9, n. 2, p. 245-259, 1967.
- [71] MARGOLIN, B. H. Design and analysis of factorial experiments via interactive computing in APL. *Technometrics*, v. 18, n. 2, p. 135-150, 1976.
- [72] MEEKER, William Q.; ESCOBAR, Luis A. *Statistical methods for reliability data*. New York : John Wiley & Sons, Inc., 1998.
- [73] MILLER, Arden. Strip-plot configurations of fractional factorials. *Technometrics*, v. 39, n. 2, p. 153-161, 1997.
- [74] MILLER, Arden. The analysis of unreplicated factorial experiments using all possible comparisons. *Technometrics*, v. 47, n. 1, p. 51-63, 2005.
- [75] MILLIKEN, George A.; SHI, Xiaoyu; MENDICINO, Michael; VASUDEV, P. K. Strip-plot design for two-step processes. *Quality and Reliability Engineering International*, v. 14, p. 197-210, 1998.
- [76] MONTGOMERY, Douglas C. *Design and analysis of experiments*. 5. ed. New York : John Wiley & Sons, Inc., 2001.
- [77] MONTGOMERY, Douglas C; KULAHCI, Murat; ALMIMI, Ashraf A. Estimation of missing observations in two-level split-plot designs. *Quality and Reliability Engineering International*, v. 24, p. 127-152, 2007.
- [78] OLGUIN, J.; FEARN, T. A new look at half-normal plots for assessing the significance of contrasts for unreplicated factorials. *Applied Statistics*, v. 46, n. 4, p. 449-462, 1997.
- [79] PASTERNAK, B; LIUZZI, A. Patterns in residuals: a test for regression model adequacy in radionuclide assay. *Technometrics*, v. 7, n. 4, p. 603-621, 1965.
- [80] PEWSEY, Arthur. Large-sample inference for the general half-normal distribution. *Communications in Statistics-Theory and Methods*, v. 31, n. 7, p. 1045-1054, 2002.
- [81] ROHATGI, V. K. *An Introduction to Probability Theory and Mathematical Statistics*. New York: J. Wiley & Sons, Inc., 1976.

- [82] SCHOEN, E. D. Designing fractional two-level experiments with nested error structures. *Journal of Applied Statistics*, v. 26, n. 4, p. 495-508, 1999.
- [83] SCHOEN, E. D. Dispersion-effects detection after screening for location effects in unreplicated two-level experiments. *Journal of Statistical Planning and Inference*, v. 126, p. 289-304, 2004.
- [84] SCHOEN, E. D.; KAUL, E. A. A. Three robust scale estimators to judge unreplicated experiments. *Journal of Quality Technology*, v. 32, n. 3, p. 276-283, 2000.
- [85] SCHOEN, E. D.; WOLFF, Kirsten. Design and analysis of a fractional $4^1 3^1 2^5$ split-plot experiment. *Journal of Applied Statistics*, v. 24, n. 4, p. 409-419, 1997.
- [86] SITTER, Randy R.; BINGHAM, Derek. Fractional factorial split-plot designs for robust parameter experiments. *Technometrics*, v. 45, n. 1, p. 80-89, 2003.
- [87] SITTER, Randy R.; LOEPPKY, J. L. Analyzing unreplicated blocked or split-plot fractional factorial designs. *Journal of Quality Technology*, v. 34, n. 3, p. 229-243, 2002.
- [88] STEFANSKY, Wilhelmine. Rejecting outliers by maximum normed residual. *The Annals of Mathematical Statistics*, v. 42, n. 1, p. 35-45, 1971.
- [89] STEINBERG, D. M.; BISGAARD, S. Technometrics: how it all started. *Technometrics*, v. 50, n. 1, p. 2-7, 2008.
- [90] STIRLING, W. Douglas. Enhancements to aid interpretation of probability plots. *The Statistician*, v. 31, n. 3, p. 211-220, 1982.
- [91] TAYLOR, G. A. R. Analysis of experiments by using half-normal plots. *The Statistician*, v. 43, n. 4, p. 529-536, 1994.
- [92] UPTON, Graham J. G.; COOK, Ian. *A Dictionary of Statistics*. 2. ed. New York: Oxford University Press, 2006.
- [93] VARIYATH, A. M.; ABRAHAM, B; CHEN, J. Analysis of performance measures in experimental designs using jackknife. *Journal of Quality Technology*, v. 37, n. 2, p. 91-100, 2005.

- [94] VENTER, J. H.; STEEL, S. J. A hypothesis-testing approach toward identifying active contrasts. *Technometrics*, v. 38, n. 2, p. 161-169, 1996.
- [95] VIVACQUA, Carla Almeida; PINHO, André Luis Santos de. *Notas de Aula de Planejamento de Experimentos*. 2008.
- [96] VOSS, D. T.; WANG, W. Simultaneous confidence intervals in the analysis of orthogonal saturated designs. *Journal of Statistical Planning and Inference*, v. 81, p. 383-392, 1999.
- [97] WEBB, Steve R. Saturated sequential factorial designs. *Technometrics*, v. 10, n. 3, p. 535-550, 1968.
- [98] WEBSTER, J. T. *Factorial experiments*. In: Encyclopedia of Statistical Sciences, edição de Samuel Kotz, Norman L. Johnson. 2. ed. v. 4, Hoboken, N.J.: Wiley-Interscience, 2006.
- [99] WILK, M. B.; GNANADESIKAN, R. Graphical methods for internal comparisons in multiresponse experiments. *The Annals of Mathematical Statistics*, v. 35, n. 2, pp. 613-631, 1964.
- [100] WILK, M. B.; GNANADESIKAN, R. Probability plotting methods for the analysis of data. *Biometrika*, v. 55, n. 1, pp. 1-17, 1968.
- [101] WILK, M. B.; GNANADESIKAN, R. Use of maximum likelihood for estimating error variance from a collection of analysis of variance mean squares. *The Annals of Mathematical Statistics*, v. 41, n. 1, pp. 292-304, 1970.
- [102] WILK, M. B.; GNANADESIKAN, R.; HUYETT, M. M. J. Probability Plots for the Gamma Distribution. *Technometrics*, v. 4, n. 1, pp. 1-20, 1962.
- [103] WU, C. F. Jeff; HAMADA, Michael. *Experiments: planning, analysis, and parameter design optimization*. New York : John Wiley & Sons, Inc., 2000.
- [104] YE, J. On measuring and correcting the effects of data mining and model selection. *Journal of the American Statistical Association*, v. 93, n. 441, p. 120-131, 1998.
- [105] YE, K. Q.; HAMADA, M. Critical values of the Lenth method for unreplicated factorial designs. *Journal of Quality Technology*, v. 32, n. 1, p. 57-66, 2000.

- [106] YE, K. Q.; HAMADA, M; WU, C. F. J. A Step-Down Lenth Method for analyzing unreplicated factorial designs. *Journal of Quality Technology*, v. 33, n. 2, p. 140-152, 2001.
- [107] ZAHN, Douglas A. Modifications of and revised critical values for the half-normal plot. *Technometrics*, v. 17, n. 2, p. 189-200, 1975.
- [108] ZAHN, Douglas A. An empirical study of the half-normal plot. *Technometrics*, v. 17, n. 2, p. 201-211, 1975.

Apêndice A

Programação

Construção do Gráfico Normal

Programa para a construção do gráfico normal. Na verdade, trata-se de uma pequena modificação na função `qqnorm()`: incluir o comando `identify()` para identificar os efeitos no gráfico, a fim de que se pudesse compará-lo com os efeitos julgados significativos no gráfico semi-normal.

```
newqqnorm <-  
function (y, ylim, labs = as.character(1:length(y)), xlab = "Quantis Normais",  
        ylab = "Efeitos Estimados", plot.it = TRUE, datax = FALSE,  
        ...)  
{  
  if (has.na <- any(ina <- is.na(y))) {  
    yN <- y  
    y <- y[!ina]  
  }  
  if (0 == (n <- length(y)))  
    stop("y is empty or has only NAs")  
  if (plot.it && missing(ylim))  
    ylim <- range(y)  
  x <- qqnorm(ppoints(n))[order(order(y))]  
  if (has.na) {  
    y <- x  
    x <- yN  
    x[!ina] <- y  
    y <- yN  
  }  
}
```

```

if (plot.it)
  if (datax)
    plot(y, x, xlab = ylab, ylab = xlab,
         xlim = ylim, ...)
  else plot(x, y, xlab = xlab, ylab = ylab,
           ylim = ylim, ...)
invisible(if (datax) list(x = y, y = x) else list(x = x,
          y = y))
identify(x, y, labels=labs)
}

```

Simulações

A função `teste()` é o programa criado para realizar as simulações usadas nesta dissertação. Com ela, se pode testar os dois critérios de comparação gráfica propostos neste trabalho: (1) calcular as distâncias dos efeitos à reta, a fim de avaliar a habilidade dos gráficos de julgar a significância dos efeitos; (2) calcular o tamanho dos *gaps* formados nos gráficos no caso da presença de valores discrepantes. Há uma versão da função `teste()` para cada número de provas utilizados no trabalho: $2^3 = 8$, $2^4 = 16$, $2^5 = 32$ e $2^6 = 64$. Por uma questão de simplicidade, mostra-se apenas o `teste3()`, que é a função para gerar os cenários simulados de um fatorial 2^3 . Para os outros casos, a programação é análoga.

```

##### CASO COM TRÊS FATORES #####
teste3 <-
function(media = 0, efA = 0, efB = 0, efC = 0, efAB = 0, efAC = 0, efBC = 0,
efABC = 0, um = 0, a = 0, b = 0, ab = 0, c = 0, ac = 0, bc = 0, abc = 0, n = 5,
despad=1)
{
library(BHH2)
library(faraway)
si <- ffFullMatrix(ffDesMatrix(3)[,1:3],x=c(1,2,3),maxInt=3)$Xa
efeitos <- NULL
efeitos.out <- NULL
distanciafull <- NULL
distanciafull.out <- NULL
distanciahalf <- NULL
distanciahalf.out <- NULL

```

```

dif.distancias <- NULL
dif.distancias.out <- NULL
resposta <- NULL
resposta.out <- NULL
gap.half <- NULL
gap.half.out <- NULL
gap.full <- NULL
gap.full.out <- NULL
diferenca.gaps <- NULL
diferenca.gaps.out <- NULL
for (i in 1:n){
y <- si%*%((0.5)*matrix(c(2*media, efA, efB, efAB, efC, efAC, efBC, efABC))) +
rnorm(8, 0, despad)
y1 <- y + matrix(c(um, a, b, ab, c, ac, bc, abc))
m <- t(matrix(c(rep(0.125, 8), rep(0.25, 56)), nrow=8, ncol=8)*si)%*%y
m1 <- t(matrix(c(rep(0.125, 8), rep(0.25, 56)), nrow=8, ncol=8)*si)%*%y1
efeitos1 <- c(m[2], m[3], m[4], m[5], m[6], m[7], m[8])
efeitos2 <- c(m1[2], m1[3], m1[4], m1[5], m1[6], m1[7], m1[8])
quanorm <-
function(y)
{
n <- length(y)
x <- qnorm(ppoints(n))[order(order(y))]
return(x)
}
qnormal <- quanorm(efeitos1)
y2 <- (2*despad/sqrt(8))*qnormal
p <- matrix(sort(y2))
p1 <- data.frame(efeitos = c("A", "B", "AB", "C", "AC", "BC", "ABC"),
estimativas=efeitos1)
p2 <- p1[order(p1$estimativas),]
p3 <- data.frame(efeitos=p2$efeitos, distâncias=abs(p2[,2]-p))
p4 <- p3[order(p3$efeitos),]
p5 <- data.frame(efeitos = c("A","B","C","AB","AC","BC","ABC"),
distâncias=c(p4[1,2], p4[5,2], p4[7,2], p4[2,2], p4[4,2], p4[6,2], p4[3,2]))
qnormal.out <- quanorm(efeitos2)
y2.out <- (2*despad/sqrt(8))*qnormal.out
p.out <- matrix(sort(y2.out))
p1.out <- data.frame(efeitos = c("A", "B", "AB", "C", "AC", "BC", "ABC"),

```

```

estimativas=efeitos2)
p2.out <- p1.out[order(p1.out$estimativas),]
p3.out <- data.frame(efeitos=p2.out$efeitos, distâncias=abs(p2.out[,2]-p.out))
p4.out <- p3.out[order(p3.out$efeitos),]
p5.out <- data.frame(efeitos = c("A","B","C","AB","AC","BC","ABC"),
distâncias=c(p4.out[1,2], p4.out[5,2], p4.out[7,2], p4.out[2,2], p4.out[4,2],
p4.out[6,2], p4.out[3,2]))
quahalf <-
function(x)
{
  x <- abs(x)
  labord <- order(x)
  x <- sort(x)
  i <- order(x)
  n <- length(x)
  ui <- qnorm((n + 1:n)/(2 * n + 1))
  return(ui)
}
efeitos3 <- (abs(efeitos1))
qhalf <- quahalf(efeitos3)
y3 <- (2*despad/sqrt(8))*qhalf
p6 <- matrix(sort(y3))
p7 <- data.frame(efeitos = c("A", "B", "AB", "C", "AC", "BC", "ABC"),
estimativas=efeitos3)
p8 <- p7[order(p7$estimativas),]
p9 <- data.frame(p8, distâncias=abs(p8[,2]-p6))
p10 <- p9[order(p9$efeitos),]
p11 <- data.frame(efeitos = c("A","B","C","AB","AC","BC","ABC"),
distâncias=c(p10[1,2], p10[5,2], p10[7,2], p10[2,2], p10[4,2], p10[6,2],
p10[3,2]))
efeitos3.out <- (abs(efeitos2))
qhalf.out <- quahalf(efeitos3.out)
y3.out <- (2*despad/sqrt(8))*qhalf.out
p6.out <- matrix(sort(y3.out))
p7.out <- data.frame(efeitos = c("A", "B", "AB", "C", "AC", "BC", "ABC"),
estimativas=efeitos3.out)
p8.out <- p7.out[order(p7.out$estimativas),]
p9.out <- data.frame(p8.out, distâncias=abs(p8.out[,2]-p6.out))
p10.out <- p9.out[order(p9.out$efeitos),]

```

```

p11.out <- data.frame(efeitos = c("A","B","C","AB","AC","BC","ABC"),
distâncias=c(p10.out[1,2], p10.out[5,2], p10.out[7,2], p10.out[2,2],
p10.out[4,2], p10.out[6,2], p10.out[3,2]))
m1 <- matrix(c(p5[1,2], p5[2,2], p5[3,2], p5[4,2], p5[5,2], p5[6,2], p5[7,2]))
m1.out <- matrix(c(p5.out[1,2], p5.out[2,2], p5.out[3,2], p5.out[4,2],
p5.out[5,2], p5.out[6,2], p5.out[7,2]))
m2 <- matrix(c(p11[1,2], p11[2,2], p11[3,2], p11[4,2], p11[5,2],
p11[6,2], p11[7,2]))
m2.out <- matrix(c(p11.out[1,2], p11.out[2,2], p11.out[3,2], p11.out[4,2],
p11.out[5,2], p11.out[6,2], p11.out[7,2]))
m3 <- m2 - m1
m3.out <- m2.out - m1.out
o <- sort(efeitos1)
gaphalf1 <- min(sort(abs(efeitos1)))
gapfull1 <-
if (min(efeitos1) < 0)
  {if (max(efeitos1) >= 0)
    {if (min(o[o >= 0])==0)
      {min(o[o >= 0])}
      else {min(o[o > 0]) - max(o[o < 0])}}
    } else {abs(max(efeitos1))}
  } else {min(efeitos1)}
o1 <- sort(efeitos2)
gaphalf2 <- min(sort(abs(efeitos2)))
gapfull2 <-
if (min(efeitos2) < 0)
  {if (max(efeitos2) >= 0)
    {if (min(o1[o1 >= 0])==0)
      {min(o1[o1 >= 0])}
      else {min(o1[o1 > 0]) - max(o1[o1 < 0])}}
    } else {abs(max(efeitos2))}
  } else {min(efeitos2)}
dif.gaps <- gapfull1 - gaphalf1
dif.gaps.out <- gapfull2 - gaphalf2
efeitos <- cbind(efeitos, efeitos1)
efeitos.out <- cbind(efeitos.out, efeitos2)
distanciafull <- cbind(distanciafull, m1)
distanciafull.out <- cbind(distanciafull.out, m1.out)
distanciahalf <- cbind(distanciahalf, m2)

```

```

distanciahalf.out <- cbind(distanciahalf.out, m2.out)
dif.distancias <- cbind(dif.distancias, m3)
dif.distancias.out <- cbind(dif.distancias.out, m3.out)
resposta <- cbind(resposta, y)
resposta.out <- cbind(resposta.out, y1)
gap.half <- cbind(gap.half, gaphalf1)
gap.half.out <- cbind(gap.half.out, gaphalf2)
gap.full <- cbind(gap.full, gapfull1)
gap.full.out <- cbind(gap.full.out, gapfull2)
diferenca.gaps <- cbind(diferenca.gaps, dif.gaps)
diferenca.gaps.out <- cbind(diferenca.gaps.out, dif.gaps.out)
}
media1 <- matrix(c(mean(distanciafull[1,]), mean(distanciafull[2,]),
mean(distanciafull[3,]), mean(distanciafull[4,]), mean(distanciafull[5,]),
mean(distanciafull[6,]), mean(distanciafull[7,])))
media2 <- matrix(c(mean(distanciafull.out[1,]), mean(distanciafull.out[2,]),
mean(distanciafull.out[3,]), mean(distanciafull.out[4,]),
mean(distanciafull.out[5,]), mean(distanciafull.out[6,]),
mean(distanciafull.out[7,])))
media3 <- matrix(c(mean(distanciahalf[1,]), mean(distanciahalf[2,]),
mean(distanciahalf[3,]), mean(distanciahalf[4,]), mean(distanciahalf[5,]),
mean(distanciahalf[6,]), mean(distanciahalf[7,])))
media4 <- matrix(c(mean(distanciahalf.out[1,]), mean(distanciahalf.out[2,]),
mean(distanciahalf.out[3,]), mean(distanciahalf.out[4,]),
mean(distanciahalf.out[5,]), mean(distanciahalf.out[6,]),
mean(distanciahalf.out[7,])))
media5 <- matrix(c(mean(dif.distancias[1,]), mean(dif.distancias[2,]),
mean(dif.distancias[3,]), mean(dif.distancias[4,]), mean(dif.distancias[5,]),
mean(dif.distancias[6,]), mean(dif.distancias[7,])))
media6 <- matrix(c(mean(dif.distancias.out[1,]), mean(dif.distancias.out[2,]),
mean(dif.distancias.out[3,]), mean(dif.distancias.out[4,]),
mean(dif.distancias.out[5,]), mean(dif.distancias.out[6,]),
mean(dif.distancias.out[7,])))
write.table(efeitos, "D:/Backup 03-02-09/Acer/WT/efeitos.txt",
quote=F)
write.table(efeitos.out, "D:/Backup 03-02-09/Acer/WT/efeitos.out.txt",
quote=F)
write.table(distanciafull, "D:/Backup 03-02-09/Acer/WT/distanciafull.txt",
quote=F)

```

```

write.table(distanciafull.out, "D:/Backup 03-02-09/Acer/WT/distanciafull.out.txt",
quote=F)
write.table(distanciahalf, "D:/Backup 03-02-09/Acer/WT/distanciahalf.txt",
quote=F)
write.table(distanciahalf.out, "D:/Backup 03-02-09/Acer/WT/distanciahalf.out.txt",
quote=F)
write.table(dif.distancias, "D:/Backup 03-02-09/Acer/WT/dif.distancias.txt",
quote=F)
write.table(dif.distancias.out, "D:/Backup 03-02-09/Acer/WT/dif.distancias.out.txt"
quote=F)
write.table(resposta, "D:/Backup 03-02-09/Acer/WT/resposta.txt", quote=F)
write.table(resposta.out,"D:/Backup 03-02-09/Acer/WT/resposta.out.txt",quote=F)
write.table(gap.half, "D:/Backup 03-02-09/Acer/WT/gap.half.txt", quote=F)
write.table(gap.half.out,"D:/Backup 03-02-09/Acer/WT/gap.half.out.txt",quote=F)
write.table(gap.full, "D:/Backup 03-02-09/Acer/WT/gap.full.txt", quote=F)
write.table(gap.full.out,"D:/Backup 03-02-09/Acer/WT/gap.full.out.txt",quote=F)
write.table(diferenca.gaps,"D:/Backup 03-02-09/Acer/WT/diferenca.gaps.txt",quote=F)
write.table(diferenca.gaps.out,"D:/Backup 03-02-09/Acer/WT/diferenca.gaps.out.txt",
print(data.frame(efeitos=c("A","B","C","AB","AC","BC","ABC"),
media.dist.semi=media3, media.dist.normal=media1, diferenca.medias=media5))
print(data.frame(efeitos=c("A","B","C","AB","AC","BC","ABC"),
media.dist.semi.out=media4, media.dist.normal.out=media2,
diferenca.medias.out=media6))
print(data.frame(mean(gap.full), mean(gap.half), mean(diferenca.gaps)))
print(data.frame(mean(gap.full.out),mean(gap.half.out),mean(diferenca.gaps.out)))
}

```

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)