

**TESTES COM ERROS FREQUENTISTAS
CONDICIONAIS E TESTES COM
INTERPRETAÇÃO BAYESIANA E
FREQUENTISTA CONDICIONAL**

CARLA REGINA GUIMARÃES BRIGHENTI

2007

CARLA REGINA GUIMARÃES BRIGHENTI

**TESTES COM ERROS FREQUENTISTAS CONDICIONAIS E
TESTES COM INTERPRETAÇÃO BAYESIANA E
FREQUENTISTA CONDICIONAL**

Tese apresentada à Universidade Federal de Lavras,
como parte das exigências do Programa de Pós-
graduação em Estatística e Experimentação
Agropecuária, para a obtenção do título de "Doutor"

Orientador:

Profº Dr. Lucas Monteiro Chaves

LAVRAS
MINAS GERAIS-BRASIL

2007

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Brighenti, Carla Regina Guimarães.

Testes com erros freqüentistas condicionais e testes com interpretação bayesiana e freqüentista condicional / Carla Regina Guimarães Brighenti. --
Lavras : UFLA, 2007.

210 p. : il.

Tese (Doutorado) – Universidade Federal de Lavras, 2007.

Orientador: Lucas Monteiro Chaves.

Bibliografia.

1. Razão de verossimilhanças. 2. Hipótese. 3. Significância. 4. Bayesiano.
5. Estatística condicionante. 6. Função particionante. 7. Região de não-decisão.
I. Universidade Federal de Lavras. II. Título.

CDD – 519.56

CARLA REGINA GUIMARÃES BRIGHENTI

**TESTES COM ERROS FREQUENTISTAS CONDICIONAIS E
TESTES COM INTERPRETAÇÃO BAYESIANA E
FREQUENTISTA CONDICIONAL**

Tese apresentada à Universidade Federal de Lavras,
como parte das exigências do Programa de Pós-
graduação em Estatística e Experimentação
Agropecuária, para a obtenção do título de "Doutor"

APROVADA em, 16 de Julho de 2007.

Prof^o. Dr. Carlos Henrique Osório Silva UFV-MG

Prof^o. Dr. Daniel Furtado Ferreira UFLA-MG

Prof^o. Dr. Mario Javier Ferrua Vivanco UFLA-MG

Prof^o. Dr. Rinaldo Artes IBMEC-SP

Prof^o Dr. Lucas Monteiro Chaves

DEX - UFLA

(Orientador)

LAVRAS

MINAS GERAIS - BRASIL

A Deus,

fonte de todas as graças e bênçãos derramadas ao longo da minha vida.

Ao meu eterno professor da vida, meu querido pai,

EURICO DE SOUZA GUIMARÃES (*in memoriam*), que mesmo não
estando mais entre nós continuou sendo meu maior inspirador.

OFEREÇO

“Apaixonar-se por Deus é o maior dos romances;
Procurá-lo, a maior das aventuras;
Encontrá-lo, a maior de todas as realizações.”
(Santo Agostinho)

A minha mãe, ZÉLIA DE OLIVEIRA GUIMARÃES, de quem recebo amor incondicional,
doação e apoio em todas as fases de minha vida,

Ao meu marido DEODORO MAGNO BRIGHENTI, guerreiro incansável, marido e pai
exemplar, o “marechal da minha vida”,

Aos meus filhos, LUCAS GUIMARÃES BRIGHENTI e
TIAGO GUIMARÃES BRIGHENTI,
meus gatinhos, que são as minhas maiores conquistas,

DEDICO

AGRADECIMENTOS

A Deus por me conceder o dom da fé.

Aos meus pais, Zélia e Eurico (*in memoriam*), pelos bons exemplos de perseverança e oração.

Ao meu especial Deodoro pelo amor, carinho, compreensão, apoio, dedicação, paciência e, principalmente, por ser um pai maravilhoso, para que eu não me sentisse angustiada nos momentos de ausência na educação dos filhos.

Aos meus familiares que torceram, rezaram e compreenderam minha ausência em muitos momentos.

Ao maninho Marcelo Ângelo Cirillo pelos “grilos” solucionados, as conversas de desabafo e o apoio de sempre.

Aos amigos que me incentivaram e ajudaram, seja no curso, com orações, e mesmo cuidando de meus filhos, em especial Juliana Vigilato, Dora, Ana, Lucimara, Imaculada, Sirene, Geraldino, Marcos Geraldo e Everton.

À Escola Preparatória de Cadetes do Ar pela liberação parcial das atividades, e principalmente às professoras Leila e Andréa pelo apoio.

Ao professor Lucas Monteiro Chaves pela orientação e pelas “doses homeopáticas” diárias de seu conhecimento.

Aos professores do Departamento de Ciências Exatas da Universidade Federal de Lavras pela confiança em mim depositada.

Aos meus ex-professores, Efraim Lázaro Reis, Valdir Mano, José Mauro, Fernando e Wellington, que um dia me deram uma carta de apresentação para que eu pudesse ter agora esta conquista, que também é deles.

A todos os colegas dos cursos de mestrado e doutorado pela partilha de conhecimentos.

Àqueles que porventura não foram aqui mencionados, mas sabem que fazem parte da minha vida, um carinho especial e muito obrigada.

BIOGRAFIA

Carla Regina Guimarães Brighenti, filha de Eurico de Souza Guimarães e Zélia de Oliveira Guimarães, nasceu em 11 de dezembro de 1975, em São João Del Rei, MG. Em março de 1993, iniciou o curso de Ciências na Fundação de Ensino Superior de São João del Rei, hoje UFSJ, diplomando-se em fevereiro de 1996 em Ciências – Licenciatura Curta e, em dezembro de 1997, em Química – Licenciatura Plena, onde também exerceu atividades de Monitoria. Concomitantemente realizou sua licenciatura em Matemática, diplomando-se em dezembro de 1996, na Universidade Vale do Rio Verde, em Três Corações, MG. No período de 1993 a 1997, foi docente no Ensino Médio das Redes Particular e Estadual de Ensino, nas disciplinas de Matemática e Química. No ano de 1997 prestou concurso público federal e ingressou como professora de Matemática na Escola Preparatória de Cadetes do Ar – EPCAR, em Barbacena/MG, local onde trabalha até os dias atuais. No ano de 1998 concluiu sua Pós-graduação “Lato sensu” em Docência Superior e ingressou no Mestrado em Agroquímica na Universidade Federal de Viçosa, em Viçosa, MG, sob a orientação do Prof. Efraim Lázaro Reis, onde exerceu atividades de Monitoria na área de Química Orgânica. Submeteu-se à defesa de sua dissertação na área de Química Analítica em agosto de 2000. Em 2001 realizou o curso de especialização em Matemática pela Universidade Federal de Viçosa, em Viçosa, MG. Em 2002 ingressou novamente no Mestrado, mas agora em Estatística e Experimentação Agropecuária, na Universidade Federal de Lavras, em Lavras, MG, sob a orientação do Prof. Lucas Monteiro Chaves, concluído em 2003, quando ingressou no Doutorado em Estatística e Experimentação Agropecuária. Durante todo o período de pós-graduação continuou trabalhando na EPCAr, mantendo apenas liberação parcial das atividades.

SUMÁRIO

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS	i
LISTA DE TABELAS.....	ii
LISTA DE FIGURAS.....	iii
RESUMO	vii
ABSTRACT	ix
CAPÍTULO 1: OS TESTES DE HIPÓTESES E DE SIGNIFICÂNCIA NO CONTEXTO FREQUENTISTA E BAYESIANO.....	01
1 Introdução Geral	03
2 Referencial Teórico	06
2.1. Teste de hipóteses	06
2.1.1. Teste de razão de verossimilhanças simples	09
2.1.2. Críticas à teoria dos testes de hipóteses de Neyman-Pearson	22
2.2. Teste de significância	24
2.2.1. Críticas ao teste de significância de Fisher	28
2.3. Teste de hipóteses no contexto bayesiano	35
2.3.1. Definições gerais	35
2.3.2. Medidas de evidência bayesiana usuais	40
2.3.3. Testes bayesianos apresentando medidas de evidência	42
3 Referências Bibliográficas	48
CAPÍTULO 2: TESTES DE HIPÓTESES COM ERROS FREQUENTISTAS CONDICIONAIS.....	51
Resumo	53
Abstract	55
1 Introdução	57
2 Referencial teórico	58
2.1. Testes frequentistas condicionais	58
2.1.1. Probabilidades de erros condicionais aos dados	61

2.2. Partição do espaço amostral	70
3 Metodologia	74
3.1. Partição contínua	74
3.2. Comparação entre estatísticas condicionantes.....	74
3.3. Região de não decisão em testes freqüentistas condicionais	75
3.4. Distribuição da estatística condicionante baseada em <i>valores-p</i>	75
3.5. Exemplo de Aplicação	75
4 Resultados e Discussão	76
4.1. Partição contínua	76
4.1.1. Partição com probabilidade de erros condicionais iguais	83
4.1.2. Partição contínua intrínseca	88
4.1.2.1. Razão de verossimilhança simétrica	90
4.1.3. Estatística condicionante ancilar	97
4.1.4. Partição baseada em <i>valores -p</i>	99
4.2. Comparação entre estatísticas condicionantes.....	103
4.2.1. Teste de Hipóteses com Razão de Verossimilhanças Simétrica	103
4.2.2. Teste de Hipóteses com Razão de Verossimilhanças Não – Simétrica	113
4.3. Região de não decisão em testes freqüentistas condicionais	130
4.4. Estudo da distribuição da estatística condicionante baseada em <i>valores-p</i>	136
4.5. Exemplo de Aplicação	140
5 Conclusões	145
6 Referências Bibliográficas	146
CAPÍTULO 3: TESTES COM INTERPRETAÇÃO FREQUENTISTA CONDICIONAL E BAYESIANA	147
Resumo	149
Abstract	151
1 Introdução	153
2 Referencial teórico	155

2.1. Teste freqüentista condicional e bayesiano unificado para hipóteses simples	155
2.1.1. Uma alteração com motivação bayesiana: a região de não-decisão.....	158
2.1.1.1. Construção da região de não-decisão	159
2.1.2. Região de não decisão em testes freqüentistas condicionais.....	165
2.2. Generalização do teste unificado	172
2.2.1. Interpretação freqüentista condicional do teste unificado generalizado	176
2.3. Comentários sobre o teste unificado	180
2.4. Considerações gerais.....	182
3 Metodologia	185
4 Resultados e Discussão	187
4.1. O teste unificado para $H_0: X \sim \text{Uniforme}(0,1)$ x $H_1: X \sim \text{Beta}(1/2,1)$	187
4.2. O teste unificado para hipóteses sobre o parâmetro da Exponencial	193
4.3. Estudo da região de não decisão no teste unificado com perdas assimétricas	196
4.4. Exemplo de aplicação	201
5 Conclusões	204
6 Referências Bibliográficas	205
ANEXOS	206

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

\mathcal{X}^n	Espaço Amostral
Θ	Espaço paramétrico
Υ	Teste de hipóteses
C_Υ	Região crítica do teste de hipóteses
δ	Função de decisão
L_i	Função de verossimilhança sob a hipótese H_i
$B(X)$	Razão de verossimilhanças
f_B^i	Função densidade de probabilidade da estatística $B(X)$ sob a hipótese H_i
F_B^i	Função distribuição de probabilidade acumulada de $B(X)$ sob a hipótese H_i
b_c	Valor crítico do teste de hipóteses em relação à distribuição da variável $B(X)$
α	Probabilidade de Erro Tipo I
β	Probabilidade de Erro Tipo II
α_x	Probabilidade de Erro Condicional Tipo I (PECI)
β_x	Probabilidade de Erro Condicional Tipo II (PECII)
Z	Estatística Condicionante
$H(x)$	Função Particionante
RND	Região de Não Decisão
π	Densidade <i>a priori</i>
ℓ	Razão entre as perdas
ρ	Razão entre as densidades <i>a priori</i>

LISTA DE TABELAS

Tabela	Página
1.1	Tipos de decisão e tipos de erro 8
1.2	Valores de b_c calculados para diferentes tamanhos de amostra com probabilidades de erros tipo I e tipo II iguais..... 22
1.3	Frequência de H_0 verdadeira no teste com $H_0: X \sim N(0,1)$ versus $H_1 \sim N(\theta,1)$ para <i>valor-p</i> no intervalo 0,049 a 0,05, mostrando a natureza não-freqüentista do <i>valor-p</i> , utilizando 10000 simulações com proporção fixada de 50 % de H_0 verdadeira..... 32
1.4	Função perda para teste de hipóteses simples x simples..... 35
2.1	Alguns possíveis valores para limitar as regiões de alta e baixa evidência e seus erros resultantes para $n=4$ e $\alpha_{ac}=0,1$ 67
2.2	Resultados da estatística condicionante Z com $B(x) = x/(7-x)$ 95
2.3	Probabilidades de erros condicionais, para alguns valores de x , de acordo com o valor crítico da decisão do respectivo teste, quando utilizadas diferentes estatísticas condicionantes.. 127
3.1	Probabilidades de erro Tipo I e II com e sem a região de não-decisão para o teste unificado para as hipóteses $H_0: X \sim \text{Uniforme}(0,1)$ x $H_1: X \sim \text{Beta}(1/2,1)$ 191
3.2	Valores críticos para o teste não-condicional de $X \sim \text{Exp}(1)$ versus $X \sim \text{Exp}(2)$ com diferentes tamanhos de amostra n e o limite inferior da RND..... 194
3.3	Probabilidades do erro Tipo I e II com e sem a inclusão da RND no teste unificado para hipóteses sobre o parâmetro da Exponencial..... 195

LISTA DE FIGURAS

Figura		Página
1.1	Gráficos exibindo as áreas correspondentes às probabilidades de erros tipos I e II iguais a 5% e 90,4%, respectivamente para o caso de $n = 1$	17
1.2	Gráficos exibindo as áreas correspondentes às probabilidades de erros tipos I e II iguais a 31,25% para o caso de $n = 1$	19
1.3	Gráficos exibindo as áreas correspondentes às probabilidades de erros tipos I e II iguais a 25,2% para o caso de $n = 4$	21
1.4	Função densidade de probabilidade sob a hipótese nula. Ambos α (a probabilidade de erro tipo I) e o <i>valor-p</i> são “áreas da cauda” sob esta curva. A área da cauda para α é fixada antes do experimento, enquanto que, a área da cauda do <i>valor-p</i> só é conhecida depois que um resultado é observado....	25
1.5	Representação das probabilidades <i>a posteriori</i> em função de razão de verossimilhanças.....	44
2.1	Representação, no espaço das realizações, de diferentes distâncias da amostra em relação à região crítica, tanto para rejeição (x_1, x_2) quanto para aceitação (x_3, x_4) da hipótese nula	59
2.2	Representação das regiões de decisão com diferentes níveis de confiança.....	62
2.3	Regiões de alta e baixa evidência para rejeição e aceitação de H_0 e respectivos erros.....	65
2.4	Desenho esquemático considerando as duas partições definidas por δ e por $Z = \{z_1, z_2\}$	71
2.5	Representação das áreas infinitesimais sob f_0	77
2.6	Desenho esquemático da relação entre as áreas associadas aos pontos x e x' em f_0 e f_1 com $k \in]x'+ \Delta x, x[$	78
2.7	Relação funcional entre x e x'	79
2.8	Simetria da Função $H(x)$ em relação à reta $y = x$	86
2.9	Probabilidade de erro condicional tipo I e II resultante da	87

	partição utilizando a função $H(x) = -\frac{2}{3} \ln\left(1 - e^{-\frac{3}{2}x}\right)$	
2.10	Relação entre as áreas sob f_0 e f_1 para a partição intrínseca.....	89
2.11	Probabilidades em $f_0 (\Delta)$ e $f_1 (\bullet)$ do lançamento de um dado..	94
2.12	Distribuição de $B(x) = x/(7-x)$ sob H_0 e sob H_1	96
2.13	Representação da área correspondente a $P_{\theta_i} \left[\max \left\{ B(x), \frac{1}{B(x)} \right\} \leq z \right]$	99
2.14	Desenho esquemático da relação entre as áreas associadas ao ponto “t” sob f_T^0 e f_T^1	100
2.15	Representação esquemática da estatística condicionante <i>valor-p</i>	101
2.16	Estatística Condicionante Ancilar associada a $B(x) = \frac{1-x}{x}$	104
2.17	Áreas correspondentes aos pontos x e $H(x) = 1-x$ para cálculo da probabilidade de erro condicional tipo I.....	107
2.18	Áreas correspondentes aos pontos x e $H(x) = 1-x$ para cálculo da probabilidade de erro condicional tipo II.....	108
2.19	Simetria da função $g(x) = [x(1-x)]^2$ em relação à reta $x = 0,5$.	110
2.20	Estatística Condicionante <i>valor-p</i> dada por $Z = \max \{1 - 6x^2 + 8x^3 - 3x^4, 4x^3 - 3x^4\}$	111
2.21	Funções Densidade de Probabilidades das Distribuições Uniforme (0,1) e Beta (1/2, 1).....	113
2.22	Estatística Condicionante $Z = \max \{2\sqrt{x}, 2 - 2\sqrt{x}\}$	115
2.23	Distribuição Acumulada da estatística ancilar Z , $F_Z(z) = z - 1$	116
2.24	Comportamento da estatística condicionante baseada nos níveis de significância intrínseca $Z = \max \left\{ 2\sqrt{x}, \frac{1}{2\sqrt{x}} \right\}$ em	119

	função dos dados.....	
2.25	Função Particionante $H(x)$ para o critério com probabilidades de erros condicionais iguais.....	122
2.26	Representação das áreas com condição do <i>valor-p</i> como condicionante.....	125
2.27	Probabilidades de Erro Condicional Tipo I e Tipo II para diferentes condicionamentos.....	128
2.28	Representação dos valores críticos resultantes das estatísticas condicionantes baseadas nos critérios de probabilidade de erros condicionais iguais e <i>valor-p</i>	132
2.29	Região de Aceitação, Rejeição e Não decisão para o teste da Uniforme (0,1) versus a Beta (1/2, 1).....	133
2.30	Representação dos valores críticos resultantes das estatísticas condicionantes com erros condicionais iguais e <i>valores-p</i>	134
2.31	Região de Aceitação, Rejeição e Não decisão para o teste da Uniforme (0,1) versus a Beta (1/2,1).....	135
2.32	Estatística Condicionante baseada em <i>valores-p</i> $Z = \max \{ F_0^T(T), 1 - F_1^T(T) \}$	138
2.33	Funções densidade de probabilidade de populações normais de média 1,05 e 1,46.....	141
2.34	Áreas correspondentes aos erros Tipo I e II sob as hipóteses nula e alternativa, respectivamente.....	143
3.1	Desenho esquemático da Região de Não Decisão para $b_c > 1$...	160
3.2	Exemplo da região de não-decisão (RND) quando $b_c > 1$	161
3.3	Esquema da construção da região de não-decisão para $b_c > 1$...	161
3.4	Gráfico de $H(z)$ para o caso (i) em que se tem $1 < b_c < H(1)$...	162
3.5	Esquema da região de não decisão para $b_c < 1$	162
3.6	Representação da região de não-decisão (RND) quando $b_c < 1$.	163
3.7	Esquema da construção da região de não-decisão para $b < 1$	164
3.8	Limites da região de não decisão em função da densidade de probabilidade.....	168

3.9	Limites da Região de Não Decisão em função da razão de verossimilhanças b	169
3.10	Probabilidades de erros Tipo I e II considerando o teste de $H_0: \theta = 1$ versus $H_1: \theta = 2$ da exponencial com valor crítico $k_1=1,367$	170
3.11	Probabilidades de erros Tipo I e II considerando o teste de $H_0: \theta = 2$ versus $H_1: \theta = 1$ da exponencial com valor crítico $k_1=3,877$	171
3.12	Região de não decisão para o teste de $H_0: \theta = 1$ versus $H_1: \theta = 2$	172
3.13	Esquema da construção da região de não-decisão no teste $\bar{Y}_{l,\rho}$ para $b_c > l\rho$	174
3.14	Esquema da Construção da Região de não decisão para $b_c > l\rho$	174
3.15	Esquema da Região de Não Decisão para $b_c > l\rho$	175
3.16	Comportamento da função particionante $H(z)$	177
3.17	Probabilidades de erro condicional associadas ao valor crítico $b_c = \rho l = 1,236$	188
3.18	Região de não-decisão para o teste unificado para as hipóteses $H_0: X \sim \text{Uniforme}(0,1)$ versus $H_1: X \sim \text{Beta}(1/2,1)$	190
3.19	Representação da Região de não-decisão em função da razão de verossimilhanças $B(x) = 2\sqrt{x}$	191
3.20	Regiões de Rejeição e Aceitação modificadas pela inclusão da RND no teste unificado para as hipóteses $H_0: X \sim \text{Uniforme}(0,1)$ versus $H_1: X \sim \text{Beta}(1/2,1)$	192
3.21	A região de não-decisão $]r,a[$ como uma função de l , para testar hipóteses sobre o parâmetro da Exponencial $H_0: \theta = 1$ contra $H_1: \theta = 2$ sob probabilidades <i>a priori</i> iguais com $n=1$	199
3.22	Função densidade de probabilidade e distribuição acumulada de $f_b^0(b) = \text{lognormal}(2,52; (3,89)^2)$ e $f_b^1(b) = \text{lognormal}(-2,523; (3,89)^2)$	202
3.23	Área correspondente a Região de não decisão.....	203

RESUMO

BRIGHENTI, CARLA REGINA GUIMARÃES. **TESTES COM ERROS FREQUENTISTAS CONDICIONAIS E TESTES COM INTERPRETAÇÃO BAYESIANA E FREQUENTISTA CONDICIONAL.** LAVRAS: UFLA, 2007. 210 p. (Tese – Doutorado em Estatística e Experimentação Agropecuária)*

Os testes de hipóteses apresentam probabilidades de erro independentes dos dados observados, o que é uma grande fonte de críticas. Nos testes de significância utiliza-se o *valor-p*, mas esta não é uma medida frequentista. Uma metodologia frequentista dependente dos dados foi formalizada por Kiefer em 1977. Nela se particiona o espaço amostral e desenvolvem-se medidas de erro frequentistas condicionais denominadas de probabilidades de erro condicionais (PEC's). O valor crítico e as PEC's de um teste dependem da partição utilizada. Berger, Brown e Wolpert (1994) perceberam que, no teste Bayesiano, para alguns valores específicos de perdas e densidades *a priori* esse seria também um legítimo teste frequentista condicional. Todavia, há casos em que podem ocorrer situações problemáticas e os autores propuseram a incorporação da “região de não-decisão”, RND, cuja função é impedir que se apresentem PEC's (ou probabilidades *a posteriori*) superiores a 0,5. No capítulo 1, objetivou-se detalhar a teoria dos testes de hipóteses e de significância, ressaltando suas principais críticas e diferenças. No capítulo 2, objetivou-se desenvolver detalhadamente toda a teoria matemática dos testes com erros frequentistas condicionais utilizando uma função $H(x)$, denominada aqui de “função particionante”, e analisar as PEC's quando utilizados quatro critérios de partição baseados na estatística ancilar, no “nível de significância intrínseco”, nas “probabilidades de erros condicionais iguais” e em *valores-p*. E, ainda, verificar a relação destas com a simetria da razão de verossimilhanças. No capítulo 3, os objetivos foram avaliar o efeito da incorporação da RND no teste unificado em alguns exemplos e estudar o seu comportamento no caso de perdas assimétricas. No caso de razão de verossimilhanças simétrica, todas as estatísticas condicionantes utilizadas concordaram. Em testes em que a razão de verossimilhança é não-simétrica ocorreram divergências. O condicionamento com PEC's iguais e o baseado em *valores-p* foram os melhores. Quanto ao tamanho da RND, concluiu-se, para os exemplos desenvolvidos, que ele é

* Orientador: Lucas Monteiro Chaves -UFLA.

dependente do tamanho da amostra e que determinados valores de perda l podem tornar o teste inadequado.

ABSTRACT

BRIGHENTI, CARLA REGINA GUIMARÃES. **TESTS WITH CONDITIONAL FREQUENTISTS ERRORS AND TESTS WITH INTERPRETATION BAYESIAN AND CONDITIONAL FREQUENTISTS.** LAVRAS: UFLA, 2007. 210 p. (Thesis - Doctorate in Statistics and Agricultural Experimentation)*

The Neyman-Pearson tests of hypotheses present probabilities of error independent of the observed data; this is an enormous source of critics. In the significance tests the p-value is used, but this is not a frequentist measuring. Kiefer (1977) formalized a frequentist methodology dependent of the data. The sample space is partitioned using a partitioning function and it is developed measure a conditional frequentist error, denominated conditional error probabilities (CEP). The critical value and CEP's of a test depend on the used partition. Berger, Brown and Wolpert (1994) observed that in the Bayesian test between two simple hypotheses, for certain types of losses function and priors densities, it would be a genuine conditional frequentist test too. But, there are some cases that some problematic situations can happen, so they proposed the incorporation of a "no-decision region", NDR, whose objective is to avoid the presence of the CEP's (or posterior probabilities) greater to 0,5. The objective of this work in chapter 1 was to review the details of the tests of hypothesis and test of significance, emphasizing the main differences and critics. In chapter 2, the objective was the development in details of all the mathematic theory of the tests with conditioning frequentists errors using one function $H(x)$, denominated "partitioning function" and to analyze the CEP's when it is used four criteria of partition based on the "ancillary" statistic, in the "intrinsic significance level", in the "equal probability continuum" and in "p-values" and also to examine the relation of them with the symmetry of the likelihood ratio. In chapter 3, the objectives were the exam of the effect of the NDR incorporation in the unified test in some examples and study its behavior in cases of asymmetrical losses. In the case of symmetry of the likelihood ratio, all the used conditioning statistics agreed. In the case when the likelihood ratio is asymmetrical, some divergences appeared. The conditioning with the equal conditional errors and based on p-values was the best. Regarding the size of the NDR, the conclusion was that the

* Advisor: Lucas Monteiro Chaves -UFLA.

developed examples are dependent of the size of the sample, and that some kind of losses l , can make the test inadequate.

CAPÍTULO 1

OS TESTES DE HIPÓTESES E DE SIGNIFICÂNCIA NO CONTEXTO FREQUENTISTA E BAYESIANO

1 INTRODUÇÃO

A Inferência Estatística tem por objetivo fazer afirmações sobre características de uma população com base em informações obtidas por dados amostrais. Em geral, essas afirmações traduzem-se em problemas fundamentais da inferência estatística como a estimação de parâmetros, mas este não é um dos problemas tratados neste trabalho.

Pode-se também dispor de determinadas hipóteses sobre a característica da população em questão, e o que se deseja é testar as hipóteses e decidir pela aceitação ou rejeição de uma delas em favor de outra, determinando-se um teste de hipóteses. No contexto freqüentista este teste de hipóteses apresenta, como probabilidade de erro, um valor definido pré-experimentalmente, que independe dos dados observados, o que é problemático e alvo de críticas.

No contexto bayesiano os testes de hipóteses são associados a algumas medidas de evidência como o fator de Bayes ou a probabilidade *a posteriori*, que são medidas dependentes dos dados. Contudo surgem outras críticas, como, por exemplo, a subjetividade da escolha das densidades *a priori*.

Um teste complementar, muito utilizado por freqüentistas, é o teste de significância em que é calculado o *valor-p*. Mas esta medida é apenas uma medida de evidência, que não tem interpretação freqüentista como probabilidade de erro.

Sendo assim, apesar de os testes de hipóteses, testes de significância e testes de hipóteses bayesianos serem os testes mais utilizados na inferência estatística, nenhum deles apresenta uma probabilidade de erro que tenha interpretação freqüentista dependente dos dados observados, e cada um deles apresenta vantagens e desvantagens questionáveis por adeptos de segmentos contrários.

Um outro teste formalizado no final da década de 70 foi denominado “teste freqüentista condicional”. Este teste foi desenvolvido com o intuito de “aperfeiçoar” os testes de hipóteses, atrelando a estes uma medida de evidência dependente dos dados que tivesse uma interpretação freqüentista.

Infelizmente, o método freqüentista condicional ainda não atingiu popularidade substancial, em parte por causa da dificuldade de se escolher uma estatística “adequada” na qual condicionar o teste e, por outro lado, acredita-se também que a matemática envolvida no procedimento proposto inicialmente por Kiefer (1977) não tenha sido totalmente esclarecedora, merecendo, por isso, especial atenção neste trabalho.

Tendo em vista o exposto apresentou-se, no capítulo 1 deste trabalho, uma apresentação teórica e uma breve discussão, com as referentes críticas, sobre os testes de hipóteses, de significância e de hipóteses no contexto bayesiano.

O objetivo do capítulo 2 foi estudar detalhadamente a construção de um teste com erros freqüentistas condicionais para o teste de hipótese simples versus simples, desenvolvendo-se passo a passo a teoria matemática envolvida e utilizando, para isso, o conceito aqui denominado de função particionante. O objetivo desse trabalho foi estudar as partições e as probabilidades de erros apresentadas quando utilizadas diferentes estatísticas condicionantes e a relação destas com a simetria da razão de verossimilhanças. Tentou-se explicar também, detalhadamente, os conceitos de estatística condicionante e probabilidade de erro condicional necessários à construção dos testes freqüentistas condicionais. Essa construção foi desenvolvida e exemplificada nos resultados, por meio de duas situações, com ou sem simetria da razão de verossimilhanças. Verificou-se o comportamento de quatro diferentes estatísticas condicionantes, com especial atenção para a estatística condicionante baseada em *valores-p*, efetuando-se um estudo sobre sua distribuição de probabilidade.

No capítulo 3 apresenta-se o teste unificado freqüentista-condicional e bayesiano proposto por Berger, Brown e Wolpert (1994) e, com os problemas que podem surgir devido às diferenças de interpretação do teste, discute-se a introdução de uma região de não decisão. A construção dessa região é feita primeiramente para o caso simples de densidades *a priori* iguais e perdas iguais; depois, essa construção é generalizada. Embora a teoria se torne muito extensa, a metodologia resultante é na verdade muito simples; de fato pode ser muito mais simples do que a metodologia de teste não condicional comum e tem como vantagem a interpretação dual freqüentista-bayesiana.

2 REFERENCIAL TEÓRICO

2.1 TESTE DE HIPÓTESES

A Inferência Estatística tem por objetivo fazer afirmações sobre características de uma população com base em informações obtidas por dados amostrais. Em geral, essas afirmações traduzem-se nos dois problemas básicos da inferência estatística: *Estimação* e *Teste de Hipóteses*. Neste trabalho, será visto o problema estatístico de *Teste de Hipóteses*.

Definindo Θ como espaço paramétrico, uma hipótese estatística H é simplesmente a especificação de um subconjunto de Θ . Se uma hipótese estatística determina totalmente a distribuição, isto é, especifica um subconjunto com apenas um ponto de Θ , é dita uma hipótese simples; caso contrário, hipótese composta.

Um teste de uma hipótese estatística, denotado por Y , é uma regra ou procedimento para decidir quando rejeitar uma determinada hipótese H .

Um teste Y de uma hipótese estatística H é dito aleatorizado se Y é definido por uma função ψ_Y , que é a probabilidade de rejeição da hipótese em função dos dados observados, ou seja,

$$\psi_Y(x_1, \dots, x_n) = P_Y [H \text{ ser rejeitada} | (x_1, \dots, x_n) \text{ foi observado}].$$

A função $\psi_Y(\cdot, \dots, \cdot)$ é denominada *função crítica* do teste Y . Observe que para efetivamente tomar a decisão é necessário um mecanismo aleatório externo ao experimento. De fato, qualquer função $\psi_Y(\cdot, \dots, \cdot)$ com domínio em \mathfrak{X}^n e contradomínio no intervalo $[0,1]$ é uma função crítica possível e define um teste aleatorizado.

Em geral, um teste é definido da seguinte maneira: Rejeite H se e somente se a amostra $(x_1, \dots, x_n) \in C_Y$, em que C_Y é um subconjunto do espaço amostral das observações. Tal teste Y é denominado um teste não-aleatorizado e C_Y a *região crítica ou de rejeição* do teste Y ; portanto, um teste não-aleatorizado de H é uma decomposição do espaço amostral das realizações \mathcal{X}^n em C_Y e \bar{C}_Y (complementar de C_Y), de modo que se $(x_1, \dots, x_n) \in C_Y$, H é rejeitada. Assim, um teste não-aleatorizado fica determinado quando especificamos sua região crítica. Desta forma, pode-se dizer que um teste não-aleatorizado é um caso particular de um teste aleatorizado cuja função crítica é dada por:

$$\psi_Y(x_1, \dots, x_n) = \begin{cases} 1 & \text{se } (x_1, \dots, x_n) \in C_Y \\ 0 & \text{caso contrário} \end{cases}$$

Em muitos problemas duas hipóteses são discutidas: uma delas será denominada hipótese nula, denotada por $H_0: \{\theta \in \Theta_0\}$, e a outra hipótese é denominada alternativa, denotada por $H_1: \{\theta \in \bar{\Theta}_0\}$, com $\Theta_0 \cup \bar{\Theta}_0 = \Theta$. (Mood, Graybill & Boes, 1974).

Ao tomar a decisão de aceitar ou recusar uma hipótese existem dois tipos de erros que se podem cometer. Pode-se rejeitar a hipótese nula H_0 quando de fato ela é verdadeira, erro tipo I; ou pode-se aceitar H_0 quando de fato ela é falsa, erro tipo II (Tabela 1.1). O tamanho do erro tipo I é definido como sendo a probabilidade de se cometer esse erro; do mesmo modo, o tamanho do erro tipo II é a probabilidade de se cometer o erro tipo II.

O ideal é que ambas as probabilidades dos erros sejam pequenas, mas geralmente, ao se tentar diminuir a probabilidade do erro tipo I, não se consegue controlar o valor da probabilidade do erro tipo II. Este dilema é resolvido

tentando-se obter testes que possuam um valor predeterminado de probabilidade de erro tipo I e o menor valor possível para a probabilidade de erro tipo II.

TABELA 1.1 - Tipos de decisão e tipos de erro

Realidade	Decisão	
	Aceitar a hipótese nula H_0	Rejeitar a hipótese nula H_0
H_0 verdadeira	Decisão correta	Erro Tipo I
H_0 falsa	Erro Tipo II	Decisão correta

A função poder do teste Υ , denotada por $\pi_\Upsilon(\theta)$, é definida como sendo a probabilidade de que H_0 seja rejeitada se a distribuição da qual a amostra foi obtida for parametrizada por θ , ou seja, $\pi_\Upsilon(\theta) = P_\theta[\text{rejeitar } H_0]$. Se Υ é um teste não-aleatorizado, então $\pi_\Upsilon(\theta) = P_\theta[(X_1, \dots, X_n) \in C_\Upsilon]$, em que C_Υ é a região crítica do teste Υ . Se Υ é um teste aleatorizado com função crítica $\psi_\Upsilon(\cdot, \dots, \cdot)$, então

$$\pi_\Upsilon(\theta) = P_\theta[\text{rejeitar } H_0] = \int \dots \int P_\theta[\text{rejeitar } H_0 | (x_1, \dots, x_n)] f_{X_1, \dots, X_n}(x_1, \dots, x_n; \theta) \prod_{i=1}^n dx_i$$

$$\pi_\Upsilon(\theta) = \int \dots \int \psi_\Upsilon(x_1, \dots, x_n) f_{X_1, \dots, X_n}(x_1, \dots, x_n; \theta) \prod_{i=1}^n dx_i = E_\theta[\psi_\Upsilon(X_1, \dots, X_n)].$$

O *tamanho do teste* Υ é definido como sendo $\sup_{\theta \in \Theta_0} [\pi_\Upsilon(\theta)] = \alpha$. Note que, no caso de hipótese simples, $H_0: \Theta_0 = \{\theta_0\}$ então, $\pi_\Upsilon(\theta_0) = \alpha$.

Um teste Υ^* de $H_0: X \sim f_0(\cdot; \theta_0)$ contra $H_1: X \sim f_1(\cdot; \theta_1)$, sendo θ_0 e θ_1 conhecidos, é dito ser um teste mais poderoso de tamanho α ($0 < \alpha < 1$) se e somente se:

$$(i) \pi_{\Upsilon^*}(\theta_0) = \alpha.$$

(ii) $\pi_{\gamma^*}(\theta_1) \geq \pi_{\gamma}(\theta_1)$ para qualquer outro teste γ para o qual $\pi_{\gamma}(\theta_0) \leq \alpha$.

Ou seja, um teste γ^* é mais poderoso de tamanho α se ele possui o tamanho de seu erro tipo I igual a α e tem menor tamanho do erro tipo II sobre todos os outros testes com tamanho de erro tipo I igual a α ou menor.

A justificativa para se fixar o tamanho do erro tipo I é que, normalmente, as hipóteses são elaboradas de forma que a hipótese nula seja mais importante do que a hipótese alternativa; sendo assim, a ocorrência do erro tipo I, recusar a hipótese nula sendo esta verdadeira, é mais grave do que o erro em se recusar a hipótese alternativa sendo esta verdadeira. Fixa-se, então, α em uma probabilidade pequena (frequentemente atribuído como 0,05 ou 0,01).

É importante ressaltar que, segundo Mood, Graybill e Boes (1974), muitos autores usam os termos “tamanho do teste” e “nível de significância” de forma equivalente. No entanto, deve-se utilizar o termo “nível de significância” para os “testes de significância”, um tipo de inferência estatística que está intimamente relacionado com os testes de hipóteses (tão intimamente ligados que muitos autores não fazem distinção) e será abordado no item 2 deste capítulo.

Neste trabalho serão considerados os testes entre hipótese nula simples contra hipótese alternativa simples. Esse caso não é muito comum na prática estatística, mas grande parte da teoria de testes de hipóteses pode ser estudada por meio dele.

2.1.1 TESTE DE RAZÃO DE VEROSSIMILHANÇAS SIMPLES

Assuma que se tenha uma amostra aleatória X_1, \dots, X_n originada de uma entre duas distribuições completamente especificadas $f_0(x)$ e $f_1(x)$.

Um teste γ de $H_0: X \sim f_0(\cdot; \theta_0)$ contra $H_1: X \sim f_1(\cdot; \theta_1)$ é definido sendo um teste de razão de verossimilhanças simples se γ é definido por:

- Rejeita-se H_0 se $B \leq b_c$,
- Aceita-se H_0 se $B > b_c$,

em que B é a razão de verossimilhanças, ou seja,

$$B = B(x_1, \dots, x_n) = \frac{\prod_{i=1}^n f_0(x_i)}{\prod_{i=1}^n f_1(x_i)} = \frac{L(\theta_0; x_1, \dots, x_n)}{L(\theta_1; x_1, \dots, x_n)} = \frac{L_0}{L_1}$$

e b_c é uma constante não-negativa e L_j é denominada função de verossimilhança para a amostra de densidade $f_j(\cdot)$, para o caso de amostras independentes. Para cada valor de b_c tem-se um teste diferente. A importância dos testes de razão de verossimilhanças se dá em consequência do resultado fundamental:

LEMA DE NEYMAN-PEARSON

Seja X_1, \dots, X_n uma amostra aleatória de uma função densidade de probabilidade $f(x; \theta)$, em que θ representa um dos dois valores conhecidos θ_0 e θ_1 , e seja $0 < \alpha < 1$ fixado.

Seja b_c uma constante positiva e C^* um subconjunto de \mathcal{X}^n tal que satisfaça

$$(i) P_{\theta_0} [(X_1, \dots, X_n) \in C^*] = \alpha.$$

$$(ii) B(x_1, \dots, x_n) = \frac{L(\theta_0; x_1, \dots, x_n)}{L(\theta_1; x_1, \dots, x_n)} = \frac{L_0}{L_1} \leq b_c \text{ se } (x_1, \dots, x_n) \in C^* \text{ e}$$

$$B(x_1, \dots, x_n) > b_c \text{ se } (x_1, \dots, x_n) \in \bar{C}^*.$$

Então, o teste γ^* correspondente à região crítica C^* é o teste mais poderoso de tamanho α relativo ao teste de hipóteses simples dado por $H_0: \theta = \theta_0$ versus $H_1: \theta = \theta_1$, como citado em Mood, Graybill e Boes (1974).

■

Para se calcular o tamanho de um teste não-aleatorizado é necessária uma integração múltipla sobre a região crítica

$$P[(X_1, \dots, X_n) \in C_r] = \int \dots \int_{C_r} \prod f(x_i; \theta) dx_i,$$

o que pode ser muito difícil.

Nos testes de razão de verossimilhanças a região crítica do teste é geralmente dada em termos de uma desigualdade que depende da razão de verossimilhanças, a estatística $B(X_1, \dots, X_n)$. Uma alternativa a este procedimento é obter a função densidade de probabilidade da estatística

$B(X_1, \dots, X_n)$ sob H_0 , $f_B^0(b)$, e fazer a integração simples $\int_0^{b_c} f_B^0(b) db$. Tal

procedimento é, em geral, também bastante complexo. O usual é obter uma desigualdade equivalente em termos de uma outra estatística suficiente S , para a qual seja mais fácil obter a distribuição de probabilidade.

$$B(X) \leq b_c \Leftrightarrow S \leq k.$$

A seguir será desenvolvido um exemplo para demonstrar como encontrar a função densidade de probabilidade da estatística $B(X)$ razão de verossimilhanças e sua distribuição acumulada sob uma hipótese nula simples e sob uma hipótese alternativa simples.

Deve-se ressaltar que é de suma importância determinar a distribuição acumulada da estatística $B(X)$ nos testes trabalhados nos capítulos 2 e 3; por isso esta ênfase é dada no capítulo que os antecede.

EXEMPLO 1(a)

Seja X_1, \dots, X_n uma amostra aleatória de uma distribuição exponencial $f(x; \theta) = \theta e^{-\theta x}$. Um teste mais poderoso de tamanho α para testar as hipóteses

simples $H_0: \theta = \theta_0$ versus $H_1: \theta = \theta_1$, sendo $\theta_0 < \theta_1$, é o de razão de verossimilhanças; portanto:

$$L_0 = \theta_0^n \exp\left\{-\theta_0 \sum_{i=1}^n x_i\right\} \text{ e } L_1 = \theta_1^n \exp\left\{-\theta_1 \sum_{i=1}^n x_i\right\}$$

e a razão de verossimilhanças:

$$B(x_1, \dots, x_n) = \frac{L_0}{L_1} = \frac{\theta_0^n \exp\left\{-\theta_0 \sum_{i=1}^n x_i\right\}}{\theta_1^n \exp\left\{-\theta_1 \sum_{i=1}^n x_i\right\}} = \left(\frac{\theta_0}{\theta_1}\right)^n \exp\left\{-(\theta_0 - \theta_1) \sum_{i=1}^n x_i\right\}.$$

A família de testes dada pela razão de verossimilhanças consiste em

$$\text{rejeitar } H_0 \text{ se } B(x_1, \dots, x_n) \leq b_c \Rightarrow \boxed{\left(\frac{\theta_0}{\theta_1}\right)^n \exp\left\{-(\theta_0 - \theta_1) \sum_{i=1}^n x_i\right\} \leq b_c}.$$

Para obter, entre esses testes, aquele de tamanho pré-determinado, por exemplo, $\alpha = 0,05$, o procedimento usual é obter a desigualdade equivalente

$$\boxed{\sum_{i=1}^n x_i \leq \frac{1}{(\theta_1 - \theta_0)} \ln \left[b_c \left(\frac{\theta_1}{\theta_0}\right)^n \right]}.$$
 Como $Y = \sum_{i=1}^n X_i$ tem uma distribuição gama

com parâmetros n e θ (Mood, Graybill & Boes, 1974); isto é,

$$f_{\sum_{i=1}^n X_i}(y) = \frac{1}{\Gamma(n)} y^{n-1} \theta^n e^{-\theta y}, \text{ basta resolver a equação integral}$$

$$P_{\theta_0} \left[\sum_{i=1}^n X_i \leq k \right] = \int_0^k \frac{1}{\Gamma(n)} y^{n-1} \theta_0^n e^{-\theta_0 y} dy = 0,05, \text{ obtendo-se o valor de } k.$$

Para os objetivos deste trabalho, será conveniente calcular diretamente a

$$\text{distribuição da variável } B(X_1, \dots, X_n) = \left(\frac{\theta_0}{\theta_1}\right)^n \exp\left\{-(\theta_0 - \theta_1) \sum_{i=1}^n X_i\right\}.$$
 Então,

para determinar a distribuição de razão de verossimilhanças $B(X)$, ou seja, a densidade de B , utiliza-se o teorema de transformação de variáveis

$f_B(b) = \left| \frac{d}{db} g^{-1}(b) \right| f_Y(g^{-1}(b))$ (Mood, Graybill & Boes, 1974). Note que

$B = g(Y) = \left(\frac{\theta_0}{\theta_1} \right)^n \exp\{-(\theta_0 - \theta_1)Y\}$ é a exponencial de uma variável com

distribuição Gama de parâmetros (n, θ) . Expressando Y como função de B , tem-se que:

$$(1) \quad g^{-1}(b) = \frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \Rightarrow \frac{d}{db} g^{-1}(b) = \frac{1}{(\theta_1 - \theta_0)b}$$

$$(2) \quad f_Y(g^{-1}(b)) = \frac{\theta^n}{\Gamma(n)} \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right\}^{n-1} e^{-\theta \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right\}}$$

Utilizando (1) e (2), tem-se:

$$f_B(b) = \left| \frac{d}{db} g^{-1}(b) \right| f_Y(g^{-1}(b)) = \frac{1}{(\theta_1 - \theta_0)b} \frac{\theta^n}{\Gamma(n)} \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right\}^{n-1} e^{-\theta \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right\}}$$

$$f_B(b) = \frac{1}{(\theta_1 - \theta_0)b} \frac{\theta^n}{\Gamma(n)} \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right\}^{n-1} e^{-\theta \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right\}} \quad (1.1)$$

Para um determinado valor α , é necessário resolver a equação integral

$$\int_0^{b_c} \frac{1}{(\theta_1 - \theta_0)b} \frac{\theta^n}{\Gamma(n)} \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right\}^{n-1} e^{-\theta \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right\}} db = \alpha$$

Usando substituição de variáveis:

$$\begin{aligned}
u &= \frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] & b = 0 &\Rightarrow u = -\infty \\
du &= \frac{1}{(\theta_1 - \theta_0)b} db & b = b_c &\Rightarrow u = \frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b_c \right] \\
& & & \int_0^{\frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right]} \frac{\theta_0^n}{\Gamma(n)} u^{n-1} e^{-\theta_0 u} du & (1.2)
\end{aligned}$$

Para $\theta = \theta_0$, tem-se:

$$F_B^0(b) = \int_0^{\frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right]} \frac{\theta_0^n}{\Gamma(n)} u^{n-1} e^{-\theta_0 u} du$$

Fazendo novamente uma substituição $w = \theta_0 u$:

$$\begin{cases} w = \theta_0 u & u = 0 \rightarrow w = 0 \\ dw = \theta_0 du & u = \frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \rightarrow w = \frac{\theta_0}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \end{cases}$$

$$F_B^0(b) = \int_0^{\frac{\theta_0}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right]} \frac{1}{\Gamma(n)} w^{n-1} e^{-w} dw = \Gamma \left(n, \frac{\theta_0}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right),$$

em que $\Gamma(n, t) = \left(\frac{1}{\Gamma(n)} \right) \int_0^t y^{n-1} \exp\{-y\} dy$ é a função gama incompleta;

$$\text{Então, } \boxed{F_B^0(b) = P_{\theta=\theta_0}(B(X) \leq b) = \Gamma \left(n, \frac{\theta_0}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right)} \quad (1.3)$$

De maneira análoga, obtém-se para $\theta = \theta_1$, fazendo uma substituição $w = \theta_1 u$ em (1.2),

$$\begin{cases} w = \theta_1 u & u \rightarrow 0 & w \rightarrow 0 \\ dw = \theta_1 du & u \rightarrow \frac{1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] & w \rightarrow \frac{\theta_1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \end{cases}$$

$$F_B^1(b) = \int_0^{\frac{\theta_1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right]} \frac{1}{\Gamma(n)} w^{n-1} e^{-w} dw = \Gamma \left(n, \frac{\theta_1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right)$$

$$\boxed{F_B^1(b) = P_{\theta=\theta_1}(B(X) \leq b) = \Gamma \left(n, \frac{\theta_1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right)} \quad (1.4)$$

Para $\alpha = 0,05$, tem-se a equação integral:

$$F_B^0(b) = P_{\theta=\theta_0}(B(X) \leq b_c) = \Gamma \left(n, \frac{\theta_0}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b_c \right] \right) = 0,05$$

- Particularizando para $n=1$, com $\theta_0 = 1$ e $\theta_1 = 2$.

$$F_B^0(b) = \Gamma \left(1, \frac{1}{(2-1)} \ln \left[\left(\frac{2}{1} \right)^1 b_c \right] \right) = 0,05 \Rightarrow \int_0^{\ln 2b_c} e^{-y} dy = \left[-e^{-y} \Big|_0^{\ln 2b_c} \right] =$$

$$\left[-e^{-\ln 2b_c} + 1 \right] = -\frac{1}{2b_c} + 1 = 0,05 \Rightarrow \frac{1}{2b_c} = 0,95 \Rightarrow \boxed{b_c = 0,526}$$

Utilizando a desigualdade equivalente, tem-se:

$$P \left[\sum_{i=1}^n X_i \leq k \right] = \int_0^k \frac{1}{\Gamma(n)} y^{n-1} \theta^n e^{-\theta y} dy = P_{\theta=1} [X \leq k] = \int_0^k e^{-y} dy = 0,05$$

$$\int_0^k e^{-y} dy = -e^{-y} \Big|_0^k = -e^{-k} + 1 = 0,05 \Rightarrow e^{-k} = 0,95 \Rightarrow k = -\ln(0,95) \Rightarrow \boxed{k = 0,051}$$

Evidentemente o valor de $b_c = 0,526$ pode ser obtido do valor de k pela equivalência das desigualdades.

$$\sum_{i=1}^n x_i \leq \frac{1}{(\theta_1 - \theta_0)} \ln \left[b_c \left(\frac{\theta_1}{\theta_0} \right)^n \right] = k \Rightarrow \boxed{\sum x_i \leq 0,051}.$$

$$b_c \left(\frac{\theta_1}{\theta_0} \right)^n = \exp\{(\theta_1 - \theta_0)0,051\} \Rightarrow b_c = \left(\frac{\theta_1}{\theta_0} \right)^{-n} \exp\{(\theta_1 - \theta_0)0,051\}$$

$$b_c = 2^{-1} \exp\{0,051\} \Rightarrow \boxed{b_c = 0,526} \Rightarrow \boxed{B(X_1, \dots, X_n) \leq 0,526}$$

Ou ainda, substituindo diretamente os valores considerados de $n = 1$, $\theta_0 = 1$ e $\theta_1 = 2$, tem-se:

Para a distribuição da razão de verossimilhanças, $B(x) = \left(\frac{1}{2} \right) e^x$

$$g^{-1}(b) = \ln(2b) \quad \Rightarrow \quad \frac{d}{db} g^{-1}(b) = \frac{1}{b}$$

$$f_Y(g^{-1}(b)) = \theta e^{-\theta \ln(2b)} = \theta e^{\ln(2b)^{-\theta}} = \theta (2b)^{-\theta}$$

$$f_B(b) = \left| \frac{d}{db} g^{-1}(b) \right| f_Y(g^{-1}(b)) = \frac{1}{b} \times \theta (2b)^{-\theta} \Rightarrow \boxed{f_B(b) = \frac{\theta}{b^{\theta+1}} \left(\frac{1}{2} \right)^\theta}$$

que corresponde à distribuição Pareto $f_B(b; 1/2, \theta)$; portanto, a probabilidade de cometer o erro tipo I será:

$$\alpha = P_{\theta_0} [\text{rejeitar } H_0] = P_{\theta_0} [B \leq b_c] = 0,05$$

$$P_{\theta_0} [B \leq b_c] = \lim_{a \rightarrow 0} \int_a^{b_c} \frac{\theta_0}{b^{\theta_0+1}} \left(\frac{1}{2} \right)^{\theta_0} db = \lim_{a \rightarrow 0} \int_a^{b_c} \frac{1}{b^2} \left(\frac{1}{2} \right) db =$$

$$\lim_{a \rightarrow 0} \int_a^{b_c} \frac{1}{2b^2} db = -\frac{1}{2b} \Big|_a^{b_c} = -\frac{1}{2b_c} + 1 = 0,05 \Rightarrow -\frac{1}{2b_c} = -0,95 \Rightarrow b_c = 0,526$$

No caso da desigualdade equivalente, tem-se que $x \leq \ln(2b_c)$; então,

$$\alpha = P_{\theta_0} [\text{rejeitar } H_0] = P_{\theta_0} [x \leq \ln(2b_c)] = 0,05$$

$$P_{\theta_0} [B \leq b_c] = \int_0^{\ln 2k} \theta_0 e^{-\theta_0 x} dx = \int_0^{\ln 2k} e^{-x} dx = -e^{-x} \Big|_0^{\ln 2k} = -e^{-\ln 2k} + 1 = 0,05$$

$$e^{-\ln 2k} = 0,95 \Rightarrow k = -\ln(0,95) = -(-0,051) \Rightarrow k = 0,051$$

Portanto, para $\alpha=0,05$ com $n=1$, o teste é dado por:

➤ Rejeite H_0 se $b \leq 0,526$ ou $\sum x_i \leq 0,051$

A taxa de erro tipo II é:

$$\beta = 1 - F_B^1(b_c) = 1 - P_{\theta=2}(B(X) \leq b_c) = 1 - \Gamma\left(1, 2 \ln[2b_c]\right)$$

$$\beta = 1 - F_B^1(0,526) = 1 - P_{\theta=2}(B(X) \leq 0,526) = 1 - \Gamma\left(1, \frac{1}{10}\right)$$

$$\beta = 1 - \int_0^{0,1} e^{-x} dx = 1 - \left(-e^{-x}\right)_0^{0,1} = e^{-0,1} \Rightarrow \boxed{\beta = 0,9048}$$

Ou seja, para garantir α pequeno, o valor do erro tipo II foi muito elevado (Figura 1.1).

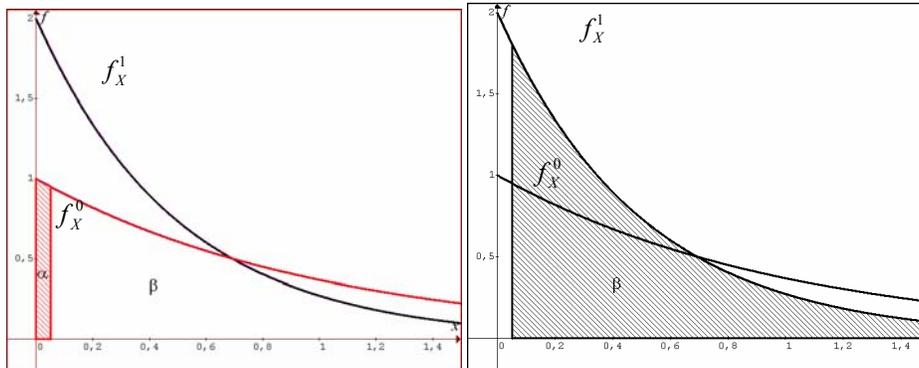


FIGURA 1.1 – Gráficos exibindo as áreas correspondentes às probabilidades de erros tipos I e II iguais a 5% e 90,4%, respectivamente para o caso de $n = 1$.

■

Neste trabalho, o principal enfoque será na situação em que é igualmente importante avaliar as taxas de erro tipo I e tipo II. Uma forma de garantir a importância de ambos os erros é obter testes em que $\alpha = \beta$, isto é:

$$\boxed{F_B^0(b_c) = 1 - F_B^1(b_c)}$$

Neste caso, garante-se que se um deles for pequeno, o outro também será e não há um valor pré-fixado para α , e sim uma condição $\alpha = \beta$ que fornece um determinado valor crítico b_c .

Uma primeira observação é que, considerando a variável aleatória B , o valor mais razoável com relação às hipóteses testadas seria $b_c = 1$, ou seja, aquele que fornece verossimilhanças iguais.

EXEMPLO 2

Considere o teste de hipóteses $H_0: X \sim \text{Beta}(2,3)$ versus $H_1: X \sim \text{Beta}(3,2)$ com $n = 1$. A razão de verossimilhanças de H_0 para H_1 é:

$$B(x) = \frac{\frac{(x)^{2-1}(1-x)^{3-1}}{B(2,3)}}{\frac{(x)^{3-1}(1-x)^{2-1}}{B(3,2)}} = \frac{x^1(1-x)^2}{x^2(1-x)} = \frac{1-x}{x}.$$

O valor crítico de $b_c = 1$ corresponde na desigualdade equivalente a $x = 0,5$. Então, o erro tipo I é dado por:

$$\begin{aligned} \alpha = 1 - F_X^0(x) &= P\left[X \geq \frac{1}{2}\right] = \int_{\frac{1}{2}}^1 \frac{(x)^{2-1}(1-x)^{3-1}}{B(2,3)} dx = \frac{1}{B(2,3)} \int_{\frac{1}{2}}^1 (x - 2x^2 + x^3) dx = \\ &= \frac{1}{B(2,3)} \left[\frac{x^2}{2} - \frac{2x^3}{3} + \frac{x^4}{4} \right]_{\frac{1}{2}}^1 = \frac{4!}{1!2!} \left[\frac{1}{2} - \frac{2}{3} + \frac{1}{4} - \frac{1}{8} + \frac{1}{12} - \frac{1}{64} \right] = 0,3125 \end{aligned}$$

e o erro tipo II é:

$$\beta = F_X^1(x) = P\left[X < \frac{1}{2}\right] = \int_0^{\frac{1}{2}} \frac{(x)^{3-1}(1-x)^{2-1}}{B(3,2)} dx = \frac{1}{B(3,2)} \left[\frac{x^3}{3} - \frac{x^4}{4} \right]_0^{\frac{1}{2}} \Rightarrow$$

$$\beta = \frac{4!}{1!2!} \left[\frac{1}{24} - \frac{1}{64} \right] = 0,3125 .$$

Assim, $F_B^1(1) = 1 - F_B^0(1) = 31,25\%$ (Figura 1.2).

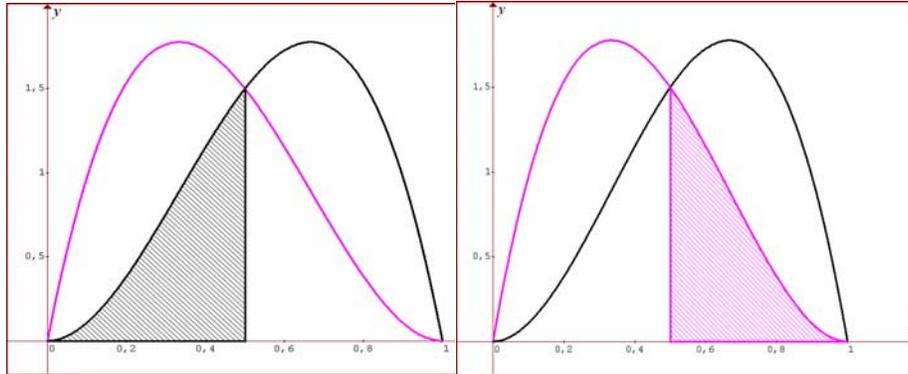


FIGURA 1.2 – Gráficos exibindo as áreas correspondentes às probabilidades de erros tipos I e II iguais a 31,25% para o caso de $n = 1$.

Mas o valor crítico $b_c = 1$ nem sempre garante taxa de erro tipo I igual à taxa de erro tipo II (exceto para o caso de razão de verossimilhanças simétrica, que será detalhado na página 54). Por exemplo,

Para $n=1$

No exemplo 1(a) tem-se que, quando $b_c = 1$, baseado nas expressões dadas por (1.3) e (1.4), o erro tipo I é dado por $\alpha = 0,5$ e o erro tipo II, por $\beta = 0,25$.

$$\alpha = \int_0^{\ln 2} e^{-x} dx = -e^{-x} \Big|_0^{\ln 2} = -e^{-\ln 2} + 1 = 0,5 \quad e$$

$$\beta = 1 - \int_0^{2\ln 2} e^{-x} dx = 1 - \left[-e^{-x} \Big|_0^{2\ln 2} \right] = 1 - [-0,25 + 1] = 0,25 .$$

Para n=4

A probabilidade de erro tipo I é dada por $\alpha = 0,197$, e a do erro tipo II, por $\beta = 0,302$.

$$\alpha = \int_0^{\ln 2^4} \frac{1}{\Gamma(4)} x^3 e^{-x} dx = -\frac{2}{3} \ln(2)^3 - \frac{1}{2} \ln(2)^2 - \frac{1}{4} \ln(2) + \frac{15}{16} \cong 0,19663$$

$$\beta = 1 - \int_0^{2 \ln 2^4} \frac{1}{\Gamma(4)} x^3 e^{-x} dx = \frac{1}{3} \ln(2)^3 + \frac{1}{8} \ln(2)^2 + \frac{1}{32} \ln(2) + \frac{1}{256} \cong 0,30197.$$

Para n=10

As probabilidades de erro tipo I e tipo II são, respectivamente, $\alpha = 0,163$ e $\beta = 0,116$.

$$\alpha = \int_0^{\ln 2^{10}} \frac{1}{\Gamma(10)} x^9 e^{-x} dx \cong 0,163 \quad e$$

$$\beta = 1 - \int_0^{2 \ln 2^{10}} \frac{1}{\Gamma(10)} x^9 e^{-x} dx \cong 0,116.$$

■

Para o exemplo 1(a), garantir $\alpha = \beta$ equivale a encontrar o valor de b_c que satisfaça a seguinte equação:

$$\Gamma\left(n, \frac{\theta_0}{(\theta_1 - \theta_0)} \ln\left[\left(\frac{\theta_1}{\theta_0}\right)^n b_c\right]\right) = 1 - \Gamma\left(n, \frac{\theta_1}{(\theta_0 - \theta_1)} \ln\left[\left(\frac{\theta_1}{\theta_0}\right)^n b_c\right]\right).$$

Utilizando $\theta_0 = 1$ e $\theta_1 = 2$, obtém-se:

$$\Gamma\left(n, \ln\left[2^n b_c\right]\right) = 1 - \Gamma\left(n, 2 \ln\left[2^n b_c\right]\right).$$

Particularizando para o valor de $n = 1$, tem-se:

$$\Gamma(1, \ln[2b_c]) = 1 - \Gamma(1, 2 \ln[2b_c]) \Rightarrow \frac{1}{\Gamma(1)} \int_0^{\ln 2b_c} e^{-y} dy = 1 - \frac{1}{\Gamma(1)} \int_0^{2 \ln 2b_c} e^{-y} dy$$

$$\left[-e^{-\ln 2b_c} + 1\right] = 1 - \left[-e^{-2 \ln 2b_c} + 1\right] \Rightarrow -\frac{1}{2b_c} + 1 = \frac{1}{4b_c^2} \Rightarrow \boxed{b_c = 0,809}$$

Conseqüentemente, as probabilidades de erro serão dadas por:

$$\alpha = \Gamma(1, \ln[2b_c]) \Rightarrow \frac{1}{\Gamma(1)} \int_0^{\ln 2(0,809)} e^{-y} dy = [-e^{-\ln 1,618} + 1] \Rightarrow \boxed{\alpha = 0,382}$$

$$\beta = 1 - \Gamma(1, 2 \ln[2b_c]) \Rightarrow 1 - \int_0^{2 \ln 2(0,809)} e^{-y} dy = e^{-2 \ln 1,618} \Rightarrow \boxed{\beta = 0,382}$$

Para obter erros iguais no caso de amostra de tamanho 4, o valor de b seria 0,797, o que forneceria uma probabilidade de erro de 25,2% (Figura 1.3).

$$\alpha = \beta \Rightarrow \int_0^{\ln 2^4(0,797)} \frac{1}{\Gamma(4)} x^3 e^{-x} dx = 1 - \int_0^{2 \ln 2^4(0,797)} \frac{1}{\Gamma(4)} x^3 e^{-x} dx = 0,252$$

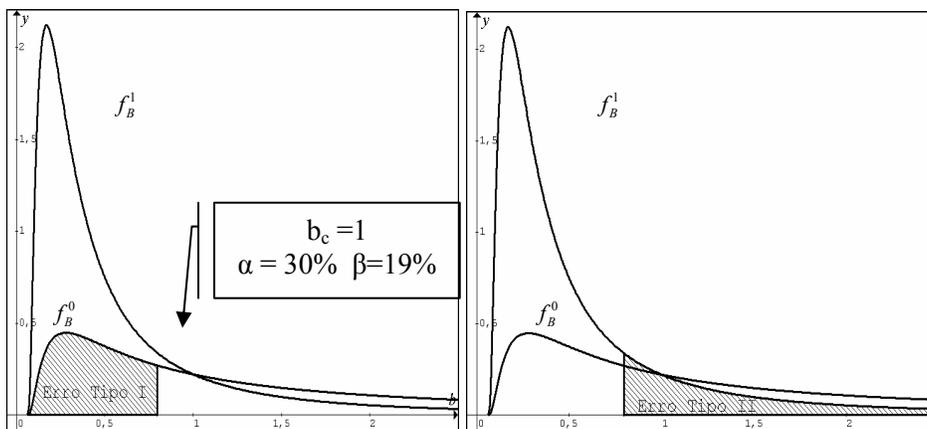


FIGURA 1.3 - Gráficos exibindo as áreas correspondentes às probabilidades de erros tipos I e II iguais a 25,2% para o caso de $n = 4$.

Utilizando o software *Maple 10*, determinou-se o valor de b_c para vários tamanhos de amostras (Tabela 1.2).

TABELA 1.2 - Valores de b_c calculados para diferentes tamanhos de amostra com probabilidades de erros tipo I e tipo II iguais.

n	1	4	10	20	30	50	100
b_c	0,8090	0,7973	0,7955	0,7948	0,7947	0,7946	0,7945
$\alpha = \beta$	38,20%	25,23%	14,07%	6,25%	2,98%	0,74%	0,028%

Como se observa nos testes de hipóteses tradicionais, a única forma de diminuir α e β é aumentando o tamanho da amostra. Ou seja, não há, em testes de hipóteses com tamanho de amostra reduzido, uma medida que possa considerar probabilidades de erros menores, mesmo quando observados dados extremos; por isso há várias críticas a este teste.

2.1.2 CRÍTICAS À TEORIA DOS TESTES DE HIPÓTESES DE NEYMAN-PEARSON

A teoria de testes de hipóteses sistematizada por Neyman e Pearson (1933¹, citado por Lehmann, 1993) consiste em definir uma região crítica tal que se a amostra obtida estiver contida nesta região, a hipótese nula é rejeitada. De forma equivalente, utilizando uma estatística teste S para um valor crítico k , a hipótese nula é recusada se $S \leq k$. Os tamanhos dos erros Tipo I e Tipo II são pré-determinados, isto é, são pré-experimentais. A justificativa para esta teoria é o princípio freqüentista, com todas as suas implicações práticas e filosóficas.

Princípio Freqüentista. “Quando um experimento é repetido indefinidamente nas mesmas condições, o erro médio real não deve ser maior

¹ Neyman, J.; Pearson, E. S. On the problem of the most efficient tests of statistical hypotheses. Philosophical Transactions of the Royal Society of London, Ser. A, 231, 289-337. 1933.

que (e seria ideal que fosse igual) a média dos erros obtidos nos experimentos (Berger, 2003)''.

Geralmente, em problemas reais uma das hipóteses é mais importante que a outra. Nesse caso, esta hipótese é tomada como a hipótese nula e a taxa de erro Tipo I tem que ser controlada. Assim, é natural considerar testes tais que a taxa de erro Tipo I não seja maior do que um tamanho especificado, digamos α , e tentar minimizar a taxa de erro Tipo II.

A maior crítica à teoria dos testes de hipóteses de Neyman-Pearson é devida ao fato de que os erros Tipo I e Tipo II não refletirem a evidência dos dados observados. Apresentando uma probabilidade de erro pré-especificado α , não se leva em conta se os dados observados estão próximos ou afastados da região crítica. Tal fato pareceu a Ronald Fisher altamente não científico. Fisher também criticou a necessidade de levar em conta uma hipótese alternativa e a dificuldade associada de se ter que lidar com uma função poder dependente de parâmetros (tipicamente desconhecidos) (Berger, 2003).

Outro problema discutido é que, para conduzir um experimento, o pesquisador propõe uma hipótese nula H_0 , uma hipótese alternativa H_1 e seleciona a taxa de erro Tipo I, α . Devido a uma idéia comum, aceita pela maioria dos periódicos, uma hipótese nula é rejeitada quando a estatística do teste fornece uma área menor que 0,05. O valor de 5% tem o objetivo de padronizar a literatura científica com relação à análise estatística; entretanto, muita importância é dada a este limite, o que presume um papel de um “limite mágico”. Segundo Tonhasca Júnior (1991), você publica se encontrá-lo; caso contrário, esquece ou começa tudo de novo. Este posicionamento leva os pesquisadores a procurar, entre muitas técnicas estatísticas, a mais conveniente (isto é, a que resulte em probabilidade menor que 5%). Por outro lado, dados interessantes podem ser ignorados porque $\alpha = 0,06$ torna os resultados “não significantes”. Uma política rigorosa sob valores α não leva em consideração

diferenças entre a significância estatística e a significância prática de outras áreas. Em muitos casos, deve-se analisar se os resultados têm algum valor prático porque “se a diferença não é diferente o bastante para fazer a diferença, qual é a diferença?”. Um experimento não é conduzido para determinar se duas variedades de um produto, ou duas substâncias, são as mesmas; muitas vezes, sabe-se isso de antemão. O objetivo de um experimento é encontrar a magnitude dessas diferenças (Tonhasca Júnior, 1991).

Quando as hipóteses H_0 e H_1 são igualmente importantes para o pesquisador, isto é, quando cometer um erro tipo II for tão grave quanto cometer um erro tipo I, pode-se questionar, então, se o teste de Neyman-Pearson seria o mais adequado. Um exemplo disso seria no caso de um teste clínico em que se pode supor um exame com resultados H_0 : $r =$ positivo versus H_1 : $r =$ negativo. Deixar de tratar um paciente que estava doente pode ser tão grave quanto tratar de um paciente sadio. Nesse caso, seria adequado pensar em um teste no qual $\alpha = \beta$. Ou seja, buscar-se-ia um equilíbrio de forma que, dependendo da estatística utilizada e do tamanho da amostra, as probabilidades de erros fossem iguais. Os erros tipo I e II têm muito em comum com o resultado falso positivo e o resultado falso negativo que ocorrem nos testes de diagnóstico. O problema nesse caso é que há situações em que os valores de α e β poderiam ser muito grandes, e para diminuir o tamanho dos erros, o tamanho da amostra deve ser aumentado excessivamente.

2.2 TESTES DE SIGNIFICÂNCIA

Um outro procedimento alternativo ao teste de hipóteses de Neyman-Pearson é o teste de significância, proposto por Ronald Fisher para testar $H_0: \theta = \theta_0$, em que é usual apresentar, como resultado inferencial final do teste, uma medida de evidência em favor da hipótese nula H_0 , denominada nível descritivo do teste (ou *valor-p*) e, com base no valor desta medida, tomar uma

decisão sobre a rejeição ou não desta hipótese de acordo com o critério do pesquisador.

A principal diferença entre o teste de hipóteses e o teste de significância está em não construir a região crítica. No teste de hipóteses especifica-se um tamanho do teste α pré-experimentalmente, enquanto, no teste de significância, o cálculo é apenas pós-experimental, sendo uma medida de evidência contra a hipótese nula (Figura 1.4). O procedimento implica em olhar para a probabilidade total do conjunto de todas as possíveis observações similares, dado que a hipótese nula é verdadeira, as quais são pelo menos tão raras quanto aquelas que foram observadas. Esta probabilidade acumulada é o *valor-p* e é conhecida também como probabilidade da cauda porque é usualmente obtida nas extremidades das distribuições. Esse procedimento resulta em um *teste de significância puro* ou *teste Fisheriano*. A particularidade desse método é que ele depende da probabilidade dos dados não-observados na cauda da distribuição, sob os quais a informação é difícil de ser obtida porque eles são raros; conseqüentemente, esse método é sensível ao modelo escolhido (Lindsey, 1996).

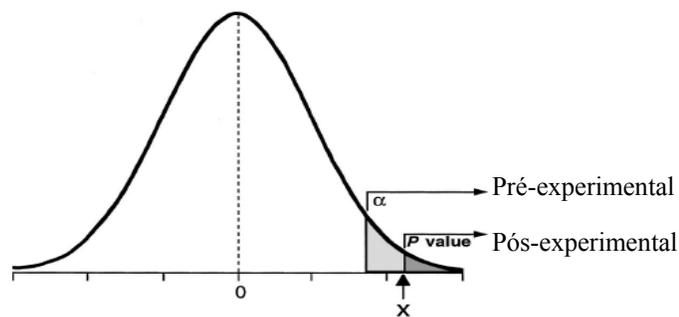


FIGURA 1.4 - Função densidade de probabilidade sob a hipótese nula. Ambos, α (a probabilidade de erro tipo I) e o *valor-p*, são “áreas da cauda” sob esta curva. A área da cauda para α é fixada antes do experimento, enquanto a área da cauda do *valor-p* só é conhecida depois que um resultado é observado.

A idéia é que, após calcular o *valor-p*, o pesquisador pode escolher o seu próprio nível de significância. Em geral, se T é uma estatística do teste e H_0 é rejeitada, então o *valor-p* é a probabilidade $P(T \geq t \mid H_0)$ em que t é o valor observado de T .

$\text{valor-p} = \text{Prob} [T \text{ pelo menos tão extremo quanto o valor observado} \mid H_0]$.

Portanto, o *valor-p* é a probabilidade de se observarem resultados mais extremos quanto aos obtidos se a hipótese nula for verdadeira. A idéia é que, se o *valor-p* for grande, ele fornece evidência de que H_0 é verdadeira, enquanto um *valor-p* pequeno indica que existe evidência nos dados contra H_0 .

Mais tecnicamente, o *valor-p*, de um valor observado t_{obs} de alguma variável aleatória T usada como uma estatística do teste, é a probabilidade de que, dado que a hipótese nula é verdadeira, T assumirá um valor tão ou mais desfavorável para a hipótese nula quanto o do valor observado t_{obs} . “Mais desfavorável para a hipótese nula” pode, em alguns casos, significar maior que; em outros casos, menor que; e em alguns casos, mais afastado de um centro especificado (teste bilateral).

As seguintes interpretações de *valores-p*, utilizadas para decidir sobre a "significância estatística", normalmente são (P-value, 2007):

$\text{valor-p} \geq 0,10$	Não existe evidência contra H_0
$\text{valor-p} < 0,10$	Fraca evidência contra H_0
$\text{valor-p} < 0,05$	Evidência significativa
$\text{valor-p} < 0,01$	Evidência altamente significativa
$\text{valor-p} < 0,001$	Evidência extremamente significativa

Quando o *valor-p* crítico é utilizado como critério para decidir se a hipótese nula deve ser rejeitada, ele normalmente é comparado ao tamanho do teste α predeterminado, escolhido de maneira um pouco arbitrária (frequentemente fixado em 0,05), sendo que:

- Se $\text{valor-}p < \alpha$, rejeita-se H_0 .
- Se $\text{valor-}p > \alpha$, não se rejeita H_0 .

EXEMPLO 3:

É realizado um experimento para determinar se um lançamento de moeda é justo (50% de chance de cair cara ou coroa) ou viesado, para cara (> 50% chance de cair cara) ou para coroa (> 50% chance de cair coroa). Suponha que, de 20 lançamentos, o resultado experimental tenha fornecido cara 14 vezes. O *valor-p* deste resultado seria a chance de em uma moeda honesta cair pelo menos 14 vezes cara entre os 20 lançamentos (pois valores maiores neste caso também são menos favoráveis à hipótese nula de uma moeda justa) ou no máximo 6 vezes cara entre os 20 lançamentos. Nesse caso, a variável aleatória T tem uma distribuição binomial. A probabilidade de que em 20 lançamentos de uma moeda honesta ocorram 14 ou mais caras é :

$$\sum_{y=14}^{20} \binom{20}{y} p^y (1-p)^{20-y} = (0,5)^{20} \sum_{y=14}^{20} \binom{20}{y} =$$

$$0,000000954 \left[\binom{20}{14} + \binom{20}{15} + \binom{20}{16} + \binom{20}{17} + \binom{20}{18} + \binom{20}{19} + \binom{20}{20} \right] =$$

$$0,000000954 [38760 + 15540 + 4845 + 1140 + 190 + 20 + 1] = 0,0577 .$$

Já que este é um teste bilateral, a probabilidade de que 20 lançamentos da moeda resultem em 14 ou mais caras ou 6 ou menos caras é 0,115.

Se for considerado o nível de significância 0,05, o *valor-p* calculado o excede e, assim, a hipótese nula, de que o resultado observado de 14 caras entre 20 lançamentos vem de uma moeda honesta, não é rejeitada. Tal fato é declarado freqüentemente como sendo “não significativo estatisticamente ao nível de significância de 5%”.

Porém, obtendo-se uma única cara a mais, o *valor-p* resultante seria

$$\sum_{y=15}^{20} \binom{20}{y} 0,5^y (1-0,5)^{20-y} =$$

$$0,000000954 \left[\binom{20}{15} + \binom{20}{16} + \binom{20}{17} + \binom{20}{18} + \binom{20}{19} + \binom{20}{20} \right] = 0,0207,$$

que no teste bilateral corresponde à probabilidade de 0,0414.

Desta vez a hipótese nula - que o resultado observado de 15 caras entre 20 provém de uma moeda honesta - é rejeitada. Tal fato é declarado freqüentemente como sendo “significativo estatisticamente ao nível de 5%”.

Note que a hipótese alternativa composta (precisa), $H_1 : p \neq \frac{1}{2}$, não foi levada em consideração. Assim como, se o problema envolvesse uma hipótese alternativa simples, esta também não seria levada em consideração no teste de significância.

2.2.1 CRÍTICAS AO TESTE DE SIGNIFICÂNCIA DE FISHER

O atrativo do *valor-p* é que ele permite fazer uma afirmação absoluta sobre um modelo à luz dos dados, sem compará-lo com qualquer outro. Tal afirmação não é possível pela função de verossimilhança, utilizada nos testes de hipóteses, que compara modelos diferentes. Há, freqüentemente, uma hipótese alternativa H_1 , mas a construção do teste de significância não permite incorporar uma hipótese alternativa específica. Obviamente, se o modelo da hipótese nula é rejeitado, outro terá que ser encontrado para substituí-lo, de forma que a suposição de ter somente um modelo disponível pode ser melhor apenas momentaneamente. Um perigo é que, se o modelo a ser testado é especificado depois de se inspecionarem os dados, o pesquisador pode escolher, até mesmo

inconscientemente, aquele que produz as observações menos raras (Lindsey, 1996).

Neyman criticou os *valores-p* por violarem o princípio frequentista (página 22), enquanto Jeffreys sentia que a lógica de basear *valores-p* em uma área de extremidade (como oposta aos dados reais) era tola [... uma hipótese possivelmente verdadeira pode ser rejeitada por não se preverem resultados observáveis que não tenham ocorrido” (Jeffreys, 1961 citado por Berger, 2003)]. Mais recentemente, com relação a ambas as críticas, tem havido grande discussão de que a má interpretação, tão comum de *valores-p* como probabilidades de erro, muito frequentemente resulta em considerável exagero da evidência contra H_0 .

Alguns dos vários equívocos mais frequentes sobre *valores-p* são dados a seguir:

1. o *valor-p* não é a probabilidade de que a hipótese nula seja verdadeira.

Comparação de abordagens Bayesianas e Fisherianas mostra que um *valor-p* pode ser muito próximo de zero, enquanto a probabilidade a posteriori da hipótese nula é muito perto de um. Este é o paradoxo de Jeffreys-Lindley.

2. o *valor-p* não é a probabilidade que decide; ele é “somente uma oposição” (justificando a “regra” de considerar *valores-p* pequenos como “significativos”).

Como o cálculo de um *valor-p* está baseado na suposição que a decisão é apenas sobre uma hipótese, ele evidentemente não pode simultaneamente ser usado para medir a probabilidade de que a suposição inicial seja verdadeira.

3. o *valor-p* não é a probabilidade de errar ao rejeitar a hipótese nula. Este equívoco é uma versão da denominada “falácia do promotor²”.

Fisher (1959, citado por Killen, 2005), que introduziu o teste de significância da hipótese nula, sabia que “tal teste de significância não permite que seja feita qualquer afirmação sobre a hipótese em questão em termos de probabilidade matemática”. Isso ocorre porque tais afirmações de interesse $p(H_0|T \geq t)$ geralmente não são iguais a $p(T \geq t | H_0)$, que é o *valor-p*, sendo t o valor observado da estatística do teste. A confusão de uma probabilidade condicional para a outra é análoga à falácia do promotor. Bayes mostrou que
$$P(H_0 | T \geq t) = \frac{P(T \geq t | H_0)P(H_0)}{P(T \geq t)}$$
. As probabilidades não-condicionais são as densidades *a priori*, e são em grande parte desconhecidas. Fisher (1959, citado por Killen, 2005) deixou claro que $p(T \geq t | H_0)$ pode “influenciar a aceitação da hipótese nula”.

4. $[1 - (\textit{valor-p})]$ não é a probabilidade de a hipótese alternativa ser verdadeira (Ver item 2).

5. o nível de significância do teste não é determinado pelo *valor-p*.

² A FALÁCIA DO PROMOTOR: você ganha o grande prêmio de loteria, mas é então acusado de ter trapaceado. No julgamento, o promotor afirma que ganhar na loteria sem trapacear é extremamente improvável, e que, então, você ser inocente deve ser comparavelmente improvável. Este raciocínio está intuitivamente falho. A falha é que o promotor não levou em conta a baixa probabilidade a priori de que você, e não algum outro, ganharia a loteria em primeiro lugar. A falácia que usa probabilidade condicional pode ser analisada: Suponha que E é a evidência observada, e I denota que o acusado é inocente. Sabe-se que $P(E|I)$ (a probabilidade que a evidência seja observada se o acusado for inocente) é mínima. O promotor conclui injustamente que $P(I|E)$ (a probabilidade que o acusado é inocente, dada a evidência E) é comparativamente mínima. Porém, $P(E|I)$ e $P(I|E)$ são bastante diferentes; usando o teorema de Bayes, tem-se que: $P(I|E) = [P(E|I).P(I)]/P(E)$. Assim, a probabilidade a priori de inocência $P(I)$ e a probabilidade total da evidência observada $P(E)$ precisam ser consideradas. Se $P(I)$ é muito maior que $P(E)$, então $P(I|E)$ pode ser grande também.

O nível de significância de um teste é um valor que deveria ser decidido pelo pesquisador que interpreta os dados antes de serem vistos, e é comparada contra o *valor-p* ou qualquer outra estatística calculada depois que o teste foi executado (P-value, 2007).

Interpretação do valor-p como uma medida de evidência não-freqüentista

A ilustração da natureza não-freqüentista dos *valores-p* pode ser vista a partir do *applet* disponível em www.stat.duke.edu/~berger. O *applet* computa a proporção de vezes que a hipótese nula é verdadeira quando o *valor-p* está em um pequeno intervalo determinado $[a,b]$. Para realizar os cálculos, é necessário estabelecer um intervalo bem próximo do *valor-p* desejado, pois a probabilidade de um ponto específico é zero. Então, quando se deseja encontrar a proporção de vezes que a hipótese nula é verdadeira para um determinado *valor-p* (por exemplo, *valor-p* = 0,05), deve-se escolher um intervalo pequeno (por exemplo, de 0,049 a 0,05).

O *applet* considera o teste de $H_0: \theta = 0$ contra $H_1: \theta \neq 0$, em que θ é a média de uma distribuição normal com desvio padrão 1. São tomadas n observações de uma população normal e o teste é baseado na estatística usual $\sqrt{n}|\bar{X}|$. O aplicativo simula uma longa série de tais testes, e simplesmente conta com que frequência H_0 é verdadeira e falsa sempre que o *valor-p* está no intervalo especificado.

Para rodar o *applet* deve-se escolher, além do intervalo para o cálculo do *valor-p*:

1º) De que forma são geradas as observações? Deve-se definir a proporção de observações geradas a partir da hipótese nula, isto é, a proporção de hipóteses nulas, H_0 , que serão escolhidas inicialmente sendo verdadeiras na sucessão de testes simulados;

2º) Como será definida a hipótese alternativa? Deve-se definir o valor da média da normal que aparece sob a hipótese alternativa, H_1 , na sucessão de testes simulados. Pode-se escolher a média da hipótese alternativa como um ponto fixo, ou gerá-la aleatoriamente de uma das distribuições disponíveis no *applet* (Normal ou Uniforme).

3º) Estabelecem-se também o tamanho da amostra e o número de simulações a serem realizadas.

Definidos os critérios, são geradas as n observações de acordo com a hipótese nula ou alternativa na proporção estabelecida. Para cada simulação é calculado o *valor-p* baseado, obviamente, na hipótese nula. Se o valor calculado pertence ao intervalo determinado, a simulação é registrada e, então, verificam-se os dados que forneceram este *valor-p* foram realmente originados da hipótese nula ou não.

A tabela 1.3 seguinte mostra os resultados de algumas simulações utilizando este *applet*.

TABELA 1.3 - Frequência de H_0 verdadeira no teste com $H_0: X \sim N(0,1)$ versus $H_1 \sim N(\theta,1)$ para *valor-p* no intervalo 0,049 a 0,05, mostrando a natureza não-freqüentista do *valor-p*, utilizando 10000 simulações com proporção fixada de 50 % de H_0 verdadeira.

Distribuição da qual θ da H_1 foi gerada	Tamanho da Amostra	% H_0 Verdadeira de acordo com o <i>valor-p</i>
N (-1,1)	10	44,32
N (3,1)	10	92,15
U (0,1)	10	29,18
U (0,3)	10	53,25
N (-1,1)	100	70,30
N (3,1)	100	99,11
U (0,1)	100	53,25
U (0,3)	100	77,92

A principal discussão é que alguns pesquisadores – nas aplicações – potencialmente utilizam *valores-p*, mas depois caem no impulso de interpretá-los como probabilidades de erro. Segundo Berger (2003), Fisher e Neyman não podem ser culpados por esta situação: Neyman foi extremamente claro ao expor que dever-se-ia usar probabilidade de erro escolhidas pré-experimentalmente se fosse desejada a validade freqüentista do teste, enquanto Fisher foi muito cuidadoso em distinguir *valores-p* de probabilidades de erro. Discussões sobre esse fato (e outros aspectos do uso inapropriado de *valores-p*) têm sido repetidamente levantadas em muitos artigos científicos (Hubbard e Armstrong, 2005; Killen, 2005; Berger, 2003; Sterne e Smith, 2001; Hubbard e Bayarri, 2003).

Tomar uma decisão é diferente de testar a significância. A diferença entre testes de significância e o teste como regra para decidir entre as hipóteses não está nos cálculos, mas no raciocínio que os motiva. Em um teste de significância focaliza-se uma única hipótese (H_0) e uma única probabilidade (o *valor-p*). O objetivo é avaliar a força de evidência amostral contra H_0 . Faz-se o cálculo do poder para verificar a sensibilidade do teste. Se não se pode rejeitar H_0 , conclui-se apenas que não há evidência suficiente contra H_0 , e não que H_0 seja efetivamente verdadeira. Se o mesmo problema de inferência é encarado como um problema de decisão, focalizam-se as duas hipóteses e propõe-se uma regra para decidir entre elas com base na evidência amostral. Neste caso, deve-se focalizar igualmente duas probabilidades, as probabilidades dos dois tipos de erro, fazendo opção por uma das duas hipóteses.

Há distinções claras entre as duas maneiras de encarar os testes estatísticos. Às vezes, entretanto, as duas abordagens se confundem. A escolha antecipada de um tamanho de teste α tem sentido quando se quer, além de descrever a força de evidência sobre os dados, tomar uma decisão.

Diferentes pesquisadores podem desejar utilizar diferentes níveis de significância. Segundo Sterne e Smith (2001), é preferível utilizar o *valor-p*, que permite que cada um decida individualmente se a evidência é suficientemente forte. Os usuários da estatística têm enfatizado, com frequência, certos níveis padrões de significância, tais como 10%, 5% e 1%. Essa ênfase reflete o tempo em que as tabelas de valores críticos, e não os programas de computador, dominavam a prática estatística. Não há linha divisória rígida entre “significante” e “não-significante”; há apenas uma evidência crescente na medida em que o *valor-p* diminui. Não há distinção prática entre os *valores-p* 0,049 e 0,051.

Há distinções claras entre as duas maneiras de encarar os testes estatísticos. Às vezes, entretanto, as duas abordagens se confundem. A teoria de Neyman-Pearson de teste de hipóteses, com taxa α de erro Tipo I, como tamanho do teste, é considerada amplamente como um teste estatístico rigoroso. O modelo de Fisher de teste de significância, em que a evidência *valor-p* denota o “nível de significância”, contudo, domina a prática de testes estatísticos. Este paradoxo aconteceu porque estas duas teorias incompatíveis de testes estatísticos clássicos estiveram juntas, anonimamente misturadas, criando a falsa impressão de um único modelo coerente de decisão estatística. Mostra-se que esta abordagem híbrida, ao testar com $\text{valor-p} < \alpha$ como critério de decisão, é utilizada em alguns livros de ensino e de pesquisa. Quer dizer, os pesquisadores que assim agem tentam o impossível interpretando, simultaneamente, o *valor-p* como uma taxa de erro Tipo I e como uma medida de evidência contra a hipótese nula. O resultado disso é que alguns ainda não sabem o que a mais apreciada, e onipresente e desejada indagação, “significância estatística”, realmente significa (Hubbard e Armstrong, 2005).

Segundo Berger (2003), é natural (e comum), em outras ciências, culpar os profissionais da estatística pela situação, apontando que livros didáticos

comuns ensinam o teste freqüentista de Neyman-Pearson e, depois, *valores-p* sem aviso suficiente de que estas são metodologias com princípios completamente diferentes. Mais ainda, segundo Berger (2003), os estatísticos não estão isentos de culpa em um sentido: não se tem feito um esforço profissional conjunto para fornecer ao mundo científico uma metodologia de teste unificada e, assim, todos se tornam cúmplices silenciosos dessa situação infeliz. Sem uma metodologia de teste unificada disponível, é tempo de se empenhar para construir e fornecer aos não-estatísticos ferramentas de teste para que eles possam efetivamente usar e entender.

2.3 TESTE DE HIPÓTESES NO CONTEXTO BAYESIANO

2.3.1 DEFINIÇÕES GERAIS

Na teoria de teste de hipóteses, no caso de hipótese simples versus simples, é usual ponderar os erros tipo I e tipo II através de uma função denominada função perda, $l(\delta_i; \theta_j)$, em que δ_i é a decisão de aceitar H_i e θ_j é o valor do parâmetro. Geralmente $l(\delta_i; \theta_i) = 0$, enquanto $l(\delta_1; \theta_0) = l_0$ e $l(\delta_0; \theta_1) = l_1$, em que l_0 e l_1 são constantes positivas (Tabela 1.4).

TABELA 1.4 – Função perda para teste de hipóteses simples x simples

	Ocorre H_0	Ocorre H_1
δ_0	0	l_1
δ_1	l_0	0

Um critério de comparação entre dois testes é naturalmente o critério de se optar pelo teste que apresenta menores perdas. Como se têm dois erros possíveis, toma-se a perda média como critério.

$$l_0 \cdot [\text{probabilidade de cometer erro Tipo I}] +$$

l_1 . [probabilidade de cometer erro Tipo II]

A função de decisão de um teste Υ é dada em termos da amostra da forma

$$\delta(X_1, \dots, X_n) = \begin{cases} \delta_1 & \text{se } (X_1, \dots, X_n) \in C_\Upsilon \\ \delta_0 & \text{se } (X_1, \dots, X_n) \notin C_\Upsilon \end{cases}.$$

Pode-se, então, definir a função risco $R_\Upsilon(\theta)$ como sendo a perda esperada.

Definição 1:

A *Função Risco* $R_\Upsilon(\theta)$ de um teste Υ é definida como:

$$R_\Upsilon(\theta) = E_\theta[l(\delta, \theta)] = \int_{C_\Upsilon} \dots \int l(\delta_1, \theta) L(x|\theta) dx + \int_{\bar{C}_\Upsilon} \dots \int l(\delta_0, \theta) L(x|\theta) dx,$$

em que $L(x|\theta) = \prod_{i=1}^n f(x_i; \theta)$.

Portanto,

$$R_\Upsilon(\theta_0) = l_0\alpha + 0 \Rightarrow R_\Upsilon(\theta_0) = l_0\alpha \quad \text{e} \quad R_\Upsilon(\theta_1) = l_1\beta.$$

É desejável escolher um teste Υ que possua a menor função risco, mas, infelizmente, tal teste raramente existe (Mood, Graybill & Boes, 1974). A dificuldade é que a função risco assume dois valores, e um teste que minimize ambos os valores geralmente não existe. Um outro critério menos desejável, porém útil, é o de minimizar o maior valor da função risco.

Definição 2:

Um teste Υ_m de $H_0: \theta = \theta_0$ versus $H_1: \theta = \theta_1$ é definido como sendo um *teste minimax* se e somente se:

$$\max[R_{\Upsilon_m}(\theta_0), R_{\Upsilon_m}(\theta_1)] \leq \max[R_\Upsilon(\theta_0), R_\Upsilon(\theta_1)]$$

para qualquer outro teste Υ .

Teorema:

Para uma amostra aleatória X_1, \dots, X_n de f_0 ou f_1 , i.i.d., considere o teste Y_m dado pela região crítica $C_{Y_m} = \{(x_1, \dots, x_n) : B \leq k_m\}$, em que B é a razão de verossimilhanças e k_m , uma constante positiva, e função perda $l = (l_0, l_1)$, tal que $R_{Y_m}(\theta_0) = R_{Y_m}(\theta_1)$, então Y_m é minimax (Mood, Graybill & Boes, 1974).

No contexto Bayesiano tem-se um outro critério para decidir sobre a optimalidade de um teste. Considerando uma distribuição a priori π , $\pi_0 = P[\theta = \theta_0] = 1 - P[\theta = \theta_1]$, deve-se utilizar a média da função risco em relação à priori π .

Definição 3:

Risco de Bayes de um teste Y com priori π é definido por:

$$r(\pi) = E^\pi [R_Y(\theta)] = (1 - \pi_0) R_Y(\theta_0) + (\pi_0) R_Y(\theta_1)$$

Definição 4:

Um teste Y_π é dito Bayesiano com respeito a uma priori π se

$$(1 - \pi_0) R_{Y_\pi}(\theta_0) + (\pi_0) R_{Y_\pi}(\theta_1) \leq (1 - \pi_0) R_Y(\theta_0) + (\pi_0) R_Y(\theta_1)$$

para qualquer outro teste Y , isto é, se seu risco de Bayes é o menor possível.

O Lema de Neyman-Pearson estabelece que um teste mais poderoso de tamanho α de H_0 versus H_1 é o teste de razão de verossimilhanças, isto é, aquele que tem a região crítica da forma

$$C = \left\{ (x_1, \dots, x_n) : B(x) = \frac{L(\theta_0; x)}{L(\theta_1; x)} \leq b_c \right\}$$

para alguma constante não-negativa b_c .

Os testes Bayesianos são também testes de razão de verossimilhanças:

LEMA 1.

Testes de Bayes e Testes de Razão de Verossimilhanças são equivalentes, isto é, todo Teste de Razão de Verossimilhanças é um Teste de

Bayes para alguma priori $\pi = (\pi_0, \pi_1)$ e todo Teste de Bayes é um Teste de Razão de Verossimilhanças com $b_c = \frac{\pi_1 l_1}{\pi_0 l_0}$.

PROVA:

O teste de Bayes define uma região crítica C_π que minimiza o risco esperado $r(\pi) = r(\pi, \delta^\pi) = \pi_0 l_0 \alpha + \pi_1 l_1 \beta$

$$\begin{aligned} &= \pi_0 l_0 \Pr[X \in C_\pi | \theta_0] + \pi_1 l_1 \Pr[X \in \bar{C}_\pi | \theta_1] = \\ &= \pi_0 l_0 \int_{C_\pi} L(\theta_0; x) dx + \pi_1 l_1 \int_{\bar{C}_\pi} L(\theta_1; x) dx \end{aligned}$$

Assumindo que X é contínuo, e que a função de verossimilhança satisfaz condições de regularidade apropriadas, tem-se:

$$\pi_0 l_0 \int_{C_\pi} L(\theta_0; x) dx + \pi_1 l_1 \left[1 - \int_{C_\pi} L(\theta_1; x) dx \right] = \pi_1 l_1 + \int_{C_\pi} [\pi_0 l_0 L(\theta_0; x) - \pi_1 l_1 L(\theta_1; x)] dx$$

Para minimizar esta expressão deve-se escolher C_π que inclua todo x para o qual o integrando é negativo e exclua todo x para o qual o integrando é positivo; conseqüentemente, $x \in C_\pi$ se e somente se

$$[\pi_0 l_0 L(\theta_0; x) - \pi_1 l_1 L(\theta_1; x)] \leq 0, \text{ isto é, } x \in C_\pi \text{ se e somente se } \frac{L(\theta_0; x)}{L(\theta_1; x)} \leq \frac{l_1 \pi_1}{l_0 \pi_0}.$$

Isso demonstra o lema, já que para qualquer b e qualquer $l_0, l_1 \geq 0$ pode-se escolher priori $\pi = (\pi_0, \pi_1)$ que satisfaça $\frac{\pi_1}{\pi_0} = b_c \frac{l_0}{l_1}$. ■

Considere o seguinte teste Bayesiano Y_π com região crítica dada por:

$$C_\pi = \left\{ (x_1, \dots, x_n) : B(x) = \frac{f(x | \theta_0)}{f(x | \theta_1)} \leq \frac{l_1 \pi_1}{l_0 \pi_0} \right\}$$

Se a razão entre as densidades a priori for $\rho = \frac{\pi_1}{\pi_0}$ e a razão entre as

perdas for $l = \frac{l_1}{l_0}$, tem-se:

$$C_\pi = \left\{ (x_1, \dots, x_n) : B(x) = \frac{f(x|\theta_0)}{f(x|\theta_1)} \leq \frac{l_1\pi_1}{l_0\pi_0} \right\} \Rightarrow C_\pi = \{ (x_1, \dots, x_n) : B(x) \leq l\rho \}$$

$$C_\pi = \left\{ (x_1, \dots, x_n) : B(x) = \frac{f(x|\theta_0)}{f(x|\theta_1)} \leq \frac{l_1\pi_1}{l_0\pi_0} \right\} \Rightarrow$$

$$C_\pi = \left\{ (x_1, \dots, x_n) : \frac{f(x|\theta_0)\pi_0}{f(x|\theta_1)\pi_1} \leq \frac{l_1}{l_0} \right\}$$

Dividindo numerador e denominador por $[f(x|\theta_0)\pi_0 + f(x|\theta_1)\pi_1]$, tem-se:

$$C_\pi = \left\{ (x_1, \dots, x_n) : \frac{\frac{f(x|\theta_0)\pi_0}{f(x|\theta_0)\pi_0 + f(x|\theta_1)\pi_1}}{\frac{f(x|\theta_1)\pi_1}{f(x|\theta_0)\pi_0 + f(x|\theta_1)\pi_1}} \leq \frac{l_1}{l_0} \right\}.$$

Mas, pelo Teorema de Bayes. $f(\theta_0|x) = \frac{f(x|\theta_0)\pi_0}{f(x|\theta_0)\pi_0 + f(x|\theta_1)\pi_1}$.

Então:

$$C_\pi = \left\{ (x_1, \dots, x_n) : \frac{f(\theta_0|x)}{f(\theta_1|x)} \leq \frac{l_1}{l_0} \right\}.$$

Ou seja, a região crítica pode ser expressa em função de razão das densidades a posteriori e de razão entre as perdas. O teste Bayesiano baseado nesta região crítica é dado da seguinte maneira:

$$Y_\pi = \begin{cases} \text{se } l_0 f(\theta_0|x) \leq l_1 f(\theta_1|x), & \text{rejeita-se } H_0 \\ \text{se } l_0 f(\theta_0|x) > l_1 f(\theta_1|x), & \text{aceita-se } H_0 \end{cases}$$

2.3.2. MEDIDAS DE EVIDÊNCIA BAYESIANA USUAIS

No problema estatístico de *Testes de Hipóteses* sob a abordagem Bayesiana após a realização do teste, que é simplesmente a regra de decisão, ou seja, a região crítica, é necessário apresentar uma medida de evidência. Na abordagem freqüentista, como comentado no item 2, é muito usado, às vezes inconvenientemente, o *valor-p* como medida de evidência, que se baseia na distribuição da estatística do teste sob H_0 . Na abordagem Bayesiana, as medidas de evidência usuais para teste de hipótese são o Fator de Bayes e a Probabilidade *a posteriori* de H_0 , que no caso de hipóteses simples se reduzem à mesma medida. Pode também ser utilizada como medida de evidência a perda esperada *a posteriori* (Berger, 1985).

Probabilidade *a posteriori* de H_0

A probabilidade *a posteriori* com relação à hipótese nula é dada por

$$f(\theta_0 | x) = \frac{f(x | \theta_0) \pi_0}{f(x | \theta_0) \pi_0 + f(x | \theta_1) \pi_1}.$$

A hipótese nula é aceita sempre que $f(\theta_0 | x)$ é grande, significando que os dados favorecem a hipótese nula, e será rejeitada caso contrário.

Perda esperada *a posteriori*

É simplesmente a ponderação da probabilidade *a posteriori* pelas perdas l_i associadas a cada decisão δ_i , de forma que:

Em caso de rejeição, apresenta-se como medida de evidência:

$$\rho(f(\theta | x), a_1) = l_0 f(\theta_0 | x) = l_0 \left[\frac{f(x | \theta_0) \pi_0}{f(x | \theta_0) \pi_0 + f(x | \theta_1) \pi_1} \right].$$

Em caso de aceitação, apresenta-se como medida de evidência:

$$\rho(f(\theta|x), a_0) = I_1 f(\theta_1|x) = I_1 \left[\frac{f(x|\theta_1)\pi_0}{f(x|\theta_0)\pi_0 + f(x|\theta_1)\pi_1} \right],$$

sendo a_i a ação de Bayes tomada.

Obs.: A função Risco é pré-experimental, pois tem-se uma integração para os valores possíveis de x , enquanto a perda esperada *a posteriori* é considerada uma medida de evidência pós-experimental.

Fator de Bayes

Sejam $H_0 : \theta \in \Theta_0$ e $H_1 : \theta \in \Theta_1$ as hipóteses a serem testadas com $\Theta_0 \cup \Theta_1 = \Theta$, $\Theta_0 \cap \Theta_1 = \emptyset$, $\Theta_0 \neq \emptyset$ e $\Theta_1 \neq \emptyset$.

O Fator de Bayes em favor da hipótese nula é dada por $FB_0(x) = \frac{m(x|H_0)}{m(x|H_1)}$, em que $m(x|H_i) = \int_{\Theta_i} L(x|\theta)\pi_i(\theta)d\theta$, ou seja, a razão

entre as preditivas. Quando o espaço paramétrico se reduz a $\Theta = \{\theta_0, \theta_1\}$, ou seja, as hipóteses nula e alternativa são simples, o Fator de Bayes se reduz a

$$FB_0(x) = \frac{L(x|\theta_0)}{L(x|\theta_1)},$$

coincidindo com a razão de verossimilhanças. Aceita-se a hipótese nula sempre que $FB_0(x)$ é grande, significando que os dados observados favorecem a hipótese nula, geralmente quando o Fator de Bayes é maior que 1, ou seja, $m(x|H_0) > m(x|H_1)$.

Recentemente, uma outra medida de evidência foi proposta por Pereira e Stern (1999). É a metodologia denominada FBST (Full Bayesian Significance Test), que é adequada para hipóteses precisas ($H_0 : \theta = \theta_0$ x $H_1 : \theta \neq \theta_0$) e, portanto, não será abordada neste trabalho.

2.3.3. TESTES BAYESIANOS APRESENTANDO MEDIDAS DE EVIDÊNCIA

No teste Υ_π , sendo incluída uma medida de evidência como a perda esperada *a posteriori*, este passa a ser apresentado da seguinte maneira:

$$\Upsilon_\pi = \begin{cases} \text{se } l_0 f(\theta_0 | x) \leq l_1 f(\theta_1 | x), \text{ rejeita-se } H_0 \text{ com perda esperada a posteriori} \\ l_0 f(\theta_0 | x) = l_0 \left[\frac{f(x|\theta_0)\pi_0}{f(x|\theta_0)\pi_0 + f(x|\theta_1)\pi_1} \right] = l_0 \left[\frac{B(x)\pi_0}{B(x)\pi_0 + \pi_1} \right] \\ \text{se } l_0 f(\theta_0 | x) > l_1 f(\theta_1 | x), \text{ aceita-se } H_0 \text{ com perda esperada a posteriori} \\ l_1 f(\theta_1 | x) = l_1 \left[\frac{f(x|\theta_1)\pi_1}{f(x|\theta_0)\pi_0 + f(x|\theta_1)\pi_1} \right] = l_1 \left[\frac{\pi_1}{B(x)\pi_0 + \pi_1} \right] \end{cases}$$

Alguns casos particulares importantes do teste Υ_π são:

- Se for utilizada uma *priori* objetiva $\pi_0 = 1/2$, lembrando que $B(x) = f(x|\theta_0)/f(x|\theta_1)$, tem-se que o teste fica da forma:

$$\Upsilon_\pi = \begin{cases} \text{se } l_0 f(\theta_0 | x) \leq l_1 f(\theta_1 | x), \\ \text{rejeita-se } H_0 \text{ com perda esperada a posteriori } l_0 f(\theta_0 | x) = l_0 \left(\frac{B(x)}{B(x) + 1} \right) \\ \text{se } l_0 f(\theta_0 | x) > l_1 f(\theta_1 | x), \\ \text{aceita-se } H_0 \text{ com perda esperada a posteriori } l_1 f(\theta_1 | x) = l_1 \left(\frac{1}{B(x) + 1} \right) \end{cases}$$

- Se forem assumidas perdas iguais, $l_0 = l_1$, para as possíveis decisões incorretas e 0 para uma decisão correta, a região crítica fica da forma:

$$C_\pi = \left\{ (x_1, \dots, x_n) : B(x) < \frac{\pi_1}{\pi_0} \text{ ou } \frac{f(\theta_0 | x)}{f(\theta_1 | x)} < 1 \right\}$$

ou, simplesmente,

$$C_\pi = \left\{ (x_1, \dots, x_n) : f(\theta_0 | x) < f(\theta_1 | x) \right\}.$$

Ou seja, a região crítica é definida em termos apenas da densidade *a posteriori*. Portanto, os testes consistem em simplesmente aceitar a hipótese que possui a maior probabilidade *a posteriori*.

No caso em que $\pi_0 \neq \pi_1$, definindo $\rho = \frac{\pi_1}{\pi_0}$, tem-se

$$\begin{aligned} f(\theta_0 | x) &= \frac{f(x | \theta_0) \pi_0}{f(x | \theta_0) \pi_0 + f(x | \theta_1) \pi_1} = \frac{\frac{f(x | \theta_0)}{f(x | \theta_1)} \pi_0}{\frac{f(x | \theta_0)}{f(x | \theta_1)} \pi_0 + \frac{f(x | \theta_1)}{f(x | \theta_1)} \pi_1} = \\ &= \frac{B(x) \pi_0}{B(x) \pi_0 + \pi_1} = \frac{B(x) \frac{\pi_0}{\pi_0}}{B(x) \frac{\pi_0}{\pi_0} + \frac{\pi_1}{\pi_0}} \Rightarrow \boxed{f(\theta_0 | x) = \frac{B(x)}{B(x) + \rho}} \end{aligned}$$

e também

$$\begin{aligned} f(\theta_1 | x) &= \frac{f(x | \theta_1) \pi_1}{f(x | \theta_1) \pi_1 + f(x | \theta_0) \pi_0} = \frac{\frac{f(x | \theta_1)}{f(x | \theta_1)} \pi_1}{\frac{f(x | \theta_1)}{f(x | \theta_1)} \pi_1 + \frac{f(x | \theta_0)}{f(x | \theta_1)} \pi_0} = \\ &= \frac{\pi_1}{\pi_1 + B(x) \pi_0} = \frac{\frac{\pi_1}{\pi_0}}{\frac{\pi_1}{\pi_0} + B(x) \frac{\pi_0}{\pi_0}} = \boxed{f(\theta_1 | x) = \frac{\rho}{\rho + B(x)}}. \end{aligned}$$

E o teste Bayesiano, Y_π , para perdas iguais e razão entre as densidades *a priori* dada por ρ , fica da forma:

$$Y_\pi = \begin{cases} \text{se } f(\theta_0 | x) < f(\theta_1 | x), \text{ rejeita-se } H_0 \\ \text{com perda esperada a posteriori } f(\theta_0 | x) = \frac{B(x)}{B(x) + \rho} \\ \text{se } f(\theta_0 | x) \geq f(\theta_1 | x), \text{ aceita-se } H_0 \\ \text{com perda esperada a posteriori } f(\theta_1 | x) = \frac{\rho}{B(x) + \rho} \end{cases}$$

■

Além de perdas iguais, se a densidade *a priori* for particularizada para $\pi_0 = 1/2$, tem-se:

$$Y_\pi = \begin{cases} \text{se } B(x) < 1 \text{ ou } f(\theta_0 | x) < f(\theta_1 | x), \text{ rejeita-se } H_0 \\ \text{com perda esperada a posteriori } f(\theta_0 | x) = \frac{B(x)}{1 + B(x)} \\ \text{se } B(x) \geq 1 \text{ ou } f(\theta_0 | x) \geq f(\theta_1 | x), \text{ aceita-se } H_0 \\ \text{com perda esperada a posteriori } f(\theta_1 | x) = \frac{1}{1 + B(x)} \end{cases}$$

Essas perdas esperadas *a posteriori* são simplesmente as probabilidades *a posteriori* (Figura 1.5).

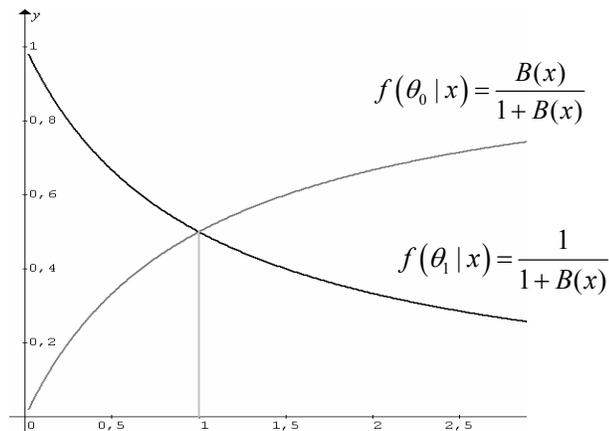


FIGURA 1.5 – Representação das probabilidades *a posteriori* em função de razão de verossimilhanças.

Retornando aos testes minimax, deve-se observar que o teste de Bayes escolhe C_π que minimiza o risco esperado, $\min \{ \pi_0 l_0 \alpha + \pi_1 l_1 \beta \}$, ou equivalentemente, a perda esperada em que a esperança é determinada com respeito à distribuição *a posteriori* de θ . Por outro lado, o teste *minimax* minimiza o risco máximo, isto é, minimiza $\{ \text{Max} \{ l_0 \alpha, l_1 \beta \} \}$ (Garthwaite, Jolliffe & Jones, 1995).

Há um resultado geral, em teoria de decisão, segundo o qual, quando há um número finito de possíveis decisões e há um único procedimento minimax, então ele é um procedimento de Bayes para a distribuição a priori menos favorável. A razão para este nome é que, entre todas as distribuições a priori possíveis, esta é a priori para o qual o risco esperado mínimo é um máximo (Ferguson, 1967).

O procedimento de Bayes, que é também um procedimento minimax, é aquele que possui risco constante para todo θ , ou seja, $l_0 \alpha = l_1 \beta$. Esse resultado é justificado por observar que α cresce quando β decresce, e vice-versa, de forma que o decrescimento de um deles levaria ao crescimento do outro conseqüentemente, cresceria a perda máxima (Garthwaite, Jolliffe & Jones, 1995).

Tem-se ainda que, como o teste minimax é um teste de Bayes para determinada priori, o minimax deve ser um teste de razão de verossimilhanças e, conseqüentemente, um teste mais poderoso para algum valor de α . Dessa forma, segue que o teste minimax é um teste de razão de verossimilhanças para o qual $l_0 \alpha = l_1 \beta$ (Garthwaite, Jolliffe & Jones, 1995).

EXEMPLO 1(b)

Dando continuidade ao exemplo 1 de se testarem as hipóteses simples $H_0: \theta = \theta_0$ versus $H_1: \theta = \theta_1$, sendo $\theta_0 < \theta_1$, $f(x; \theta) = \theta e^{-\theta x}$, utilizado para ilustrar o lema de Neyman-Pearson, deseja-se agora verificar o que ocorre quando são utilizadas funções perda.

Pelo Lema 1 (página 38), um teste de Bayes é um Teste de razão de Verossimilhanças com $b_c = \frac{\pi_1 l_1}{\pi_0 l_0}$.

Considere $\theta_0 = 1$ e $\theta_1 = 2$ com tamanho de amostra $n = 4$ e perdas $l_0 = 2$ e $l_1 = 1$ e priori $\pi_0 = 1/4$ e $\pi_1 = 3/4$. Então, o teste Bayesiano tem a região crítica, dada por:

$$B(x) = \frac{f(x | \theta_0)}{f(x | \theta_1)} \leq \frac{l_1 \pi_1}{l_0 \pi_0} \Rightarrow \sum_{i=1}^n x_i \leq \frac{1}{(\theta_1 - \theta_0)} \ln \left[\frac{l_1 \pi_1}{l_0 \pi_0} \left(\frac{\theta_1}{\theta_0} \right)^n \right]$$
$$\sum_{i=1}^4 x_i \leq \frac{1}{(2-1)} \ln \left[\frac{(\frac{3}{4})1}{(\frac{1}{4})2} \left(\frac{2}{1} \right)^4 \right] = \ln [3(2)^3] \Rightarrow \sum_{i=1}^4 x_i \leq 3,178$$

Como $\sum_{i=1}^4 X_i \sim Gama(4,1)$ e

$$\alpha = P_{\theta_0} \left[\sum_{i=1}^4 X_i \leq 3,178 \right] = \int_0^{3,178} \frac{1}{\Gamma(4)} y^3 e^{-y} dy$$
$$\alpha = \frac{1}{6} \int_0^{3,178} y^3 e^{-y} dy = 0,3926 \Rightarrow \boxed{\alpha = 39,26\%}$$

e para a hipótese alternativa $\sum_{i=1}^4 X_i \sim Gama(4,2)$

$$\beta = P_{\theta_1} \left[\sum_{i=1}^4 X_i > 3,178 \right] = 1 - \int_0^{3,178} \frac{1}{\Gamma(4)} y^3 2^4 e^{-2y} dy$$

$$\beta = 1 - \left[\frac{8}{3} \int_0^{3,178} y^3 e^{-2y} dy \right] = 1 - 0,8778 \Rightarrow \boxed{\beta = 12,22\%}$$

O risco de Bayes é de

$$r(\pi) = \pi_0 l_0 \alpha + \pi_1 l_1 \beta = \frac{1}{4}(1)0,3926 + \frac{3}{4}(2)0,1222 = 0,09815 + 0,1833 = 0,2815.$$

Usando a abordagem freqüentista, isto é, considerando simplesmente o teste de Razão de Verossimilhanças, fixando $\alpha = 0,05$, o valor crítico para a região de rejeição é de

$$\alpha = P_{\theta_0} \left[\sum_{i=1}^4 X_i \leq k \right] = \int_0^k \frac{1}{\Gamma(4)} y^3 e^{-y} dy = 0,05 \Rightarrow k = 1,367,$$

que corresponde a uma probabilidade de erro tipo II dada por:

$$\beta = P_{\theta_1} \left[\sum_{i=1}^4 X_i > 1,367 \right] = 1 - \int_0^{1,367} \frac{1}{\Gamma(4)} y^3 2^4 e^{-2y} dy = 1 - 0,293 \Rightarrow \boxed{\beta = 0,707}.$$

O Teste de Bayes aumentou, então, consideravelmente α além de quaisquer dos níveis fixos habituais, mas, ao mesmo tempo, β foi muito reduzido. O fato de que $\beta < \alpha$ segue da desigualdade $\frac{\pi_1 l_1}{\pi_0 l_0} = \frac{\frac{3}{4}1}{\frac{1}{4}2} > 1$.

Se as informações a priori forem ignoradas e utilizado o procedimento minimax com risco constante, em que $\alpha l_0 = \beta l_1$, o qual é um teste de Bayes para um priori adequada π , tem-se $2\alpha = \beta$ e região crítica $\sum_{i=1}^4 x_i \leq 2,1975$, sendo $\alpha = 0,18015$ e $\beta = 0,3603$, e o risco de Bayes sobe para $r(\pi) = l_0 \alpha = l_1 \beta = 0,3603$.

■

2 REFERENCIAL TEÓRICO

2.1 TESTES FREQUENTISTAS CONDICIONAIS

Os testes freqüentistas tradicionais ou não-condicionais apresentam probabilidades de erro independentes dos dados; assim, para um teste de tamanho $\alpha = 0,05$, apresenta-se a mesma probabilidade de erro na rejeição se os dados estão apenas no limite da rejeição ou mais internamente na região de rejeição. Após definida a região crítica C_Y , a decisão e os erros tipo I e II são os mesmos para amostras obtidas em diferentes regiões do espaço amostral, respectivamente, dentro e fora da região crítica. Ou seja, o teste de hipóteses tradicional não depende da distância dos valores observados em relação à fronteira da região crítica. Na Figura 2.1 pode-se perceber que apesar de as observações x_1 e x_3 estarem muito próximas em relação ao espaço amostral, corresponderiam a decisões diferentes, enquanto, entre x_1 e x_2 , ambas seriam rejeitadas no teste Y de tamanho α , não apresentando nenhuma medida que leve em consideração que x_2 está mais afastada da fronteira do que x_1 . O mesmo ocorre com x_3 e x_4 , que seriam aceitos com mesmo nível de confiança apesar das distâncias em relação à fronteira da região crítica.

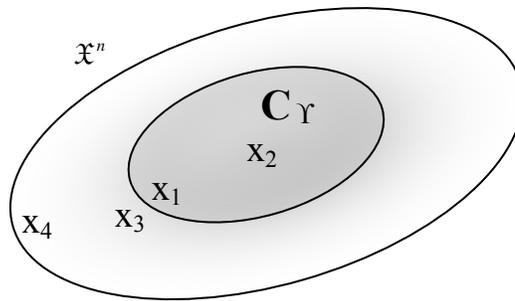


FIGURA 2.1 – Representação, no espaço das realizações, de diferentes distâncias da amostra em relação à região crítica, tanto para rejeição (x_1, x_2) quanto para aceitação (x_3, x_4) da hipótese nula.

Como o objetivo de um experimento é encontrar a magnitude das diferenças entre as hipóteses, a forma tradicional de contornar este problema é através de um teste de significância ou de um valor p . Observe, entretanto, como foi descrito no item 2 do capítulo 1, que *valores-p* não são verdadeiras medidas de erro freqüentistas. Assim, ao se desejar uma solução freqüentista para as críticas feitas aos testes de hipóteses, o *valor-p* não é o recurso adequado. De fato, há uma intensa discussão sobre como os *valores-p* podem ser altamente enganosos quando interpretados erroneamente como razões de erro freqüentista no teste de hipóteses (Berger, 2003).

Kiefer (1977) propôs que os freqüentistas tentassem empregar a metodologia denominada “*teste freqüentista condicional*” para superar estes problemas. O artigo de Kiefer (1977) foi de grande importância e também muito debatido, merecendo comentários de outros estatísticos, como George A. Barnard, Lawrence D. Brown, Robert J. Buehler, Arthur P. Dempster, Oscar Kempthorne, Dennis V. Lindley e J. Wolfowitz.

A idéia básica por trás desta abordagem freqüentista condicional é construir uma medida estatística de “força de evidência” nos dados. A idéia é particionar o espaço amostral, \mathfrak{X}^n , utilizando uma estatística condicionante Z , obtendo-se um subespaço C_z , tal que

$$\mathfrak{X}^n = \bigcup_{Z=z} C_z, \text{ com } C_z \cap C_{z'} = \emptyset,$$

e então desenvolver medidas freqüentistas condicionais em C_z .

É importante ressaltar que as medidas freqüentistas condicionais construídas dessa maneira são totalmente freqüentistas e provavelmente, para a maioria das pessoas, claramente mais fáceis de interpretar do que as medidas freqüentistas não-condicionais. Elas têm a mesma propriedade não-condicional, contudo podem fornecer mais garantia sobre a decisão a ser tomada, pois dependem do conjunto de dados obtido (Berger, 2003).

A teoria desenvolvida pode ser exemplificada em uma situação bastante simples.

EXEMPLO 1 (c)

Dando continuidade ao exemplo 1(a), para $f(x; \theta) = \theta e^{-\theta x}$, se forem testadas as hipóteses simples $H_0: \theta = 1$ versus $H_1: \theta = 2$, considerando o tamanho de amostra $n = 4$, tem-se que o menor valor possível para a variável $B(x)$ é dado por

$$B(x) = \frac{1^4 e^{-\sum_{i=1}^4 x_i}}{2^4 e^{-2\sum_{i=1}^4 x_i}} = \left(\frac{1}{2}\right)^4 e^{\sum_{i=1}^4 x_i} \text{ e, quando } \sum_{i=1}^4 x_i \sim 0, \text{ tem-se}$$

$$B(x) \sim \left(\frac{1}{2}\right)^4 = 0,0625.$$

O teste freqüentista não-condicional (teste de razão de verossimilhanças) com o critério de apresentar uma probabilidade de erro tipo I igual ao erro tipo II é definido pelo valor crítico $B(x) \leq 0,797$ com $\alpha = \beta = 25,2\%$, isto é, consiste em rejeitar a hipótese nula se o valor observado para $B(x)$ for menor ou igual a $0,797$ (Tabela 1.2). É intuitivo que, nesta situação, ao se realizar um experimento, e encontrar um valor de razão de verossimilhanças b igual a $0,69$, tem-se uma menor evidência para recusar a hipótese nula do que no caso em que se obteve uma amostra que forneça b igual a $0,15$, por exemplo. Neste caso, parece razoável subdividir a região crítica em “sub-regiões” de tal modo que seja possível estabelecer probabilidades de erros dependes do fato de que o valor observado pertença ou não a determinada “sub-região”.

2.1.1 Probabilidades de Erros Tipo I e II Condicionais aos Dados

Deseja-se, agora, apresentar uma medida de evidência que será denominada probabilidade de erro condicional. Essa medida de erros tipo I e II será calculada condicionalmente ao fato de os dados estarem numa “subregião”, denominada aqui de região de baixa ou alta evidência, contrária ou a favor de H_0 , determinada pela partição do espaço amostral.

Para explicar tal partição, considere o problema apresentado no exemplo 1 (c). Deseja-se estabelecer uma partição de $[0, \infty)$ de forma que se tenham dois níveis de confiança relativos a rejeitar ou não a hipótese nula. A idéia é, então, obter dois valores c e d para os quais se tem uma alta evidência, tanto para recusar a hipótese nula se $b \leq c$ quanto para aceitá-la se $d \leq b$, e uma evidência menor em rejeitar a hipótese nula para o caso em que $c < b \leq 0,797$ ou aceitá-la, quando $0,797 < b < d$ (Figura 2.2).

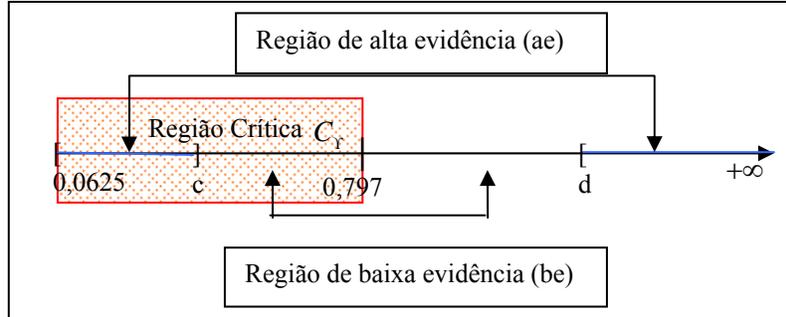


FIGURA 2.2 - Representação das regiões de decisão com diferentes níveis de confiança.

Para o exemplo 1 (c), eles serão definidos da seguinte maneira:

Dado que o valor de b resultante da amostra observada (x_1, \dots, x_4) está na região de alta evidência (ae), ou seja, $b \in \left(\left[(0,5)^4, c \right] \cup [d, \infty) \right)$, a probabilidade de que ele esteja no intervalo $\left[(0,5)^4, c \right]$ define a taxa de erro tipo I condicional α_{ae} , dada por

$$\alpha_{ae} = P_{\theta=1} \left\{ b \in \left[(0,5)^4, c \right] \mid b \in \left(\left[(0,5)^4, c \right] \cup [d, \infty) \right) \right\} = \frac{F_B^0(c)}{F_B^0(c) + (1 - F_B^0(d))}.$$

Utilizando as expressões (1.3) e (1.4), deduzidas na página 13,

$$F_B^0(b) = \Gamma \left(n, \frac{\theta_0}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right) \text{ e } F_B^1(b) = \Gamma \left(n, \frac{\theta_1}{(\theta_1 - \theta_0)} \ln \left[\left(\frac{\theta_1}{\theta_0} \right)^n b \right] \right),$$

tem-se:

$$\alpha_{ae} = \frac{\int_0^{\ln[2^4 c]} \frac{1}{\Gamma(4)} y^3 e^{-y} dy}{\int_0^{\ln[2^4 c]} \frac{1}{\Gamma(4)} y^3 e^{-y} dy + \int_{\ln[2^4 d]}^{\infty} \frac{1}{\Gamma(4)} y^3 e^{-y} dy} \Rightarrow$$

$$\alpha_{ae} = \frac{\int_0^{\ln[16c]} y^3 e^{-y} dy}{\int_0^{\ln[16c]} y^3 e^{-y} dy + \int_{\ln[16d]}^{\infty} y^3 e^{-y} dy} \quad (2.1).$$

E na região de baixa evidência (be), tem-se:

$$\alpha_{be} = P_{\theta=1} \{b \in]c; 0,797] \mid b \in (]c; 0,797] \cup]0,797; d[)\} \Rightarrow$$

$$\alpha_{be} = \frac{\int_{\ln(16 \times 0,797)}^{\ln(16 \times 0,797)} y^3 e^{-y} dy}{\int_{\ln 16c}^{\ln(16 \times 0,797)} y^3 e^{-y} dy + \int_{\ln(16 \times 0,797)}^{\ln 16d} y^3 e^{-y} dy} \quad (2.2).$$

Como o interesse principal, neste trabalho, são os testes de hipóteses em que os erros tipo I e tipo II são igualmente importantes, é necessário avaliar, também, as probabilidades de erros condicionais tipo II resultantes deste procedimento, que são dadas por:

$$\beta_{ae} = P_{\theta=2} \{b \in [d, \infty) \mid b \in ((0,5)^4, c] \cup [d, \infty)\} = \frac{1 - F_B^1(d)}{F_B^1(c) + (1 - F_B^1(d))} \Rightarrow$$

$$\beta_{ae} = \frac{\int_{2\ln[16 \times d]}^{\infty} y^3 e^{-y} dy}{\int_0^{2\ln[16c]} y^3 e^{-y} dy + \int_{2\ln[16 \times d]}^{\infty} y^3 e^{-y} dy} \quad (2.3)$$

e

$$\beta_{be} = P_{\theta=2} \{b \in]0, 797; d[[b \in (]c; 0, 797] \cup]0, 797; d[]\} \Rightarrow$$

$$\beta_{be} = \frac{\int_0^{2\ln(16d)} y^3 e^{-y} dy}{\int_{2\ln(16 \times 0,797)}^{2\ln(16 \times 0,797)} y^3 e^{-y} dy + \int_{2\ln 16c}^{2\ln 16d} y^3 e^{-y} dy} \quad (2.4).$$

A escolha dos valores de c e d é muito subjetiva da maneira como foi definida. Por exemplo, para uma taxa de erro tipo I condicional fixada em $\alpha_{ae} = 0,10$, na região de alta evidência, pode-se considerar, por exemplo, o valor de $c = 0,2$, isto é,

]rejeita-se H_0 se

$$B(x) \leq c \Rightarrow \frac{L_0}{L_1} \leq c \Rightarrow \frac{L_0}{L_1} \leq 0,2 \Rightarrow L_0 \leq 0,2L_1 \Rightarrow 5L_0 \leq L_1.$$

Ou seja, se a hipótese alternativa é, no mínimo, cinco vezes mais verossímil do que a hipótese nula.

Para este valor sugerido de c e com a taxa de erro condicional tipo I definida em 10%, tem-se o valor de d obtido resolvendo-se a equação integral (2.1)

$$\frac{\int_0^{\ln[2^4, 0,2]} \frac{1}{\Gamma(4)} y^3 e^{-y} dy}{\int_0^{\ln[2^4, 0,2]} \frac{1}{\Gamma(4)} y^3 e^{-y} dy + \int_{\ln[2^4, d]}^{\infty} \frac{1}{\Gamma(4)} y^3 e^{-y} dy} = 0,10,$$

que resulta em d igual a 8,592, que, por sua vez, fornece uma probabilidade de erro condicional tipo I na região de baixa evidência $\alpha_{be} = 0,32$, dada pela integral (2.2).

Para a probabilidade de erro condicional tipo II, resolvendo as integrais (2.3) e (2.4), tem-se que $\beta_{ae} = 0,05$ e $\beta_{be} = 0,31$. É claro que a probabilidade de

ocorrência de determinada amostra na região de alta ou baixa evidência é dependente da distribuição de $B(X)$ sob H_0 e H_1 e influencia diretamente no cálculo da esperança da probabilidade de erro condicional tipo I e II (Figura 2.3).

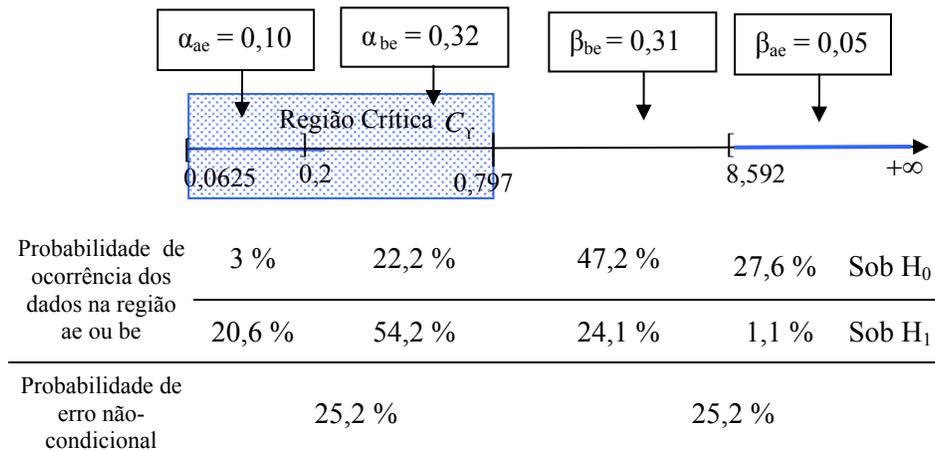


FIGURA 2.3 - Regiões de alta e baixa evidência para rejeição e aceitação de H_0 e respectivos erros.

Conseqüentemente, se for usada esta partição simples no exemplo 1(c) com $c = 0,2$, e for observado o valor de $b = 0,69$, diz-se que “ H_0 é rejeitada com erro condicional de 32%”; porém, se for observado $b = 0,15$, diz-se que “ H_0 é rejeitada com probabilidade de erro condicional de 10 %”.

Então, o pesquisador deve decidir se é desejável “particionar” os 25,2 % de probabilidade de tomar uma decisão errada, obtidos por um teste freqüentista não-condicional, em uma partição na qual se têm duas probabilidades de erro tipo I, sendo 3% a chance de ocorrência de dados na região em que se tem

probabilidade de 10% de erro ao rejeitar (inferior ao $\alpha = 25,2\%$) e 22,2 % a chance de ocorrência de dados na região em que se tem probabilidade de rejeitar com erro de 32 % (superior ao $\alpha = 25,2\%$, inicialmente estabelecido com $b_c = 0,797$). O mesmo ocorre com a probabilidade de aceitar, em que os erros são “particionados” em 1,1 % de chance de ocorrência de dados com α igual a 1,1 % (inferior ao $\beta = 25,2\%$) e 24,1% de chance de ocorrência de dados quando se apresenta probabilidade de erro igual a 31 % (superior ao $\beta = 25,2\%$). Isso reflete o objetivo do teste com taxas de erros freqüentistas condicionais de declarar uma medida de maior confiança para valores mais afastados da fronteira da região crítica. Observe que a região crítica continua sendo estabelecida pré-experimentalmente, mas a diferença básica é que, nesse tipo de teste, a probabilidade de erro apresentada não é aquela fixada para definir a região crítica, e sim uma taxa de erro dependente do conjunto de dados observado e, por isso, condicional. O que a teoria de testes com erros freqüentistas condicionais propõe é uma partição da região crítica (e também da região de aceitação) de modo que a probabilidade de erro apresentada seja dependente da distância dos dados observados em relação à fronteira da região crítica.

E importante lembrar que um freqüentista, ao realizar um teste de hipótese tradicional, rejeitaria a hipótese nula nos dois casos comentados no exemplo 1 (c) ($b = 0,69$ ou $b = 0,15$), pois $b_c = 0,797$, mas com o tamanho do teste fixado em $\alpha = \beta = 25,2\%$ para $n = 4$, e se fosse realizado também um teste de significância, o *valor-p* calculado para $b = 0,69$ seria igual a 0,2216, e para $b = 0,15$, a 0,0123, indicando maior evidência para rejeição de H_0 no caso em que $b = 0,15$. Mas os *valores-p* não têm interpretação como probabilidades de erro freqüentistas e esse pode ser um procedimento inadequado em alguns casos, como o de realizar o teste híbrido comparando α com o *valor-p*, como comentado no item 2 do capítulo 1.

Outros possíveis valores de c , para o exemplo 1(c) em que $n = 4$, foram estabelecidos e os valores de d , α_{be} , β_{ae} e β_{be} resultantes foram calculados fixando a probabilidade de erro condicional tipo I em $\alpha_{ae} = 0,10$. Os resultados estão exibidos na Tabela 2.1 seguinte.

TABELA 2.1 - Alguns possíveis valores para limitar as regiões de alta e baixa evidência e seus erros resultantes para $n = 4$ e $\alpha_{ae} = 0,1$.

c	0,07	0,08	0,1	0,2	0,25	0,3	0,31	0,315	0,319
d	102×10^4	24806	1048	8,592	2,846	1,129	0,943	0,861	0,799
α_{be}	0,2522	0,2524	0,2544	0,3195	0,4184	0,6981	0,8251	0,911	0,996
β_{ae}	0,00	0,00	0,00	0,0531	0,1518	0,3086	0,3454	0,3643	0,3799
β_{be}	0,2524	0,2528	0,2564	0,3079	0,3081	0,1825	0,1082	0,0558	0,0020

Nota-se que, para o valor de α_{ae} estabelecido, uma das menores probabilidades de erro condicional tipo I obtidas para a região de baixa evidência foi próxima ao próprio erro usual para este caso, que é de 25,2%, com um valor de d muito afastado do valor crítico $b_c = 0,797$. Da mesma forma que a probabilidade de erro tipo I na região de baixa evidência tende a ser muito maior à medida que d se aproxima do valor crítico 0,797 estabelecido para o teste freqüentista. Observa-se também que, como o tamanho de amostra é pequeno ($n = 4$), o teste se torna muito robusto, pois quando o valor de c passa de 0,2 para 0,319, a probabilidade de erro tipo I na região de baixa evidência é praticamente 1.

O valor de d deve ser maior que o valor crítico, ou seja, $0,797 < d < \infty$.
 Como o valor de c está relacionado ao valor de d , este também fica limitado.
 Assim, para esse caso, o valor máximo para c é:

$$\alpha_{ae} = 0,1 = \frac{\int_0^{\ln[16c]} y^3 e^{-y} dy}{\int_0^{\ln[16c]} y^3 e^{-y} dy + \int_{\ln[16d]}^{\infty} y^3 e^{-y} dy} \Rightarrow$$

$$0,9 \left(\int_0^{\ln[16c]} y^3 e^{-y} dy \right) = 0,1 \left(\int_{\ln[16(0,797)]}^{\infty} y^3 e^{-y} dy \right) \Rightarrow$$

$$\int_0^{\ln[16c]} y^3 e^{-y} dy = \frac{1}{9} \left(\int_{2,546}^{\infty} y^3 e^{-y} dy \right) \approx 0,0831 \Rightarrow c \cong 0,32$$

Portanto, o maior valor admissível para c deve ser menor que 0,32, o que explica o $\alpha_{be} = 0,996$ quando isso ocorre.

Para valores do limite c da região de alta evidência inferiores a 0,1, o intervalo resultante $[c,d]$ corresponde a mais de 99% da distribuição de $F_B^0(b)$.
 Como consequência, obtêm-se probabilidades de erros praticamente iguais às do teste não-condicional, sendo $\alpha_{be} \cong \alpha$ e $\beta_{be} \cong \beta$, enquanto $\alpha_{ae} \cong 0$ e $\beta_{ae} \cong 0$.

Uma observação interessante é que, ponderando os dois erros pela probabilidade da região que os define, obtêm-se a taxa de erro tipo I e II do teste freqüentista não-condicional, que é de 25,2%.

$$\alpha_{ae} \times P_{\theta_0}([0,0625;c] \cup [d;\infty)) + \alpha_{be} \times P_{\theta_0}([c;0,797] \cup]0,797;d[) =$$

$$= \frac{P_{\theta}([0,0625;c])}{P_{\theta}([0,0625;c] \cup [d;\infty))} P_{\theta}([0,0625;c] \cup [d;\infty)) +$$

$$\frac{P_{\theta}([c;0,797])}{P_{\theta}([c;0,797] \cup]0,797;d[)} P_{\theta}([c;0,797] \cup]0,797;d[) =$$

$$= P_{\theta}([0, 0,625; c]) + P_{\theta}(]c; 0,797]) = P_{\theta}([0, 0,625; 0,797]) =$$

$$\int_{\ln[2^4 \times 0,0625]}^{\ln[2^4 \times 0,797]} \frac{1}{\Gamma(4)} y^3 e^{-y} dy = P_{\theta}(\text{região crítica } C_r) = 0,252$$

Do mesmo modo, tem-se que a taxa de erro tipo II média será:

$$\beta_{ae} \times P_{\theta_1}([0, 0,625; c] \cup [d; \infty)) + \beta_{be} \times P_{\theta_1}(]c; 0,797] \cup]0,797; d[) =$$

$$= \frac{P_{\theta_1}([d; \infty))}{P_{\theta_1}([0, 0,625; c] \cup [d; \infty))} P_{\theta_1}([0, 0,625; c] \cup [d; \infty)) +$$

$$+ \frac{P_{\theta_1}(]0,797; d[)}{P_{\theta_1}(]c; 0,797] \cup]0,797; d[)} P_{\theta_1}(]c; 0,797] \cup]0,797; d[) =$$

$$P([d; \infty)) + P(]0,797; d[) = 1 - P_{\theta_1}(\{([0, 0,625; c]) + P_{\theta_1}(]c; 0,797])\}) =$$

$$= 1 - P_{\theta_1}([0, 0,625; 0,797]) = 1 - \int_{2\ln[2^4 \times 0,0625]}^{2\ln[2^4 \times 0,797]} \frac{1}{\Gamma(4)} y^3 e^{-y} dy = 1 - 0,748 = 0,252$$

Assim, a partição da região crítica e da região de aceitação fornecem testes com probabilidades de erro condicionais aos dados, mas que, em média, ou seja, em termos freqüentistas, apresentam a mesma probabilidade de erro estabelecida inicialmente pelo valor crítico utilizado.

No exemplo apresentado na Figura 2.3 tem-se que a taxa de erro tipo I será, em média:

$$\alpha_{ae} \times P_{\theta_0}([0, 0,625; c] \cup [d; \infty)) + \alpha_{be} \times P_{\theta_0}(]c; 0,797] \cup]0,797; d[) =$$

$$0,10 \times (3\% + 27,6\%) + 0,32 \times (22,2\% + 47,2\%) \cong 25,2\%$$

E a taxa de erro tipo II será, em média:

$$\beta_{ae} \times P_{\theta_0}([0, 0,625; c] \cup [d; \infty)) + \beta_{be} \times P_{\theta_0}(]c; 0,797] \cup]0,797; d[) =$$

$$0,05 \times (1,1\% + 20,6\%) + 0,31 \times (24,1\% + 54,2\%) \cong 25,2\%$$

É importante ressaltar que a partição explicada neste exemplo foi uma partição simples, em que o problema de os erros serem condicionais aos dados

foi parcialmente resolvido, pois seria adequada agora a pergunta: E se fossem observadas duas amostras pertencentes à região de baixa evidência? A probabilidade de erro apresentada seria a mesma para as duas? Isso mostra que o teste ainda deve ser apurado. A idéia inicial foi apresentada nesta seção 2.1.1, mas o ideal é que se tenha uma partição completamente dependente dos dados, ou seja, uma “partição contínua”.

2.2 PARTIÇÃO DO ESPAÇO AMOSTRAL

Um teste estatístico Υ com região crítica C_Υ define uma regra de decisão $\delta(x)$, que é uma função do espaço amostral \mathfrak{X}^n no espaço das decisões, de forma que: $\delta(x) = \text{rejeitar } H_0$ se $x \in C_\Upsilon$, $\delta(x) = \text{aceitar } H_0$ se $x \in \overline{C_\Upsilon}$.

Considerando agora uma estatística Z , que será denominada de *estatística condicionante*, em um primeiro momento será considerado o caso em que a estatística Z assume apenas um número finito de valores z_1, \dots, z_m .

Considere também:

- $C^{z_i} = \{x \in \mathfrak{X}^n ; Z(x) = z_i\}$ o conjunto das amostras para as quais a estatística Z vale z_i .
- $C_0^{z_i} = C^{z_i} \cap \overline{C_\Upsilon}$, isto é, conjunto das amostras para as quais a estatística condicionante vale z_i e a decisão tomada é aceitar a hipótese nula.
- $C_1^{z_i} = C^{z_i} \cap C_\Upsilon$, isto é, conjunto das amostras para as quais a estatística condicionante vale z_i e a decisão tomada é recusar H_0 (Figura 2.4).

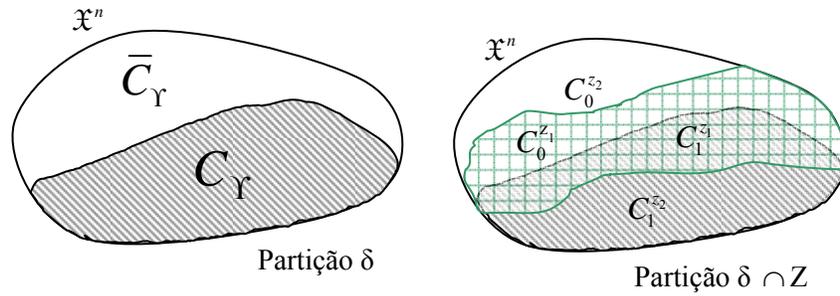


FIGURA 2.4 - Desenho esquemático considerando as duas partições definidas por δ e por $Z = \{z_1, z_2\}$.

Os valores da estatística condicionante z_1 e z_2 corresponderiam aos limites das regiões de alta e baixa evidência apresentadas no exemplo 1(c). Lembrando que para baixa evidência, isto é, quando $Z = z_1$, poderia ocorrer aceitação ($C_0^{z_1}$) ou rejeição ($C_1^{z_1}$) da hipótese nula (área quadriculada). O mesmo ocorrendo para $Z = z_2$, que corresponde às regiões de alta evidência para aceitação ($C_0^{z_2}$) e rejeição ($C_1^{z_2}$).

Então, a cada conjunto C^{z_i} está associada uma taxa de erro condicional tipo I, dada por $\alpha_{z_i} = \frac{P_{\theta_0} [C_1^{z_i}]}{P_{\theta_0} [C^{z_i}]}$, e uma taxa de erro condicional tipo II

$$\beta_{z_i} = \frac{P_{\theta_1} [C_0^{z_i}]}{P_{\theta_1} [C^{z_i}]}.$$

As denominações:

α_{z_i} = Probabilidade de se cometer um erro Tipo I dado que $Z = z$ foi observada $P_{\theta_0}[(\text{Rejeitar } H_0 \mid C^{z_i})]$ e

β_{z_i} = Probabilidade de se cometer um erro Tipo II dado que $Z = z$ foi observada $P_{\theta_1}[(\text{Aceitar } H_0 | C^{z_i})]$
serão chamadas de “*probabilidades de erro condicionais*” e abreviadas por “*PEC*”.

A idéia é que a taxa de erro não condicional é distribuída através da partição definida pela estatística Z em erros condicionais. Esta distribuição é tal que, ponderando os erros condicionais α_{z_i} pela probabilidade de C^{z_i} ocorrer, tem-se a probabilidade de erro não condicional. De fato, como para cada valor da variável aleatória $Z = z_i$ tem-se uma probabilidade de erro condicional α_{z_i} , pode-se considerar estas como variáveis aleatórias α_{z_i} e β_{z_i} . Tem-se:

Proposição:

$$E_{\theta_0} [\alpha_{z_i}] = \alpha \quad \text{e} \quad E_{\theta_1} [\beta_{z_i}] = \beta.$$

Demonstração:

Sendo $\alpha_{z_i} = P_{\theta_0} [\text{Erro Tipo I} | C^{z_i}] = \frac{P_{\theta_0} [C_Y \cap C^{z_i}]}{P_{\theta_0} [C^{z_i}]}$, tem-se:

$$E_{\theta_0} [\alpha_{z_i}] = \sum_z \frac{P_{\theta_0} [C_Y \cap C^{z_i}]}{P_{\theta_0} [C^{z_i}]} \cdot P_{\theta_0} [C^{z_i}] = P_{\theta_0} [C_Y] = P_{\theta_0} [\text{Erro Tipo I}] = \alpha$$

Para o erro tipo II, tem-se:

$$\beta_{z_i} = P_{\theta_1} [\text{Erro Tipo II} | C^{z_i}] = \frac{P_{\theta_1} [\bar{C}_Y \cap C^{z_i}]}{P_{\theta_1} [C^{z_i}]} \quad \text{e}$$

$$E_{\theta_1} [\beta_{z_i}] = \sum_z \frac{P_{\theta_1} [\bar{C}_Y \cap C^{z_i}]}{P_{\theta_1} [C^{z_i}]} \cdot P_{\theta_1} [C^{z_i}] = P_{\theta_1} [\bar{C}_Y] = P_{\theta_1} [\text{Erro Tipo II}] = \beta.$$

Ou seja, a esperança da probabilidade de erro condicional é a própria probabilidade de erro freqüentista estipulada no teste freqüentista de Neyman-Pearson.

Esta proposição vale também para o caso em que Z é contínua, o que será demonstrado posteriormente.

Considerar que a estatística condicionante Z assume apenas um número finito de valores é muito restritivo. Para os casos em que Z assume valores reais seria necessário discretizar o problema, o que leva a uma situação de aproximação. Em razão disso, Kiefer (1977) desenvolve sua teoria considerando uma “partição contínua” definida por estatísticas que assumem valores reais. Todavia, ele não desenvolve todas as passagens matemáticas dessa construção no artigo, no qual uma explicitação mais detalhada certamente tornaria seu trabalho mais acessível à comunidade científica em geral, o que tornaria o impacto prático de sua teoria mais acentuado. Esse trabalho pretende, como contribuição, explicitar ao máximo as passagens matemáticas necessárias para desenvolver o conceito de partição contínua e recomendar uma estrutura padrão para testes freqüentistas condicionais baseados em diferentes estatísticas condicionantes.

3 METODOLOGIA

3.1 Partição Contínua

Desenvolveu-se detalhadamente toda a teoria matemática de testes freqüentistas condicionais utilizando quatro diferentes estatísticas condicionantes que fornecem diferentes partições contínuas. As partições resultam das estatísticas condicionantes baseadas nas seguintes situações:

3.1.1 Estatística Ancilar;

3.1.2 Níveis de Significância Intrínseca;

3.1.3 Probabilidade de Erros Condicionais Iguais;

3.1.4 *Valores-p*.

Forneceu-se uma estrutura padrão para testes freqüentistas condicionais baseados em diferentes estatísticas condicionantes a partir da denominada função particionante $H(x)$.

3.2 Comparação entre estatísticas condicionantes

Para explicar algumas possibilidades de condicionamento foram utilizados dois exemplos explorando as quatro formas de estatísticas condicionantes e a construção da partição em cada um dos casos desenvolvidos no item 3.1.

O primeiro exemplo foi desenvolvido para razão de verossimilhanças simétrica considerando um teste de hipóteses $H_0: X \sim \text{Beta}(2,3)$ versus $H_1: X \sim \text{Beta}(3,2)$. Para o segundo exemplo, com razão de verossimilhanças não-simétrica, foi desenvolvido o teste de hipóteses $H_0: X \sim \text{Uniforme}(0,1)$ versus $H_1: X \sim \text{Beta}\left(\frac{1}{2}, 1\right)$.

Como o objetivo principal desse trabalho é estudar a teoria de testes freqüentistas condicionais e o desenvolvimento da teoria utilizando o conceito

de função particionante, os dois exemplos foram resolvidos para tamanho de amostra 1. No caso de amostra de tamanho maior, deve-se considerar a utilização de recursos computacionais para o desenvolvimento da distribuição da razão de verossimilhanças, o que não é a finalidade do presente estudo.

3.3 Região de Não Decisão em Testes Frequentistas Condicionais

Em caso de discordância entre as decisões nos testes utilizando as diferentes formas de condicionamento, propôs-se a inclusão de uma região de não-decisão (RND), restringindo o espaço amostral e considerando apenas a região em que há concordância entre as decisões. Para comprovar como a inclusão da região de não-decisão afeta a decisão final, foi utilizado um programa no SAS (1990) (Anexo A), em que o processo foi repetido 100.000 vezes e a probabilidade de ocorrência da RND foi calculada e comparada com a calculada via integração.

3.4 Distribuição da Estatística Condicionante baseada em *Valores-p*

Foi realizado um estudo sobre a distribuição da estatística condicionante baseada em *valores-p*.

3.5. Exemplo de Aplicação

A teoria estudada foi aplicada a um exemplo com dados reais para o teste de hipóteses de duas variedades de café. Foram utilizadas amostras obtidas de um experimento realizado pelo Setor de Fisiologia Vegetal do Departamento de Biologia da Universidade Federal de Lavras, em 2006. Entre as várias análises realizadas no experimento, selecionou-se a variável “teor de amido presente na raiz da planta”, estudada em três plantas das variedades de café Catuaí e Icatu.

4 RESULTADOS E DISCUSSÃO

4.1. Partição Contínua

Considere uma estatística Z com valores em \mathbb{R} . Uma partição em \mathcal{X}^n é, então, equivalente a uma partição em \mathbb{R} . O primeiro problema com a teoria é que uma partição de \mathbb{R} é uma situação muito geral para definir boas propriedades condicionais. É necessário obter partições com boas propriedades.

O fato é que uma partição adequada estará intimamente relacionada a uma boa estatística condicionante. A construção de uma partição adequada pode ser exemplificada, considerando que se deve decidir entre $f(\cdot|\theta_0) = f_0$ e

$f(\cdot|\theta_1) = f_1$, e será utilizado um teste de razão de verossimilhanças, sendo

$B(x) = \frac{f_0(x)}{f_1(x)}$ com região crítica $b \leq b_c$. Em muitas situações essa desigualdade

é equivalente à desigualdade $T(x) \leq k$, em que $T(X)$ é a estatística suficiente.

Tem-se o seguinte:

Considere o caso em que $n = 1$ e o espaço amostral \mathcal{X} tem dimensão 1, isto é, $\mathcal{X} = \mathbb{R}$. A idéia é partir \mathbb{R} em subconjuntos formados, por exemplo, por dois pontos $\{x', x\}$, tais que $x' < k$ e $x > k$, sendo k o valor crítico. Se x' ocorre, H_0 é rejeitada e H_0 é aceita se x ocorre.

Quando x' é observado, a hipótese nula é rejeitada. O erro ocorre se f_0 é a verdadeira distribuição. Então, se f_0 é verdadeira, qual é a probabilidade de se obter x' dado que x ou x' poderiam ter ocorrido?

Para que possam ser definidas probabilidades, considerem-se os intervalos $[x, x + \Delta x]$ e $[x', x' + \Delta x]$. Dessa maneira, podem ser definidas as probabilidades A_0, A_0', A_1 e A_1' , em que A_i corresponde à área sob a curva f_i no

intervalo $[x, x + \Delta x]$ e A_0' corresponde à área sob a curva f_0 no intervalo $[x', x' + \Delta x]$ (Figura 2.5).

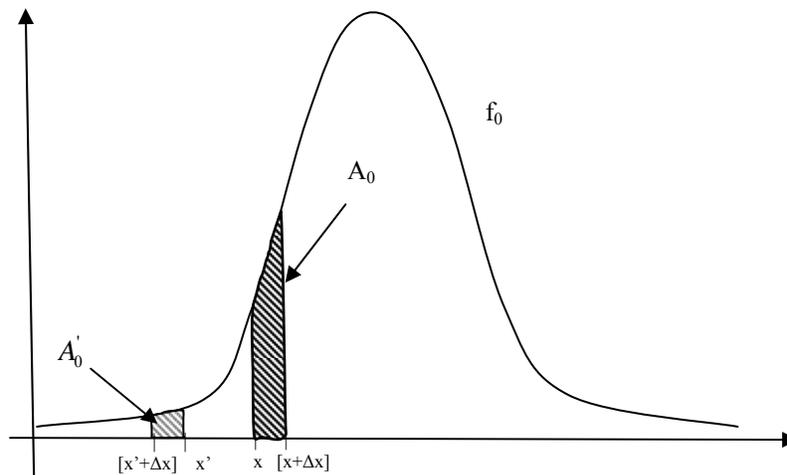


FIGURA 2.5 - Representação das áreas infinitesimais sob f_0 .

Nesse caso, a Probabilidade de Erro Condicional Tipo I (PEC I) é dada por $\alpha_{x'} = \frac{A_0'}{A_0' + A_0}$; ou seja, sob a hipótese nula, o erro condicional é dado pela

probabilidade de ocorrer uma amostra no intervalo $[x', x' + \Delta x]$, dado que poderiam ter ocorrido $[x', x' + \Delta x]$ ou $[x, x + \Delta x]$. Da mesma forma, a Probabilidade de Erro Condicional Tipo II (PEC II), que acontece se x é

observado e aceita-se H_0 quando esta é falsa, é dada por $\beta_x = \frac{A_1}{A_1' + A_1}$

(Figura 2.6).

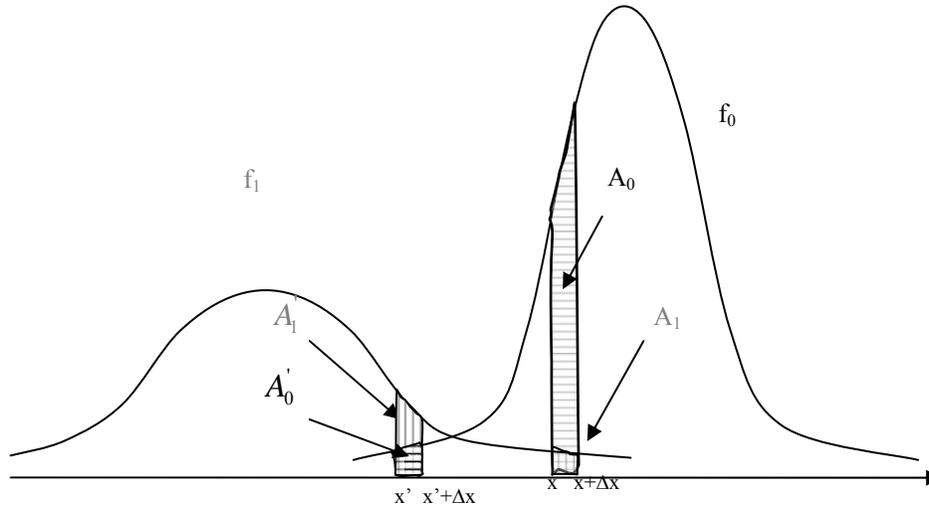


FIGURA 2.6 - Desenho esquemático da relação entre as áreas associadas aos pontos x e x' em f_0 e f_1 com $k \in]x'+\Delta x, x[$.

Para o caso em que as áreas são pequenas, pode-se utilizar as aproximações $A_0 \approx f_0(x)\Delta x$ e $A'_0 \approx f_0(x')\Delta x$. Então

$$\alpha_{x'} \approx \frac{f_0(x')\Delta x}{f_0(x)\Delta x + f_0(x')\Delta x} = \frac{f_0(x')}{f_0(x) + f_0(x')} \quad \text{e}$$

$$\beta_x \approx \frac{f_1(x)\Delta x}{f_1(x)\Delta x + f_1(x')\Delta x} = \frac{f_1(x)}{f_1(x) + f_1(x')}.$$

Entretanto, construídos dessa forma, x e x' não têm dependência nenhuma. Uma maneira mais apropriada de se obterem os pontos $\{x, x'\}$ para definir uma partição do espaço amostral é estabelecer uma relação funcional

entre x e x' , de modo que x' seja uma função de x , isto é, $x' = H(x)$. em que $H(x)$ será denominada “*função particionante*” (Figura 2.7).

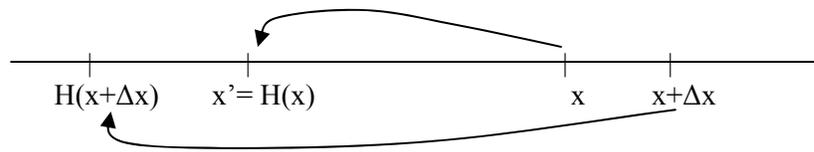


FIGURA 2.7 – Relação funcional entre x e x' .

Então, a “*função particionante*” $H(x)$ é a função responsável pelo critério de associação entre o ponto x , que pertence à “*região de aceitação*”, e o ponto x' , que pertence à “*região de rejeição*”. Ou seja, é a função que define o critério para uma determinada partição contínua do espaço amostral. Note que uma propriedade desejável da função H é que ela seja estritamente decrescente.

Definida uma relação funcional entre x e x' , fica também estabelecida uma relação funcional entre as áreas A_0 e A_0' e A_1 e A_1' . Tal fato ocorre porque, ao se tomar o intervalo $[x, x + \Delta x]$, o intervalo em x' fica determinado como $[H(x), H(x + \Delta x)]$.

A relação entre as áreas é

$$A_0 = \int_x^{x+\Delta x} f_0(u)du \quad \text{e} \quad A_0' = \int_{H(x)}^{H(x+\Delta x)} f_0(u)du$$

$A_0' = |F_0(H(x + \Delta x)) - F_0(H(x))|$ e, pelo Teorema do Valor Médio para Integrais (Larson, Hostetler & Edwards, 2001), existe $\xi \in [H(x + \Delta x), H(x)]$, tal que

$$A_0' = |F_0(H(x + \Delta x)) - F_0(H(x))| = f_0(\xi) |H(x + \Delta x) - H(x)| \Rightarrow$$

$$A_0' = f_0(\xi) \left[\frac{|H(x + \Delta x) - H(x)|}{\Delta x} \right] \Delta x$$

O valor absoluto é necessário, pois $H(x)$ é decrescente e, portanto, $H'(x)$ é negativa.

A função $H(x)$ define, então, uma “partição contínua” e será denominada função particionante. Considerando esta partição, as probabilidades de erros condicionais são dadas por

$$\alpha_{x'} = \frac{A_0'}{A_0' + A_0} = \frac{f_0(\xi) \left[\frac{|H(x + \Delta x) - H(x)|}{\Delta x} \right] \cancel{\Delta x}}{f_0(\xi) \left[\frac{|H(x + \Delta x) - H(x)|}{\Delta x} \right] \cancel{\Delta x} + f_0(x) \cancel{\Delta x}} \Rightarrow$$

$$\alpha_{x'} = \lim_{\Delta x \rightarrow 0} \frac{f_0(\xi) \left[\frac{|H(x + \Delta x) - H(x)|}{\Delta x} \right]}{f_0(\xi) \left[\frac{|H(x + \Delta x) - H(x)|}{\Delta x} \right] + f_0(x)} \Rightarrow$$

$$\Rightarrow \alpha_{x'} = \frac{f_0(H(x)) |H'(x)|}{f_0(H(x)) |H'(x)| + f_0(x)} \quad (2.5)$$

e

$$\beta_x = \frac{A_1}{A_1 + A_1'} = \frac{f_1(x) \cancel{\Delta x}}{f_1(x) \cancel{\Delta x} + f_1(\xi) \left[\frac{|H(x + \Delta x) - H(x)|}{\Delta x} \right] \cancel{\Delta x}} \Rightarrow$$

$$\beta_x = \lim_{\Delta x \rightarrow 0} \frac{f_1(x)}{f_1(x) + f_1(\xi) \left[\frac{|H(x + \Delta x) - H(x)|}{\Delta x} \right]} \Rightarrow$$

$$\Rightarrow \boxed{\beta_x = \frac{f_1(x)}{f_1(x) + f_1(H(x))|H'(x)|}} \quad (2.6).$$

Então, o teste freqüentista condicional Υ^c é dado pela estrutura geral apresentada a seguir:

$$\Upsilon^c = \begin{cases} \text{se } x \leq k, \text{ rejeita-se } H_0 \text{ e apresenta-se} \\ \text{PEC I } \alpha_x = \left(\frac{f_0(x)}{f_0(x) + f_0(H(x))|H'(x)|} \right) \\ \text{se } x > k, \text{ aceita-se } H_0 \text{ e apresenta-se} \\ \text{PEC II } \beta_x = \left(\frac{f_1(x)}{f_1(x) + f_1(H(x))|H'(x)|} \right) \end{cases}$$

- Uma observação sobre a notação:

Ao denotar a PEC I por α_x e PEC II por β_x , deve-se notar que apenas um tipo de erro pode ser cometido. Se $x \leq k$, apenas o erro tipo I pode ocorrer, e se $x > k$, apenas o erro tipo II.

Utilizando as definições das PEC's, pode-se demonstrar que a $E_{\theta_0}[\alpha_Z] = \alpha$ e $E_{\theta_1}[\beta_Z] = \beta$ para o caso em que Z é uma partição contínua.

Demonstração:

Para a probabilidade de erro tipo I, sendo k o valor crítico do teste, tem-se:

Sendo, pela expressão (2.5), $\alpha_{x'} = \frac{f_0(H(x))H'(x)}{f_0(H(x))H'(x) + f_0(x)}$ e

$$P_{\theta_0} [A_0 + A_0'] = f_0(H(x))H'(x)\Delta x + f_0(x)\Delta x, \text{ tem-se:}$$

$$E_{\theta_0} [\alpha_{x'}] = \int_k^{\infty} \left(\frac{f_0(H(x))H'(x)}{f_0(H(x))H'(x) + f_0(x)} \sum_{i=1}^{\infty} (f_0(H(x))H'(x) + f_0(x))\Delta x_i \right) dx$$

$$E_{\theta_0} [\alpha_{x'}] = \int_k^{\infty} \left(\frac{f_0(H(x))H'(x)}{f_0(H(x))H'(x) + f_0(x)} \cdot \frac{(f_0(H(x))H'(x) + f_0(x)) \sum_{i=1}^{\infty} (\Delta x_i)}{(f_0(H(x))H'(x) + f_0(x)) \sum_{i=1}^{\infty} (\Delta x_i)} \right) dx$$

Como $\sum_{i=1}^{\infty} (\Delta x_i)$ é a área total sob f_0 , ou seja, 1, então:

$$E_{\theta_0} [\alpha_{x'}] = \int_k^{\infty} (f_0(H(x))H'(x))dx = 1 - F_0(k) = P_{\theta_0} [C_Y] = \text{Erro Tipo I} = \alpha$$

E para a probabilidade de erro tipo II, tem-se, de acordo com a expressão (2.6), $\beta_x = \frac{f_1(x)}{f_1(H(x))H'(x) + f_1(x)}$ e

$$P_{\theta_1} [A_1 + A_1'] = f_1(H(x))H'(x)\Delta x + f_1(x)\Delta x$$

$$E_{\theta_1} [\beta_x] = \int_{-\infty}^k \left(\frac{f_1(x)}{f_1(H(x))H'(x) + f_1(x)} \sum_{i=1}^{\infty} (f_1(H(x))H'(x) + f_1(x))\Delta x_i \right) dx$$

$$E_{\theta_1} [\beta_x] = \int_{-\infty}^k \left(\frac{f_1(x)}{f_1(H(x))H'(x) + f_1(x)} \cdot \frac{(f_1(H(x))H'(x) + f_1(x)) \sum_{i=1}^{\infty} (\Delta x_i)}{(f_1(H(x))H'(x) + f_1(x)) \sum_{i=1}^{\infty} (\Delta x_i)} \right) dx$$

$$E_{\theta_1} [\beta_x] = \int_{-\infty}^k f_1(x)dx = F_1(k) = P_{\theta_1} [\bar{C}_Y] = \text{Erro Tipo II} = \beta$$

■

4.1.1. Partição com Probabilidades de Erros Condicionais Iguais

Construídas no nível de generalidade apresentado no item 4.1, a função particionante $H(x)$ define partições que podem não possuir boas propriedades para as probabilidades de erros condicionais. Portanto, é necessário procurar alguma restrição. Kiefer (1977) recomendou, em seu artigo, após toda a explicação sobre a teoria da partição do espaço amostral, uma partição denominada por ele de “partição contínua de probabilidade igual”, isto é, uma partição que fornece valores numericamente iguais para as PEC’s Tipo I e II, ou seja,

$$\alpha_x = \beta_{x'}.$$

Assim sendo, a propriedade de interesse é que dentro de cada conjunto da partição se tenham pontos para os quais ocorra a “mesma força de evidência”, tanto para rejeitar quanto para aceitar a hipótese nula. Entenda-se “força de evidência” como uma grandeza que indica, por meio de cálculos de probabilidade, a existência de indícios para rejeição ou aceitação de uma hipótese.

Nesse caso, tem-se, de acordo com as expressões (2.5) e (2.6):

$$\begin{aligned} \frac{A_0'}{A_0' + A_0} &= \frac{A_1}{A_1' + A_1} \Rightarrow \frac{f_0(H(x))|H'(x)|}{f_0(H(x))|H'(x)| + f_0(x)} = \frac{f_1(x)}{f_1(H(x))|H'(x)| + f_1(x)} \\ f_0(H(x))|H'(x)|f_1(x) + f_0(H(x))|H'(x)|f_1(H(x))|H'(x)| &= \\ f_1(x)f_0(H(x))|H'(x)| + f_1(x)f_0(x) & \\ \frac{f_0(H(x))|H'(x)|f_1(x) + f_1(H(x))f_0(H(x))|H'(x)|^2}{f_1(x)f_0(H(x))|H'(x)| + f_0(x)f_1(x)} &= \\ \frac{f_1(x)f_0(H(x))|H'(x)| + f_0(x)f_1(x)}{f_1(H(x))f_0(H(x))|H'(x)|^2} &= f_0(x)f_1(x) \end{aligned}$$

$$\boxed{|H'(x)| = \sqrt{\frac{f_1(x)f_0(x)}{f_1(H(x))f_0(H(x))}} .$$

Integrando a equação diferencial

$$\sqrt{f_1(H(x))f_0(H(x))}|H'(x)| = \sqrt{f_1(x)f_0(x)} ,$$

tem-se

$$\int_x^\infty \sqrt{f_1(H(u))f_0(H(u))}|H'(u)| du = \int_x^\infty \sqrt{f_1(u)f_0(u)} du .$$

Portanto, $H(x)$ é solução da equação diferencial acima que pode ser colocada na forma de uma equação integral.

Fazendo a substituição de variáveis

$$\begin{aligned} H(u) = t & & u = x & \Rightarrow t = H(x) \\ H'(u) du = dt & & u \rightarrow \infty & \Rightarrow t = H(u) \rightarrow -\infty , \\ |H'(u)| du = -dt & & & \end{aligned}$$

tem-se:

$$\begin{aligned} - \int_{H(x)}^{-\infty} \sqrt{f_1(t)f_0(t)} dt &= \int_x^\infty \sqrt{f_1(u)f_0(u)} du \Rightarrow \\ \boxed{\int_{-\infty}^{H(x)} \sqrt{f_1(t)f_0(t)} dt} &= \int_x^\infty \sqrt{f_1(u)f_0(u)} du . \end{aligned} \quad (2.7)$$

Esta equação integral define implicitamente a função particionante $H(x)$.

Como $x' = H(x)$ é razoável exigir, pela simetria da construção, que $H(x') = x$. Dessa forma, $H(x)$ fica definida em \mathbb{R} (ou \mathbb{R}^+) e $H(H(x)) = x$. Como H deve ser decrescente, é necessário que exista um valor a^* tal que $H(a^*) = a^*$. Esse valor a^* será denominado valor crítico e dado por:

$$\boxed{\int_{-\infty}^{a^*} \sqrt{f_1(H(x))f_0(H(x))}|H'(x)| dx} = \int_{a^*}^\infty \sqrt{f_1(x)f_0(x)} dx . \quad (2.8)$$

Obs.: A existência do ponto fixo $H(a^*) = a$ decorre do fato que $H(x)$ tem que interceptar a reta $y = x$. Note que $H(x) \rightarrow -\infty$ quando $x \rightarrow \infty$ e, portanto, sempre intercepta a reta $y = x$.

EXEMPLO 1(d)

Para exemplificar a partição com probabilidades de erros condicionais iguais, suponha as densidades exponenciais

$$f_0(x) = e^{-x}, \text{ para } x > 0, \text{ e } f_1(x) = 2e^{-2x}, \text{ para } x > 0.$$

É necessário, para determinação do valor crítico, solucionar a equação integral (2.8):

$$\begin{aligned} \int_0^{a^*} [2e^{-2x} \cdot e^{-x}]^{\frac{1}{2}} dx &= \int_{a^*}^{\infty} [2e^{-3x}]^{\frac{1}{2}} dx \Rightarrow \int_0^{a^*} e^{-\frac{3}{2}x} dx = \int_{a^*}^{\infty} e^{-\frac{3}{2}x} dx \\ \Rightarrow \int_0^{a^*} e^{-\frac{3}{2}x} dx &= \frac{2}{3} \int_{a^*}^{\infty} \frac{3}{2} e^{-\frac{3}{2}x} dx = \frac{2}{3} \left[1 - \int_0^{a^*} \frac{3}{2} e^{-\frac{3}{2}x} dx \right] \Rightarrow \int_0^{a^*} e^{-\frac{3}{2}x} dx = \frac{2}{3} - \int_0^{a^*} e^{-\frac{3}{2}x} dx \\ \Rightarrow \int_0^{a^*} e^{-\frac{3}{2}x} dx &= \frac{2}{3} - \int_0^{a^*} e^{-\frac{3}{2}x} dx \Rightarrow 2 \int_0^{a^*} e^{-\frac{3}{2}x} dx = \frac{2}{3} \Rightarrow \int_0^{a^*} e^{-\frac{3}{2}x} dx = \frac{1}{3} \\ \frac{2e^{-\frac{3}{2}x}}{-3} \Big|_0^{a^*} &= \frac{1}{3} \Rightarrow e^{-\frac{3}{2}x} \Big|_0^{a^*} = -\frac{1}{2} \Rightarrow e^{-\frac{3}{2}a^*} - 1 = -\frac{1}{2} \Rightarrow \ln \frac{1}{2} = -\frac{3}{2} a^* \Rightarrow \boxed{a^* = 0,462}. \end{aligned}$$

Para determinação da função particionante, deve-se resolver a equação integral dada por (2.7):

$$\begin{aligned} \int_0^{H(x)} [2e^{-2y} \cdot e^{-y}]^{\frac{1}{2}} dy &= \int_x^{\infty} [2e^{-3y}]^{\frac{1}{2}} dy \Rightarrow \int_0^{H(x)} e^{-\frac{3}{2}y} dy = \int_x^{\infty} e^{-\frac{3}{2}y} dy \\ \frac{2e^{-\frac{3}{2}y}}{-3} \Big|_0^{H(x)} &= \frac{2}{3} - \left[\frac{2e^{-\frac{3}{2}y}}{-3} \Big|_0^x \right] \Rightarrow \frac{2}{-3} \left(e^{-\frac{3}{2}H(x)} - 1 \right) = \frac{2}{3} - \left[\frac{2}{-3} \left(e^{-\frac{3}{2}x} - 1 \right) \right] \end{aligned}$$

$$\Rightarrow -\left(e^{-\frac{3}{2}H(x)} - 1\right) = \left[1 + \left(e^{-\frac{3}{2}x} - 1\right)\right] \Rightarrow 1 - e^{-\frac{3}{2}H(x)} = e^{-\frac{3}{2}x} \Rightarrow \boxed{H(x) = -\frac{2}{3} \ln\left(1 - e^{-\frac{3}{2}x}\right)}$$

Note que:

$$H(a^*) = -\frac{2}{3} \ln\left(1 - e^{-\frac{3}{2}a^*}\right) = -\frac{2}{3} \ln\left(1 - e^{-\frac{3}{2}(0,462)}\right) = -\frac{2}{3} \ln(0,499) = 0,4621 = a^*$$

(Figura 2.8).

Esse valor de $x = a^* = 0,462$ implica no valor crítico utilizando a razão de verossimilhanças $B(x) = 0,5e^x$, $b_c = 0,794$.

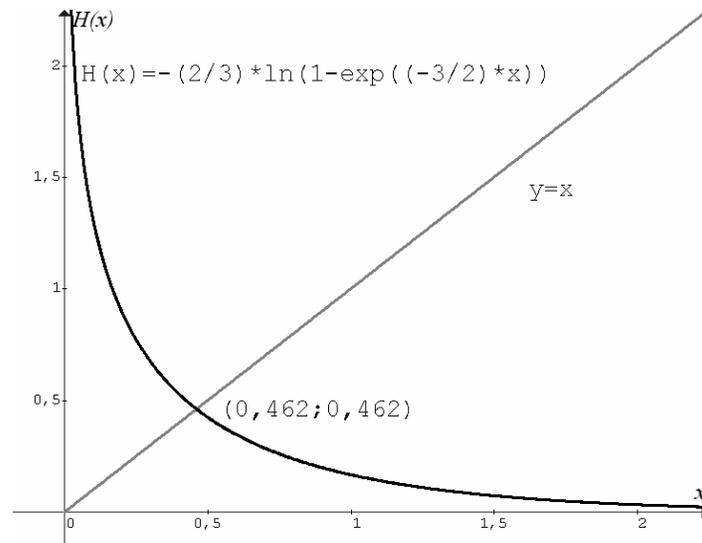


FIGURA 2.8 – Simetria da Função $H(x)$ em relação à reta $y = x$.

Então, com $n = 1$, para o exemplo apresentado, o teste frequentista condicional Y^C , seguindo a estrutura geral de teste apresentada na página 82, é

$$Y^c = \begin{cases} \text{se } b \leq 0,794 \text{ ou } x \leq 0,462, \text{ rejeita-se } H_0 \text{ e apresenta-se} \\ \text{PEC I } \alpha_x = \left(1 + e^{-\frac{x}{2}} \left(1 - e^{-\frac{3}{2}x} \right)^{\frac{1}{3}} \right)^{-1} \\ \text{se } b > 0,794 \text{ ou } x > 0,462, \text{ aceita-se } H_0 \text{ e apresenta-se} \\ \text{PEC II } \beta_x = \left(1 + e^{\frac{x}{2}} \left(1 - e^{-\frac{3}{2}x} \right)^{\frac{1}{3}} \right)^{-1} \end{cases}$$

A representação gráfica das PEC's I e II é dada na Figura 2.9.

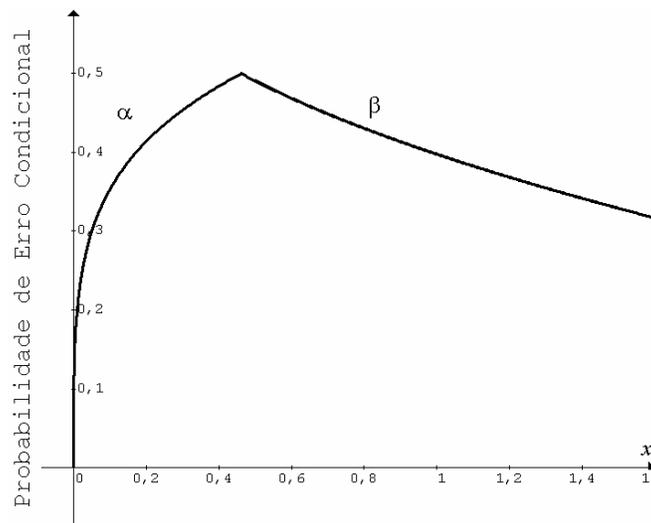


FIGURA 2.9 – Probabilidade de erro condicional tipo I e II resultante da

partição utilizando a função $H(x) = -\frac{2}{3} \ln \left(1 - e^{-\frac{3}{2}x} \right)$.

Supondo que seja observado o valor $x = 0,8$, então a hipótese nula é aceita e a probabilidade de erro condicional Tipo II, apresentada utilizando a expressão (2.6), com $H(0,8) = 0,239$, é

$$\beta_x = \frac{f_1(0,8)}{f_1(0,239)|H'(0,8)| + f_1(0,8)} = \frac{e^{-1,6}}{e^{-0,478} \left| (1 - e^{-1,2})^{-1} e^{-0,4} \right| + e^{-1,6}} = 0,4303.$$

Ou ainda, utilizando a expressão simplificada apresentada no teste Y^C anterior,

$$\text{tem-se } \beta_x = \left(1 + e^{\frac{(0,8)}{2}} \left(1 - e^{-\frac{3}{2}(0,8)} \right)^{\frac{1}{3}} \right)^{-1} = \left(1 + e^{0,4} \left(1 - e^{-1,2} \right)^{\frac{1}{3}} \right)^{-1} = 0,4303.$$

Uma observação importante é que $\alpha_x = \beta_x \neq \alpha = \beta$, ou seja, garantir probabilidades de erros condicionais iguais não garante probabilidades de erros não-condicionais iguais. Esse fato pode ser verificado pela Tabela 1.2, em que o valor crítico difere do valor utilizado neste teste, sendo $b_c = 0,809$ para o critério $\alpha = \beta$ e $b_c = 0,794$ para $\alpha_x = \beta_x$.

4.1.2 PARTIÇÃO CONTÍNUA INTRÍNSECA

Tem-se uma grande liberdade para definir a função particionante $H(x)$. Essencialmente, $H(x)$ deve ser decrescente, $H(H(x)) = x$ e $H(a^*) = a^*$. Uma vez definida a função $H(x)$, as taxas de erros condicionais são:

$$\alpha_x = \frac{f_0(x)}{f_0(H(x))|H'(x)| + f_0(x)} \quad \text{e} \quad \beta_x = \frac{f_1(x)}{f_1(x) + f_1(H(x))|H'(x)|}.$$

Analisando estas equações, uma condição interessante a ser imposta em $H(x)$, proposta por Birnbaum (1961), é de que $H(x)$ satisfaça o sistema de equações diferenciais de tal modo que se obtenham áreas iguais (Figura 2.10):

$$\begin{cases} f_0(H(x'))|H'(x')| = f_1(x') & x' < a^* \\ f_1(H(x))|H'(x)| = f_0(x) & x > a^* \end{cases} \quad (2.9)$$

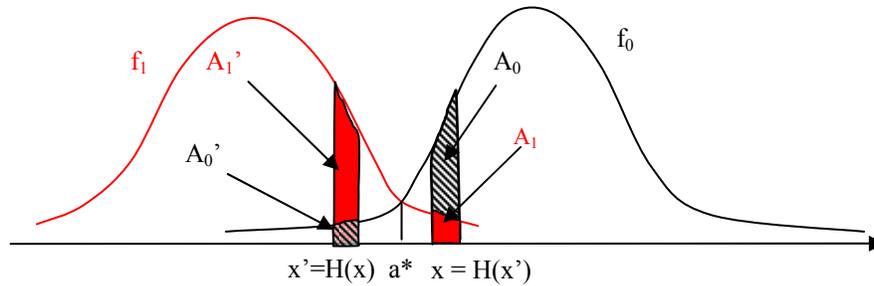


FIGURA 2.10 - Relação entre as áreas sob f_0 e f_1 para a partição intrínseca.

Assim, as PEC's serão simplificadas para:

$$\alpha_{x'} = \frac{A_0'}{A_0' + A_0} = \frac{f_0(x')}{f_0(H(x'))|H'(x')| + f_0(x')} = \frac{f_0(x')}{f_1(x') + f_0(x')} = \frac{\frac{f_0(x')}{f_1(x')}}{\frac{f_1(x')}{f_1(x')} + \frac{f_0(x')}{f_1(x')}} \Rightarrow \alpha_{x'} = \frac{B(x')}{1 + B(x')}$$

e

$$\beta_x = \frac{A_1}{A_1' + A_1} = \frac{f_1(x)}{f_1(H(x))|H'(x)| + f_1(x)} = \frac{f_1(x)}{f_0(x) + f_1(x)} \Rightarrow \beta_x = \frac{1}{B(x) + 1}.$$

Portanto, as probabilidades de erros condicionais são expressas em termos da razão de verossimilhanças.

$H(x)$ é obtida integrando-se o sistema de equações diferenciais.

$$\begin{cases} f_0(H(x'))|H'(x')| = f_1(x') & x' < a^* \\ f_1(H(x))|H'(x)| = f_0(x) & x > a^* \end{cases}$$

Integrando $f_0(H(u))|H'(u)| = f_1(u)$ para $x > a^*$

$$\int_x^\infty f_0(H(u))|H'(u)| du = \int_x^\infty f_1(u) du .$$

Fazendo a substituição de variáveis $t = H(u)$ $dt = H'(u)du$

$$\begin{aligned} u = x &\Rightarrow t = H(x) \\ u \rightarrow \infty &\Rightarrow t = H(u) \rightarrow -\infty \end{aligned} , \text{ tem-se}$$

$$- \int_{H(x)}^{-\infty} f_0(t) dt = \int_x^\infty f_1(u) du \Rightarrow \int_{-\infty}^{H(x)} f_0(t) dt = \int_x^\infty f_1(u) du$$

$$\Rightarrow F_0(H(x)) - 0 = 1 - F_1(x) \Rightarrow$$

$$\boxed{H(x) = F_0^{-1}(1 - F_1(x))} \quad \text{para } x > a^* .$$

Da mesma forma, para $x < a^*$, $H(x)$ é dada por

$$\boxed{H(x) = F_1^{-1}(1 - F_0(x))} . \quad (2.10)$$

4.1.2.1 RAZÃO DE VEROSSIMILHANÇAS SIMÉTRICA

Todos os cálculos para definir as partições e as taxas de erros condicionais foram realizados considerando uma amostra de tamanho 1, isto é, uma única observação x . Nesse caso, as funções densidade de probabilidade envolvidas são $f_0(x) = f(x; \theta_0)$ e $f_1(x) = f(x; \theta_1)$. Igualmente importante é

considerar a estatística razão de verossimilhanças $B(X) = \frac{f_0(X)}{f_1(X)}$. Nesse caso,

as funções de densidade de probabilidade envolvidas são $f_B^0(b)$ e $f_B^1(b)$ (densidades de probabilidades da estatística $B(X)$ supondo, respectivamente, H_0 e H_1 verdadeiras).

PROPOSIÇÃO: $\boxed{f_B^0(b) = bf_B^1(b)}$.

Como $B(X) = \frac{f_0(X)}{f_1(X)}$ usando o teorema de transformação de variáveis, tem-se:

$$f_B^0(b) = f_0(B^{-1}(b)) \left| \frac{d}{db} B^{-1}(b) \right| \quad f_B^1(b) = f_1(B^{-1}(b)) \left| \frac{d}{db} B^{-1}(b) \right|.$$

Logo:

$$\frac{f_B^0(b)}{f_B^1(b)} = \frac{f_0(B^{-1}(b)) \left| \frac{d}{db} B^{-1}(b) \right|}{f_1(B^{-1}(b)) \left| \frac{d}{db} B^{-1}(b) \right|} = \frac{f_0(B^{-1}(b))}{f_1(B^{-1}(b))} = B(B^{-1}(b)) = b.$$

■

Utilizando a variável aleatória $B(X)$ pode-se definir a função particionante $H(b)$, ou seja, considerando a tomada de decisão em relação aos valores observados b da razão de verossimilhanças. As expressões para as PEC's dadas em (2.5) e (2.6) ficam da seguinte maneira:

$$\alpha_{b'} = \frac{f_0(H(b)) |H'(b)|}{f_0(H(b)) |H'(b)| + f_0(b)} \quad \text{e} \quad \beta_b = \frac{f_1(b)}{f_1(b) + f_1(H(b)) |H'(b)|}.$$

Empregando o critério de probabilidade de erros condicionais iguais ($\alpha_{H(b)} = \beta_b$) dado em (2.9), $f_B^1(H(b)) |H'(b)| = f_B^0(b)$, e a proposição anterior, tem-se a seguinte simplificação:

$$\alpha_{b'} = \frac{f_B^0(H(b))|H'(b)|}{f_B^0(H(b))|H'(b)| + f_B^0(b)} = \frac{\frac{f_B^0(H(b))|H'(b)|}{f_B^1(H(b))|H'(b)|}}{\frac{f_B^0(H(b))|H'(b)|}{f_B^1(H(b))|H'(b)|} + \frac{f_B^0(b)}{f_B^1(H(b))|H'(b)|}} =$$

$$\alpha_{b'} = \frac{H(b)}{H(b)+1} = \frac{b'}{b'+1}$$

e

$$\beta_b = \frac{f_B^1(b)}{f_B^1(H(b))H'(b) + f_B^1(b)} = \frac{f_B^1(b)}{f_B^0(b) + f_B^1(b)} \frac{\frac{f_B^1(b)}{f_B^1(b)}}{\frac{f_B^0(b)}{f_B^1(b)} + \frac{f_B^1(b)}{f_B^1(b)}} = \frac{1}{b+1}.$$

Ou seja, fornece $\alpha_b = \frac{b}{1+b}$ e $\beta_b = \frac{1}{1+b}$, uma partição que define as probabilidades de erros condicionais em função de valores observados da própria razão de verossimilhanças, que são os denominados níveis de significância intrínseca (Birnbaum, 1961).

Note que nesse caso têm-se, de fato, probabilidades de erros condicionais iguais, pois, para os pontos b e $\frac{1}{b}$, observa-se que

$$\left. \begin{array}{l} \text{PEC I em } b \Rightarrow \alpha_b = \frac{b}{b+1} \\ \text{PEC II em } \frac{1}{b} \Rightarrow \beta_{\frac{1}{b}} = \frac{1}{\frac{1}{b}+1} = \frac{b}{b+1} \end{array} \right\} \alpha_b = \beta_{\frac{1}{b}}.$$

Quando a distribuição de $B(X)$ sob H_0 é igual à distribuição de $\frac{1}{B(X)}$ sob H_1 , ou seja, $f_B^0(x) = f_{1/B}^1(x)$, a estatística de razão de verossimilhanças $B(X)$ é dita simétrica.

PROPOSIÇÃO: Se a razão de verossimilhanças é simétrica, então

$$\boxed{H(b) = \frac{1}{b}} \quad (2.11)$$

é solução de $f_B^1(H(b))|H'(b)| = f_B^0(b)$.

Prova:

Seja $Z = g(B) = \frac{1}{B}$. Utilizando o teorema de transformação de variáveis:

$$f_Z^1(z) = \left| \frac{d}{dz} g^{-1}(z) \right| f_B^1(g^{-1}(z)) = \frac{1}{z^2} f_B^1\left(\frac{1}{z}\right) \Rightarrow f_{1/B}^1(z) = \frac{1}{z^2} f_B^1\left(\frac{1}{z}\right).$$

Como se trata de razão de verossimilhanças simétrica,

$$f_{1/B}^1(z) = f_B^0(z) \Rightarrow f_B^0(z) = \frac{1}{z^2} f_B^1\left(\frac{1}{z}\right) \Rightarrow f_B^0(z) = |H'(z)| f_B^1(H(z)).$$

■

Uma questão interessante é, de posse da função particionante, obter a estatística condicionante. No caso $H(b) = \frac{1}{b}$, a estatística é $Z = \max\left\{B, \frac{1}{B}\right\}$.

Teorema: Em testes de razão de verossimilhanças, na ocorrência de simetria ($f_B^0 = f_{1/B}^1$) com probabilidade de erro Tipo I e Tipo II iguais ($\alpha = \beta$), tem-se, de acordo com (2.10) ($F_B^0(b) = 1 - F_B^1(b)$), então, $b_c = 1$, isto é, o teste de razão de verossimilhanças é dado por: Rejeita-se H_0 se $b \leq 1$.

Prova:

$$F_B^0(b) = \int_0^b f_B^0(y) dy = \int_0^b f_{1/B}^1(y) dy = P_1 \left[\frac{1}{B} \leq b \right] = P_1 \left[B \geq \frac{1}{b} \right].$$

Então
$$F_B^0(b) = 1 - P_1 \left(B \leq \frac{1}{b} \right) \Rightarrow F_B^0(b) = 1 - F_B^1\left(\frac{1}{b}\right).$$

Com $\alpha = \beta$, tem-se:

$$F_B^0(b) = 1 - F_B^1(b) \Rightarrow 1 - F_B^1\left(\frac{1}{b}\right) = 1 - F_B^1(b) \Rightarrow F_B^1\left(\frac{1}{b}\right) = F_B^1(b) .$$

Como F_B é crescente $\Rightarrow b = \frac{1}{b} \Rightarrow b = 1$.

■

Note que a equação $F_0(1) = 1 - F_1(1)$ implica em $H(1) = F_0^{-1}(1 - F_1(1)) = F_0^{-1}(F_0(1)) = 1$. Assim, a condição $\alpha = \beta$ fica bastante simplificada para o caso de razão de verossimilhanças simétrica.

EXEMPLO 4

Seja o lançamento de um dado viciado com função de probabilidade dada por $H_0: f_0(x) = \frac{x}{21}$ ou por $H_1: f_1(x) = \frac{7-x}{21}$. Assim sendo, têm-se as seguintes probabilidades de ocorrência no lançamento do dado (Figura 2.11):

x	$f_0(x)$	$f_1(x)$
1	1/21	6/21
2	2/21	5/21
3	3/21	4/21
4	4/21	3/21
5	5/21	2/21
6	6/21	1/21

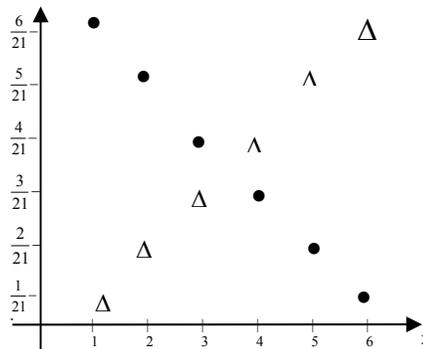


FIGURA 2.11 – Probabilidades em f_0 (Δ) e f_1 (\bullet) do lançamento de um dado

A razão de probabilidade nesse caso será $B(x) = \frac{f_0(x)}{f_1(x)} = \frac{x}{7-x}$. Este é

um caso de razão de verossimilhanças simétrica, pois:

$$f_B^0(b) = \left| \frac{d}{db} g^{-1}(b) \right| f_0(g^{-1}(b)) = f_B^0(b) = \left| \frac{7}{(b+1)^2} \right| f_0\left(\frac{7b}{b+1}\right) \Rightarrow$$

$$f_B^0(b) = \frac{7}{(b+1)^2} \frac{1}{21} \frac{7b}{b+1} \Rightarrow f_B^0(b) = \frac{7b}{3(b+1)^3}.$$

E também:

$$f_{1/B}^1(b) = \left| \frac{d}{db} g^{-1}(b) \right| f_1(g^{-1}(b)) = \left| \frac{-7}{(b+1)^2} \right| f_1\left(\frac{7}{b+1}\right) \Rightarrow f_{1/B}^1(b) = \frac{7b}{3(b+1)^3}.$$

Então, a estatística condicionante adequada é dada pela estatística

$Z = \max\left\{B(x), \frac{1}{B(x)}\right\}$, cuja partição corresponde à utilização da função

particionante $H(x) = \frac{1}{B(x)}$, resultando em $f_0(H(x)) = f_1(x)$. Têm-se, então,

os seguintes resultados da estatística condicionante (Tabela 2.2).

TABELA 2.2 – Resultados da estatística condicionante Z com $B(x) = x/(7-x)$.

x	$B(x)$	$1/B(x)$	Z
1	0,18	6	6
2	0,4	2,5	2,5
3	0,75	1,33	1,33
4	1,33	0,75	1,33
5	2,5	0,4	2,5
6	6	0,18	6

As distribuições de $B(x)$ sob H_0 e sob H_1 podem ser representadas pelos gráficos da Figura 2.12.

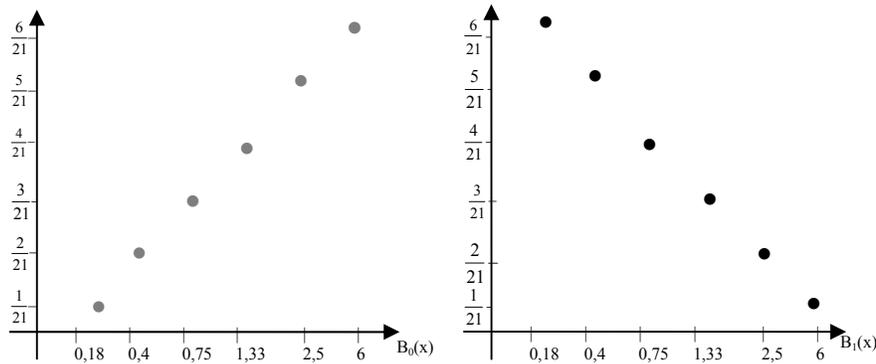


FIGURA 2.12 - Distribuição de $B(x) = x/(7-x)$ sob H_0 e sob H_1 .

Como se trata de razão de verossimilhança simétrica, o teste seria dado por: Rejeite H_0 se $B(x) > 1$ (ou $x > 3$). Supondo que o valor da estatística condicionante Z obtido seja $z = 2,5$. Logo, $B(x) = 2,5$ ou $\frac{1}{B(X)} = 2,5$, que equivale a $B(x) = 0,4$. Como $B(x) = x/(7-x)$, tem-se que, para $B(x) = x/(7-x) = 2,5 \Rightarrow x = 5$, rejeita-se H_0 com

$$\text{PEC I} = \alpha_z = \frac{f_0(2)}{f_0(5) + f_0(2)} = \frac{\frac{2}{21}}{\frac{2}{21} + \frac{5}{21}} = 0,286,$$

e para $B(x) = 0,4 \Rightarrow x = 2$, aceita-se com PEC II

$$\beta_z = \frac{f_1(5)}{f_1(5) + f_1(2)} = \frac{\frac{2}{21}}{\frac{2}{21} + \frac{5}{21}} = 0,286.$$

Ou seja, em caso de partição intrínseca, quando a razão de verossimilhança é simétrica, tem-se $\alpha = \beta$ (probabilidade de erros não-condicionais iguais), e também $\alpha_z = \beta_z$ (PEC's iguais).

4.1.3 ESTATÍSTICA CONDICIONANTE ANCILAR

Uma forma mais tradicional de condicionamento, defendida por alguns estatísticos como Lindley (citado por Berger, Boukai & Wang, 1997), é aquela baseada em uma estatística ancilar, definida como uma estatística Z em que $f_Z^0(x) = f_Z^1(x)$, ou seja, independente das hipóteses testadas.

Teorema:

Para o caso de razão de verossimilhanças simétrica, isto é, $f_B^0 = f_{1/B}^1$,

$$Z = \max \left\{ B(x), \frac{1}{B(x)} \right\} \quad (2.12)$$

é uma estatística ancilar (Figura 2.13).

Demonstração:

$$F_Z^0(z) = P_{\theta_0} [Z \leq z] = P_{\theta_0} \left[\max \left\{ B(x), \frac{1}{B(x)} \right\} \leq z \right]. \quad (1)$$

Todavia, o $\max \left\{ B(x), \frac{1}{B(x)} \right\}$ é a união de dois eventos, $\left\{ \frac{1}{z} \leq B(x) \leq 1 \right\} \cup \{1 \leq B(x) \leq z\}$, como se observa na Figura 2.13, ou simplesmente $\left\{ \frac{1}{z} \leq B(x) \leq z \right\}$. Substituindo em (1), tem-se:

$$F_Z^0(z) = P_{\theta_0} [Z \leq z] = P_{\theta_0} \left[\frac{1}{z} \leq B(x) \leq z \right] \Rightarrow \boxed{F_Z^0(z) = F_B^0(z) - F_B^0\left(\frac{1}{z}\right)}. \quad (2)$$

Como $F_{1/B}^0(z) = P_{\theta_0} \left[\frac{1}{B(x)} \leq z \right] = P_{\theta_0} \left[\frac{1}{z} \leq B(x) \right] = 1 - F_B^0 \left(\frac{1}{z} \right)$, então em

(2) tem-se:

$$F_Z^0(z) = F_B^0(z) - \left[1 - F_{1/B}^0(z) \right].$$

Da mesma forma, para a hipótese alternativa, tem-se

$$\boxed{F_Z^1(z) = F_B^1(z) - F_B^1 \left(\frac{1}{z} \right)}. \quad (3)$$

Mas, como se trata de razão de verossimilhanças simétrica, $F_B^0 = F_{1/B}^1$, então:

$$F_B^0(z) = F_{1/B}^1(z) = P_{\theta_1} \left[\frac{1}{B(x)} \leq z \right] = P_{\theta_1} \left[\frac{1}{z} \leq B(x) \right] = 1 - F_B^1 \left(\frac{1}{z} \right) \text{ e}$$

$$F_B^0 \left(\frac{1}{z} \right) = F_{1/B}^1 \left(\frac{1}{z} \right) = P_{\theta_1} \left[\frac{1}{B(x)} \leq \frac{1}{z} \right] = P_{\theta_1} [z \leq B(x)] = 1 - F_B^1(z).$$

Substituindo em (2) e comparando-se com (3), tem-se:

$$F_Z^0(z) = F_B^0(z) - F_B^0 \left(\frac{1}{z} \right) = \left[1 - F_B^1 \left(\frac{1}{z} \right) \right] - \left[1 - F_B^1(z) \right] = F_B^1(z) - F_B^1 \left(\frac{1}{z} \right) = F_Z^1(z)$$

■

Portanto, $Z = \max \left\{ B(x), \frac{1}{B(x)} \right\}$ é ancilar e

$$Z = \max \left\{ c + \frac{1}{B(x)}, c + B(x) \right\} = c + \max \left\{ \frac{1}{B(x)}, B(x) \right\}$$

também é ancilar, já que a constante c não altera a distribuição de Z .

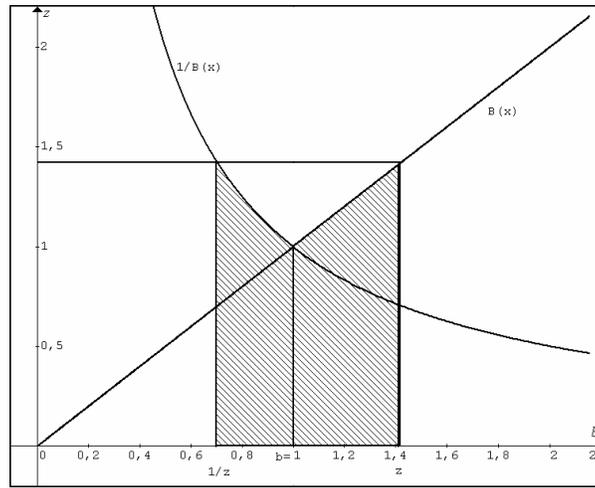


FIGURA 2.13 - Representação da área correspondente a

$$P_{\theta_1} \left[\max \left\{ B(x), \frac{1}{B(x)} \right\} \leq z \right].$$

Para o caso de razão de verossimilhanças não simétrica, deve-se encontrar uma estatística que satisfaça a condição de ancilaridade. Berger, Boukai & Wang (1997) comentam que o condicionamento em estatísticas ancilares pode ser inadequado em algumas situações e podem, até mesmo, ocorrer situações em que não seja possível encontrar uma estatística ancilar.

4.1.4 PARTIÇÃO BASEADA EM VALORES-P

Berger, Brown e Wolpert (1994) abordaram a questão da escolha da estatística condicionante na perspectiva de buscar uma unificação entre o teste freqüentista tradicional e o teste bayesiano, a qual é discutida em detalhes no capítulo 3 desse trabalho. Um outro teste discutido por Paulo (2002) propõe uma unificação do teste freqüentista condicional e do teste bayesiano, e ainda uma

potencial unificação com o teste Fisheriano com a utilização de *valores-p* como estatísticas condicionantes. O teste proposto é o seguinte:

Considere uma estatística unidimensional T e seja $F_T^i(\cdot)$, $i = 0, 1$ a sua função distribuição acumulada, em que estas são funções diferenciáveis e inversíveis no domínio de T . Considere a estatística condicionante

$$Z = \max \{ F_T^0(T), 1 - F_T^1(T) \}. \quad (2.13)$$

Observe que $F_T^0(t)$ e $1 - F_T^1(t)$ são *valores-p* calculados, respectivamente, quando é testada a hipótese H_0 contra H_1 e vice-versa, usando T como uma estatística do teste. Então, Z compara pontos amostrais “pareados” com “a mesma força de evidência” - como medida por *valores-p* - tanto contra H_0 quanto contra H_1 (Figura 2.14).

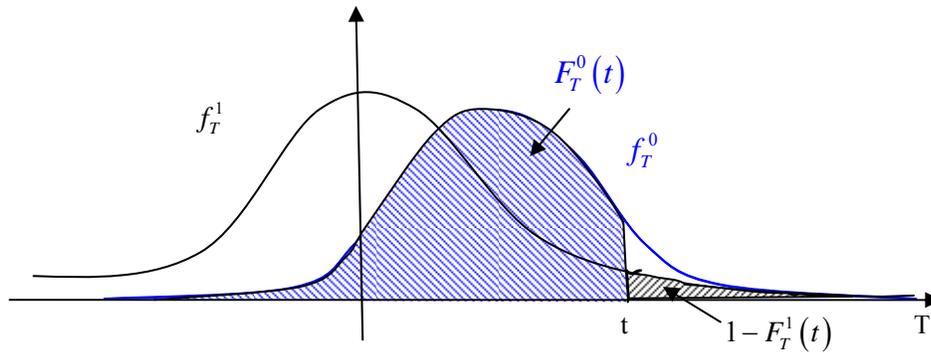


FIGURA 2.14 - Desenho esquemático da relação entre as áreas associadas ao ponto “ t ” sob f_T^0 e f_T^1 .

Dado um valor da estatística $Z = z$, há dois possíveis valores de t . São eles $t = [F_T^0]^{-1}(z)$ ou $t' = [F_T^1]^{-1}(1 - z)$ (Figura 2.15).

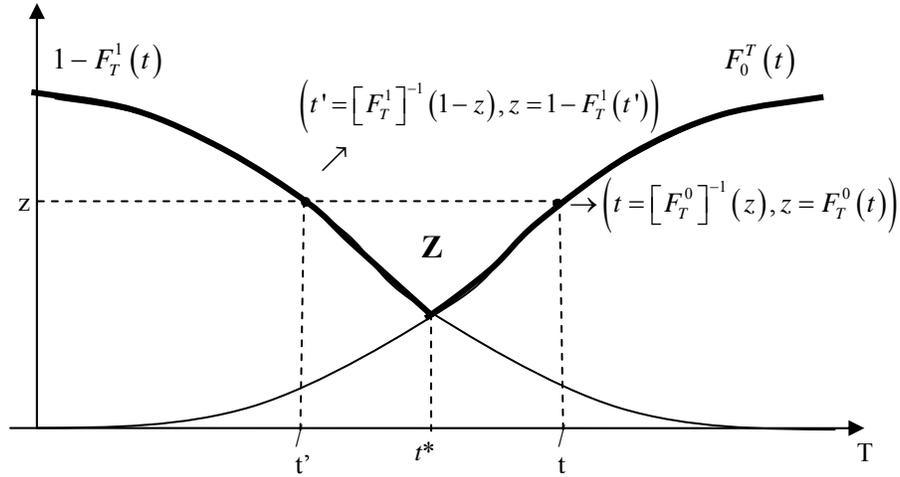


FIGURA 2.15 - Representação esquemática da estatística condicionante *valor-p*.

PROPOSIÇÃO: Sob as pressuposições anteriores, denote por t^* a solução da equação $F_T^0(t) = 1 - F_T^1(t)$. Então, o teste freqüentista condicional com região de rejeição $\{T < t^*\}$ e estatística condicionante $Z = \max\{F_T^0(T), 1 - F_T^1(T)\}$ apresenta as seguintes PEC, tal que, sob a rejeição de H_0 ,

$$\alpha_z \equiv P_0 \{ \text{rejeitar } H_0 \mid Z = z \} \Rightarrow \alpha_z = \frac{B(t)}{1 + B(t)},$$

e sob rejeição de H_1

$$\beta_z \equiv P_1 \{ \text{rejeitar } H_1 \mid Z = z \} \Rightarrow \beta_z = \frac{1}{1 + B(t)},$$

em que t é o valor observado de T e $B(t) = \frac{f_T^0(t)}{f_T^1(t)}$.

Prova: De acordo com a representação esquemática na Figura 2.15, tem-se que:

A função particionante $H(t)$ é obtida a partir do par de pontos $(t, t' = H(t))$ que fornece a mesma “força de evidência”, medidos pelos “valores- p ”, tal que: $F_T^0(H(t)) = 1 - F_T^1(t)$. Tem-se, então, a função particionante $H(t) = [F_T^0]^{-1}(1 - F_T^1(t))$. Esta função particionante (expressão 2.10 na página 91) correspondente à função da partição intrínseca. Portanto, a partição baseada em *valores- p* fornece as mesmas expressões para as probabilidades de erros condicionais dadas na partição intrínseca.

Daí pode-se afirmar que as probabilidades de erros condicionais serão dadas por:

1) Sob a rejeição de H_0 ,

$$\alpha_z = \frac{f_T^0(H(t))H'(t)}{f_T^0(H(t))H'(t) + f_T^0(t)} = \frac{\frac{f_T^0(H(t))}{f_T^1(H(t))}}{\frac{f_T^0(H(t))}{f_T^1(H(t))} + \frac{f_T^0(t)}{f_T^1(H(t))H'(t)}} \Rightarrow$$

$$\alpha_z = \frac{B(H(t))}{B(H(t)) + 1} \Rightarrow \boxed{\alpha_{t'} = \frac{B(t')}{1 + B(t')}}.$$

2) Sob rejeição de H_1 ,

$$\beta_z = \frac{f_T^1(t)}{f_T^1(H(t))H'(t) + f_T^1(t)} = \frac{\frac{f_T^1(t)}{f_T^1(t)}}{\frac{f_T^0(t)}{f_T^1(t)} + \frac{f_T^1(t)}{f_T^1(t)}} \Rightarrow \boxed{\beta_t = \frac{1}{B(t) + 1}}.$$

Note que os *valores- p* não são usados diretamente como medidas de evidência. A proposta de utilização dos *valores- p* na construção da estatística condicionante fornece um elo entre os testes de significância e os testes freqüentistas condicionais.

4.2 COMPARAÇÃO ENTRE ESTATÍSTICAS CONDICIONANTES

4.2.1 Teste de Hipóteses com Razão de Verossimilhanças Simétrica

1.1 – Considere o teste de hipóteses $H_0: X \sim \text{Beta}(2,3)$ versus $H_1: X \sim \text{Beta}(3,2)$ (ver exemplo 2, página 16) . A razão de verossimilhanças de

H_0 para H_1 é igual a $B(x) = \frac{1-x}{x}$ para $n = 1$.

A distribuição de $B(X)$ sob a hipótese nula é dada por:

$$\begin{aligned} f_B^0(b) &= \left| \frac{d}{db} g^{-1}(b) \right| f_Y(g^{-1}(b)) = \frac{1}{(1+b)^2} \frac{[g^{-1}(b)]^1 [1-g^{-1}(b)]^2}{B(2,3)} = \\ &= \frac{1}{(1+b)^2} \frac{4!(1+b)^{-1} [1-(1+b)^{-1}]^2}{1!2!} = \frac{12}{(1+b)^3} \frac{b^2}{(1+b)^2} \Rightarrow \boxed{f_B^0(b) = 12b^2(1+b)^{-5}} \end{aligned}$$

E sob a hipótese alternativa, é:

$$\begin{aligned} f_B^1(b) &= \left| \frac{d}{db} g^{-1}(b) \right| f_Y(g^{-1}(b)) = \frac{1}{(1+b)^2} \frac{[g^{-1}(b)]^2 [1-g^{-1}(b)]^1}{B(3,2)} = \\ &= \frac{1}{(1+b)^2} \frac{4!(1+b)^{-2} [1-(1+b)^{-1}]^1}{1!2!} = \frac{12}{(1+b)^4} \frac{b}{(1+b)} \Rightarrow \boxed{f_B^1(b) = 12b(1+b)^{-5}} \end{aligned}$$

Como

$$\begin{aligned} f_{1/B}^1(b) &= \left| \frac{d}{db} g^{-1}(b) \right| f_Y(g^{-1}(b)) = \frac{1}{b^2} 12 \left(\frac{1}{b} \right) \left(1 + \left(\frac{1}{b} \right) \right)^{-5} = \frac{12}{b^3} \left(\frac{b+1}{b} \right)^{-5} \\ &\Rightarrow f_{1/B}^1(b) = 12b^2(b+1)^{-5} = f_B^0(b) \end{aligned}$$

Tem-se que a razão de verossimilhanças é simétrica.

1ª Estatística Condicionante: Estatística Ancilar

$$\text{Seja } Z = \min \left\{ \frac{1}{B(x)}, B(x) \right\} = \min \left\{ \left(\frac{1-x}{x} \right)^{-1}, \left(\frac{1-x}{x} \right) \right\}. \text{ (Figura 2.16)}$$

Z é ancilar (página 98 item 4.1.3).

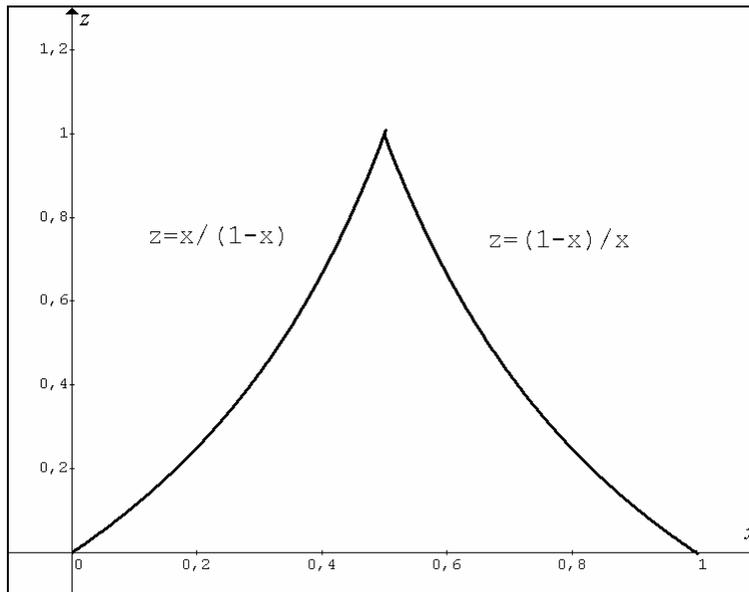


FIGURA 2.16 – Estatística Condicionante Ancilar associada a $B(x) = \frac{1-x}{x}$.

$$\text{A intersecção ocorre em } \left(\frac{1-x}{x} \right)^{-1} = \left(\frac{1-x}{x} \right) \Rightarrow \boxed{x = \frac{1}{2}}.$$

Observe que, para valores menores que $1/2$, a estatística condicionante será dada

$$\text{por } z = \frac{x}{1-x}, \text{ e para valores maiores, por } z = \frac{1-x}{x}.$$

Então, dado um valor da estatística $Z = z$, há dois possíveis valores de x , $x = (1+z)^{-1}$ e $x' = 1 - (1+z)^{-1} = 1 - x$. Tem-se, então, a função particionante $H(x)$ dada por $\boxed{H(x)=1-x}$. Utilizando estes pontos x_1 e x_2 determinados, pode-se fazer a verificação de Z como uma estatística ancilar, pois:

➤ Quando H_0 é verdadeira, ou seja, admite uma distribuição Beta (2,3), a distribuição acumulada é dada por:

$$\begin{aligned}
 F_Z^0(y_0) &= P[Z \leq y_0] = \int_{x_1}^{x_2} \frac{(x)^{2-1} (1-x)^{3-1}}{B(2,3)} dx = \frac{1}{B(2,3)} \int_{\left(\frac{y_0}{1+y_0}\right)}^{\left(\frac{1}{1+y_0}\right)} (x - 2x^2 + x^3) dx = \\
 &= \frac{1}{B(2,3)} \left[\frac{x^2}{2} - \frac{2x^3}{3} + \frac{x^4}{4} \right]_{\left(\frac{y_0}{1+y_0}\right)}^{\left(\frac{1}{1+y_0}\right)} = \\
 &= 12 \left[\left(\frac{1}{1+y_0}\right)^2 \frac{1}{2} - \left(\frac{1}{1+y_0}\right)^3 \frac{2}{3} + \left(\frac{1}{1+y_0}\right)^4 \frac{1}{4} - \left(\frac{y_0}{1+y_0}\right)^2 \frac{1}{2} + \left(\frac{y_0}{1+y_0}\right)^3 \frac{2}{3} - \left(\frac{y_0}{1+y_0}\right)^4 \frac{1}{4} \right] \\
 &= 12 \left[\frac{1}{2} \left[\left(\frac{1}{1+y_0}\right)^2 - \left(\frac{y_0}{1+y_0}\right)^2 \right] - \frac{2}{3} \left[\left(\frac{1}{1+y_0}\right)^3 - \left(\frac{y_0}{1+y_0}\right)^3 \right] + \frac{1}{4} \left[\left(\frac{1}{1+y_0}\right)^4 - \left(\frac{y_0}{1+y_0}\right)^4 \right] \right] \\
 F_Z^0(y_0) &= 12 \left[\frac{1}{2} \left[\frac{1-y_0^2}{(1+y_0)^2} \right] - \frac{2}{3} \left[\frac{1-y_0^3}{(1+y_0)^3} \right] + \frac{1}{4} \left[\frac{1-y_0^4}{(1+y_0)^4} \right] \right] = \frac{y_0^4 + 4y_0^3 - 4y_0 - 1}{(1+y_0)^4}.
 \end{aligned}$$

➤ Quando H_1 é verdadeira, ou seja, admite uma distribuição Beta (3,2), a distribuição acumulada é dada por:

$$F_Z^1(y_0) = P[Z \leq y_0] = \int_{x_1}^{x_2} \frac{(x)^{3-1} (1-x)^{2-1}}{B(3,2)} dx = \frac{1}{B(3,2)} \int_{\left(\frac{y_0}{1+y_0}\right)}^{\left(\frac{1}{1+y_0}\right)} (x^2 - x^3) dx =$$

$$12 \left[\frac{x^3}{3} - \frac{x^4}{4} \left(\frac{1}{1+y_0} \right) \right]_{\left(\frac{y_0}{1+y_0} \right)} = 12 \left[\left(\frac{1}{1+y_0} \right)^3 \frac{1}{3} - \left(\frac{1}{1+y_0} \right)^4 \frac{1}{4} - \left(\frac{y_0}{1+y_0} \right)^3 \frac{1}{3} + \left(\frac{y_0}{1+y_0} \right)^4 \frac{1}{4} \right] =$$

$$= 12 \left[\frac{1-y_0^3}{3(1+y_0)^3} - \frac{1-y_0^4}{4(1+y_0)^4} \right] = \frac{y_0^4 + 4y_0^3 - 4y_0 - 1}{(1+y_0)^4}.$$

De onde segue que $F_Z^0(y_0) = F_Z^1(y_0)$ e, portanto, Z é ancilar.

Obs.: Sugestão de Escolha da Estatística Ancilar

Para escolha de uma estatística ancilar para o teste, é possível utilizar a própria condição $F_Z^0(y_0) = F_Z^1(y_0)$. Podem ser utilizados os limites x_1 e x_2 na

integral $F_Z^i(y_0) = P[Z \leq y_0] = \int_{x_1}^{x_2} f_i(x) dx$, de modo que se tenha uma relação

entre x_1 e x_2 que satisfaça a condição anterior. Neste exemplo, tem-se:

$$\frac{1}{B(2,3)} \left[\frac{x^2}{2} - \frac{2x^3}{3} + \frac{x^4}{4} \right]_{(x_1)}^{(x_2)} = \frac{1}{B(3,2)} \left[\frac{x^3}{3} - \frac{x^4}{4} \right]_{(x_1)}^{(x_2)}$$

Daí chega-se à equação $x_1^2(1-x_1)^2 = x_2^2(1-x_2)^2$, que tem como soluções:

$$x_2 = x_1 \quad \text{ou} \quad x_2 = 1 - x_1 \quad \text{ou} \quad x_2 = \frac{1}{2} \pm \frac{1}{2} \sqrt{1 - 4x_1^2 + 4x_1}.$$

Note que a segunda solução corresponde à estatística ancilar utilizada no exemplo. ■

Calculando agora os erros condicionais, tem-se:

- Considerando primeiramente a distribuição Beta (2,3), ou seja, H_0 verdadeira, calcula-se a probabilidade de erro condicional tipo I dado pelas áreas determinadas pelos intervalos $[x, x + \Delta x]$ e $[H(x), H(x+\Delta x)]$ na equação (2.5) (Figura 2.17):

$$\alpha_x = \frac{f_0(x)}{f_0(x) + f_0(H(x))|H'(x)|} = \frac{f_0(x)}{f_0(x) + f_0(1-x)} =$$

$$= \frac{\frac{x^{2-1}(1-x)^{3-1}}{B(2,3)}}{\frac{x^{2-1}(1-x)^{3-1}}{B(2,3)} + \frac{(1-x)^{2-1}x^{3-1}}{B(2,3)}} = \frac{(1-x)^2 x}{(1-x)^2 x + x^2(1-x)} \Rightarrow \boxed{\alpha_x = 1-x}$$

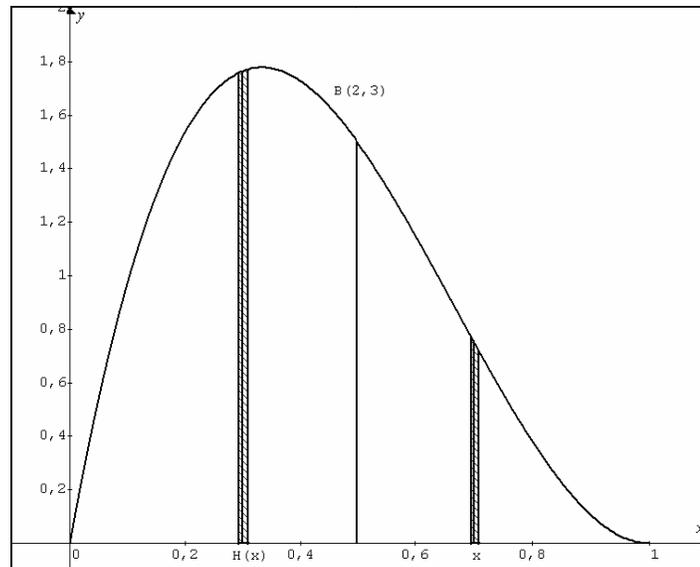


FIGURA 2.17 - Áreas correspondentes aos pontos x e $H(x) = 1-x$ para cálculo da probabilidade de erro condicional tipo I.

- Com a distribuição Beta (3,2) como verdadeira, calcula-se a probabilidade de erro condicional tipo II também utilizando a equação (2.5) (Figura 2.18):

$$\beta_x = \frac{f_1(x)}{f_1(x) + f_1(H(x))|H'(x)|} = \frac{f_1(x)}{f_1(x) + f_1(1-x)} =$$

$$\frac{\frac{x^{3-1}(1-x)^{2-1}}{B(3,2)}}{\frac{x^{3-1}(1-x)^{2-1}}{B(3,2)} + \frac{(1-x)^{3-1}(1-(1-x))^{2-1}}{B(3,2)}} = \frac{x^2(1-x)}{x^2(1-x) + (1-x)^2 x} \Rightarrow \boxed{\beta_x = x}$$

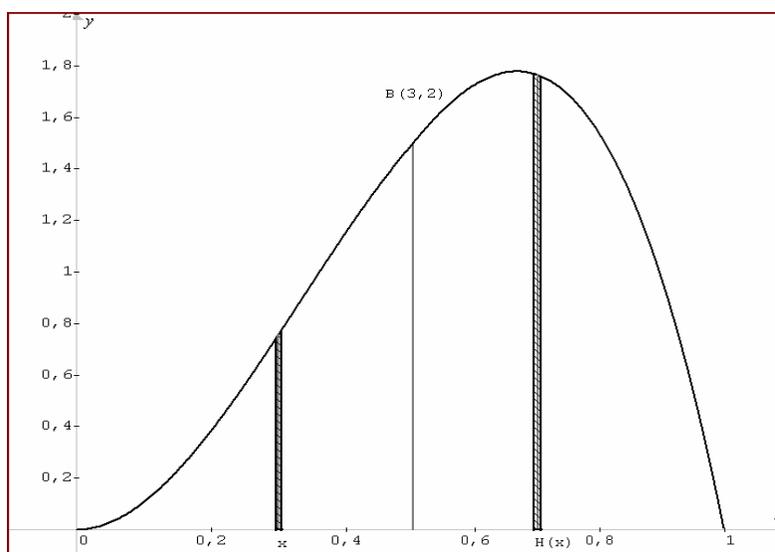


FIGURA 2.18 - Áreas correspondentes aos pontos x e $H(x) = 1-x$ para cálculo da probabilidade de erro condicional tipo II.

O teste com taxas de erros freqüentistas condicionais fica, então, da forma:

$$Y = \begin{cases} \text{se } x \geq \frac{1}{2}, \text{ rejeita-se } H_0 \text{ e apresenta-se PEC I } \alpha_x = 1 - x \\ \text{se } x < \frac{1}{2}, \text{ aceita-se } H_0 \text{ e apresenta-se PEC II } \beta_x = x \end{cases}$$

2ª Estatística Condicionante: Significância Intrínseca

A estatística condicionante utilizada por Birnbaum (1961, citada por Kiefer, 1977) é definida por $Z = \min\{B(X), B(X)^{-1}\}$. Nesse caso, tem-se $Z = \min\left\{\frac{1-x}{x}, \frac{x}{1-x}\right\}$, e, portanto, a abordagem de Birnbaum é idêntica à anterior.

Exemplo numérico:

Suponha que o valor da estatística condicionante Z obtido seja $z = \frac{3}{7} \cong 0,429$. Logo, $B(x) = 0,429$ ou $1/B(x) = 0,429$, que equivale a $B(x) = 2,33$. Como $B(x) = \frac{1-x}{x}$, tem-se que, para $B(x) = 0,429 \Rightarrow x = 0,7$ com PEC I $\alpha_z = 0,3$, e para $B(x) = 2,33 \Rightarrow x = 0,3$ com PEC II $\beta_z = 0,3$.

3ª Estatística Condicionante: Partição Contínua com erros condicionais iguais.

Kiefer (1977) sugere a escolha da estatística condicionante Z tal que $\alpha_z = \beta_z$. Com isso, a partição será dada por:

$$\int_{-\infty}^{a^*} \sqrt{f_0(x) f_1(x)} dx = \int_{a^*}^{+\infty} \sqrt{f_0(x) f_1(x)} dx \Rightarrow$$

$$\Rightarrow \int_0^{a^*} \sqrt{\frac{x(1-x)^2}{B(2,3)} \cdot \frac{x^2(1-x)}{B(3,2)}} dx = \int_{a^*}^1 \sqrt{\frac{x^3(1-x)^3}{B(2,3)B(3,2)}} dx$$

$$\frac{1}{B(2,3)} \int_0^{a^*} [x(1-x)]^{\frac{3}{2}} dx = \frac{1}{B(3,2)} \int_{a^*}^1 [x-x^2]^{\frac{3}{2}} dx \Rightarrow \int_0^{a^*} [x-x^2]^{\frac{3}{2}} dx = \int_{a^*}^1 [x-x^2]^{\frac{3}{2}} dx$$

Geometricamente, como a função $g(x) = [x(1-x)]^{3/2}$ é simétrica em relação a 0,5, então $\int_0^1 [x-x^2]^{3/2} dx = 2 \int_0^{0,5} [x-x^2]^{3/2} dx$; daí tem-se que $a^* = 0,5$ (Figura 2.19).

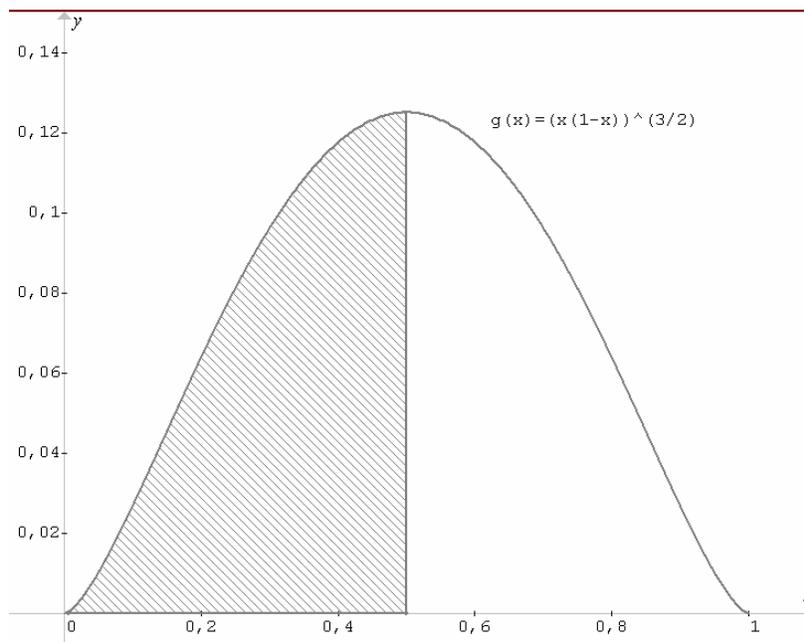


FIGURA 2.19 – Simetria da função $g(x) = [x(1-x)]^{3/2}$ em relação à reta $x = 0,5$.

A função particionante $H(x)$ é dada pela solução da equação:

$$\int_{-\infty}^{H(x)} \sqrt{f_0(y)f_1(y)} dy = \int_x^{+\infty} \sqrt{f_0(y)f_1(y)} dy \Rightarrow \int_0^{H(x)} [y(1-y)]^{3/2} dy = \int_x^1 [y-y^2]^{3/2} dy$$

Novamente pela da simetria, pode-se concluir que $H(x) = (1-x)$.

Como a função $H(x)$ e o valor crítico são os mesmos do problema anterior, o teste neste caso é igual aos casos anteriores.

4º Estatística Condicionante: valores-p

Considere a estatística condicionante $Z = \max \{p_0, p_1\}$, sendo:

$$p_0 = P_{\theta_0} [X \geq x] = 1 - F_0(x) = \int_0^x \frac{(y)^{2-1} (1-y)^{3-1}}{B(2,3)} dy = 12 \left[\frac{y^2}{2} - \frac{2y^3}{3} + \frac{y^4}{4} \right]_0^x \Rightarrow$$

$$\boxed{p_0 = 1 - 6x^2 + 8x^3 - 3x^4} \text{ e}$$

$$p_1 = P_{\theta_1} [X \leq x] = F_1(x) = 12 \int_0^x (y^2 - y^3) dy = 12 \left[\frac{y^3}{3} - \frac{y^4}{4} \right]_0^x \Rightarrow \boxed{p_1 = 4x^3 - 3x^4}.$$

Como $Z = \max \{1 - 6x^2 + 8x^3 - 3x^4, 4x^3 - 3x^4\}$, o valor crítico que define a fronteira da região crítica é:

$$1 - 6a^{*2} + 8a^{*3} - 3a^{*4} = 4a^{*3} - 3a^{*4} \Rightarrow a^* = 0,5 \quad (\text{Figura 2.20}), \text{ ou}$$

mesmo $\int_{-\infty}^{a^*} f_1(H(x))H'(x)dx = \int_{a^*}^{+\infty} f_0(x)dx$, que neste exemplo se resume a:

$$\int_0^{a^*} \frac{u^2(1-u)}{B(3,2)} du = \int_{a^*}^1 \frac{x(1-x)^2}{B(2,3)} dx \Rightarrow 4a^{*3} - 6a^{*2} + 1 = 0 \Rightarrow \boxed{a^* = 0,5}.$$

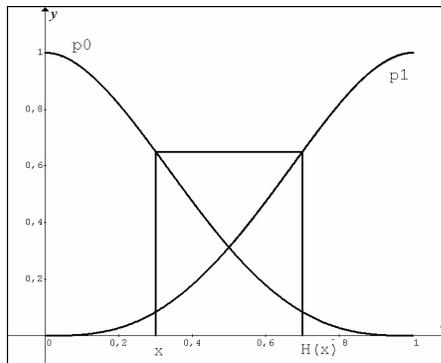


FIGURA 2.20 - Estatística Condicionante *valor-p* dada por

$$Z = \max \{1 - 6x^2 + 8x^3 - 3x^4, 4x^3 - 3x^4\}.$$

A função $H(x)$ que define a partição é dada pela solução da equação:

$$\int_{-\infty}^{H(x)} f_1(H(x))H'(x)dx = \int_x^{+\infty} f_0(x)dx \quad .$$

Nesse exemplo têm-se os limites de integração entre 0 e 1, que fornece:

$$\int_0^{H(x)} f_1(H(y))H'(y)dy = \int_x^1 f_0(y)dy \Rightarrow 4y^3 - 3y^4 \Big|_0^{H(x)} = 6u^2 - 8u^3 + 3u^4 \Big|_x \Rightarrow$$

$$\boxed{1 - 6x^2 + 8x^3 - 3x^4 = 4H(x)^3 - 3H(x)^4}$$

,que tem uma solução dada por $H(x) = 1 - x$, como nos casos anteriores, fornecendo o mesmo teste. Assim, é declarado, por exemplo, que $x = 0,3$ na região de aceitação tem tanta força de evidência quanto $x = 0,7$ na região de rejeição, já que eles levariam ao mesmo *valor-p*.

Para o teste de hipóteses clássico, o valor crítico 0,5 corresponderia a probabilidades de erro dadas por:

$$\alpha = P_{\theta_0} [\text{rejeitar } H_0] = P_{\theta_0} [x \geq 0,5] = 1 - \int_0^{0,5} \frac{x(1-x)^2}{B(2,3)} dx = 1 - 12 \left[\frac{x^2}{2} - \frac{2x^3}{3} + \frac{x^4}{4} \Big|_0^{0,5} \right] =$$

$$= 1 - 6(0,5)^2 + 8(0,5)^3 - 3(0,5)^4 = 0,3125$$

$$\beta = P_{\theta_1} [\text{aceitar } H_0] = P_{\theta_1} [x < 0,5] = \int_0^{0,5} \frac{x^2(1-x)}{B(3,2)} dx = 12 \left[\frac{x^2}{2} - \frac{2x^3}{3} + \frac{x^4}{4} \Big|_0^{0,5} \right] =$$

$$= 12 \left[\frac{x^3}{3} - \frac{x^4}{4} \Big|_0^{0,5} \right] = 4(0,5)^3 - 3(0,5)^4 = 0,3125$$

Observação:

Os testes de hipóteses referente ao parâmetro média μ da população de uma distribuição normal, cujo desvio padrão σ possui o mesmo valor sob as hipóteses H_0 e H_1 ($\sigma_0 = \sigma_1 = \sigma$), são testes com razão de verossimilhanças simétrica. Esse fato pode ser verificado a partir da simetria das distribuições sob $H_0 : \mu = \mu_0$ e sob $H_1 : \mu = \mu_1$. Portanto, as estatísticas condicionantes estudadas

nesse trabalho são todas adequadas para esse tipo de teste, que foi exemplificado no item 4.2.1.

4.2.2 Teste de Hipóteses com Razão de Verossimilhanças Não-Simétrica

2.1. Deseja-se testar as hipóteses simples $H_0: X \sim \text{Uniforme}(0,1)$ versus $H_1: X \sim \text{Beta}(1/2,1)$ (Figura 2.21).

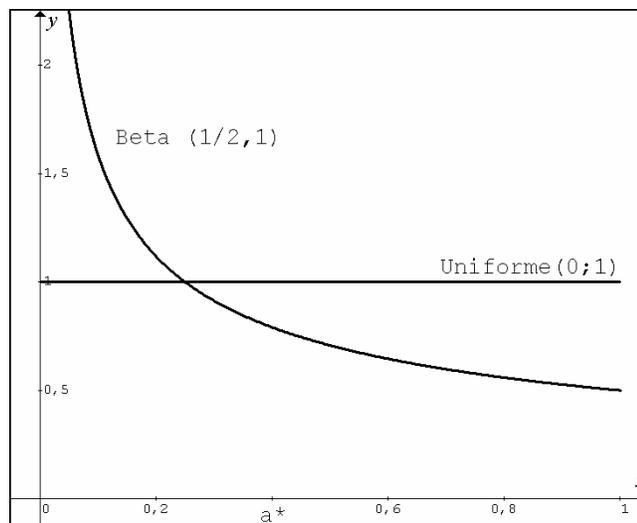


FIGURA 2.21 – Funções Densidade de Probabilidades das Distribuições Uniforme (0,1) e Beta (1/2, 1).

A razão de verossimilhanças de H_0 para H_1 é: $B(x) = \frac{1}{(2\sqrt{x})^{-1}} = 2\sqrt{x}$

sendo não simétrico, pois a distribuição de $B(x)$ será dada por:

$$f_B^0(b) = \left| \frac{d}{db} g^{-1}(b) \right| f_x(g^{-1}(b)) = \frac{b}{2} f_x\left(\frac{b^2}{4}\right) = \frac{b}{2} \quad e$$

$$f_B^1(b) = \left| \frac{d}{db} g^{-1}(b) \right| f_X(g^{-1}(b)) = \frac{b}{2} f_X\left(\frac{b^2}{4}\right) = \frac{b}{2} \frac{1}{2\sqrt{\left(\frac{b^2}{4}\right)}} = \frac{b}{2} \frac{1}{\left(\frac{2b}{2}\right)} = \frac{1}{2}$$

Então:

$$F_B^0(y) = P[b \leq y] = \int_0^y \frac{x}{2} dx = \frac{y^2}{4} \quad \text{e} \quad F_B^1(y) = P[b \leq y] = \int_0^y \frac{1}{2} dx = \frac{y}{2}.$$

Portanto,

$$F_{1/B}^1(y) = P[1/b \leq y] = P\left[b \geq \frac{1}{y}\right] = 1 - \int_0^{1/y} \frac{1}{2} dx = 1 - \frac{2}{y} = \frac{y-2}{y},$$

ou seja, $F_B^0(y) \neq F_{1/B}^1(y)$.

1ª Estatística Condicionante: Estatística Ancilar

Para obter uma estatística condicionante que seja ancilar, neste exemplo deve-se ter:

$$F_Z^0(y_0) = F_Z^1(y_0) \Rightarrow \int_{x_1}^{x_2} 1 dx = \int_{x_1}^{x_2} \frac{1}{2} x^{-\frac{1}{2}} dx \Rightarrow x \Big|_{x_1}^{x_2} = x^{\frac{1}{2}} \Big|_{x_1}^{x_2}; \text{ então, a relação}$$

entre x_1 e x_2 deve ser $\boxed{x_2 - x_1 = \sqrt{x_2} - \sqrt{x_1}}$.

Uma possibilidade de estatística condicionante Ancilar Z é $Z = \max\{B(X), 2 - B(X)\}$, que se reduz a $Z = \max\{2\sqrt{x}, 2 - 2\sqrt{x}\}$. O valor mínimo de Z ocorre em

$$2\sqrt{x} = 2 - 2\sqrt{x} \Rightarrow 4\sqrt{x} = 2 \Rightarrow \sqrt{x} = \frac{1}{2} \Rightarrow \boxed{x = 0,25} \text{ (Figura 2.22).}$$

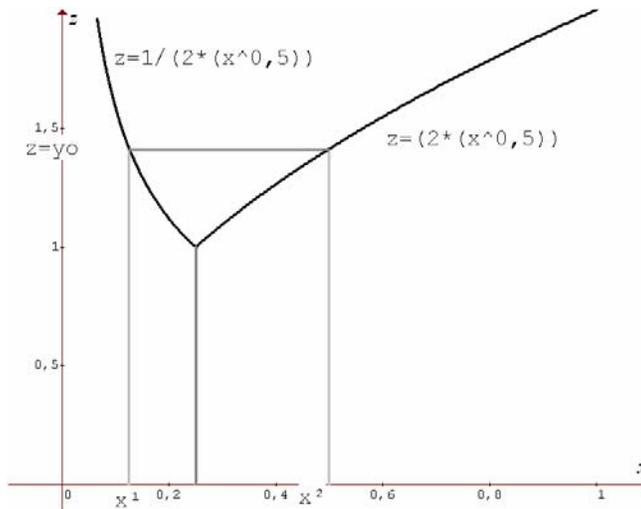


FIGURA 2.22 – Estatística Condicionante $Z = \max\{2\sqrt{x}, 2 - 2\sqrt{x}\}$.

A partir do ponto mínimo de Z pode-se definir as regiões de decisão em $x \leq 0,25$ ou $x > 0,25$. Como $0 \leq X \leq 1$ e, conseqüentemente, $1 \leq Z \leq 2$, escolhendo-se um valor tal que $z \leq y_0 \Rightarrow x_1 \leq x \leq x_2$. Então, na primeira função determina-se x_1 , sendo:

$$y_0 = 2 - 2\sqrt{x_1} \Rightarrow 2\sqrt{x_1} = 2 - y_0 \Rightarrow \sqrt{x_1} = \frac{2 - y_0}{2} \Rightarrow x_1 = \left(\frac{2 - y_0}{2}\right)^2.$$

E para a segunda função, encontra-se o ponto x_2 definido por:

$$y_0 = 2\sqrt{x_2} \Rightarrow \sqrt{x_2} = \frac{y_0}{2} \Rightarrow x_2 = \left(\frac{y_0}{2}\right)^2 \Rightarrow x_2 = \frac{y_0^2}{4}.$$

Afirmção: A estatística Z é ancilar.

Utilizando os pontos x_1 e x_2 determinados, pode-se fazer a verificação que Z é uma estatística ancilar, pois:

- Quando H_0 é verdadeira, ou seja, admite uma distribuição Uniforme $(0,1)$, a distribuição acumulada é dada por:

$$F_Z^0(y_0) = P[Z \leq y_0] = \int_{x_1}^{x_2} 1 dx = \int_{\left(\frac{2-y_0}{2}\right)^2}^{\frac{y_0^2}{4}} 1 dx = x \Big|_{\left(\frac{2-y_0}{2}\right)^2}^{\frac{y_0^2}{4}} =$$

$$= \frac{y_0^2}{4} - \frac{(2-y_0)^2}{4} = \frac{y_0^2 - 4 + 4y_0 - y_0^2}{4} = y_0 - 1$$

- Quando H_1 é verdadeira, ou seja, admite uma distribuição Beta $(1/2,1)$, a distribuição acumulada é dada por:

$$F_Z^1(y_0) = \int_{\left(\frac{2-y_0}{2}\right)^2}^{\frac{y_0^2}{4}} \frac{1}{2} x^{-\frac{1}{2}} dx = x^{\frac{1}{2}} \Big|_{\left(\frac{2-y_0}{2}\right)^2}^{\frac{y_0^2}{4}} = \frac{y_0}{2} - \frac{2-y_0}{2} = y_0 - 1.$$

Dessa forma, a distribuição da estatística Z independe do parâmetro assumido por cada uma das hipóteses inicialmente, sendo uma estatística ancilar. Note que a estatística ancilar Z tem distribuição uniforme (Figura 2.23).

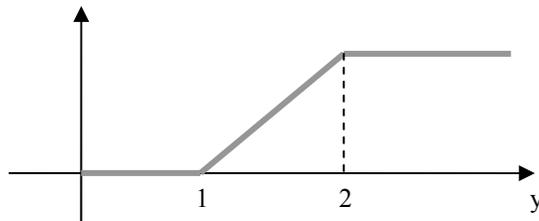


FIGURA 2.23 – Distribuição Acumulada da estatística ancilar Z ,
 $F_Z(z) = z - 1$.

■

Então, dado um valor da estatística $Z = z$, têm-se dois possíveis valores de x : $x = \left(\frac{2-z}{2}\right)^2$ e $x = \frac{z^2}{4}$. Explicitando z , tem-se $z = 2\sqrt{x}$ na segunda expressão; com isso, fica definida a função $H(x)$ dada por

$$x = \left(\frac{2-z}{2}\right)^2 \Rightarrow 4x = 4 - 4z + z^2 \Rightarrow H(x) = \frac{4 - 4(2\sqrt{x}) + (2\sqrt{x})^2}{4}$$

$$\boxed{H(x) = 1 - 2\sqrt{x} + x}.$$

Tem-se, então, que $\frac{d}{dx}(1 - 2\sqrt{x} + x) = \boxed{H'(x) = 1 - \frac{1}{\sqrt{x}}}$.

Utilizando $H(x)$ para construção da partição, têm-se as probabilidades de erros condicionais calculadas de acordo com as expressões (2.5) e (2.6). Estas são rerepresentadas aqui para comodidade do leitor:

$$\alpha_x = \frac{f_0(H(x))|H'(x)|}{f_0(H(x))|H'(x)| + f_0(x)} \quad \text{e} \quad \beta_x = \frac{f_1(x)}{f_1(x) + f_1(H(x))|H'(x)|}.$$

➤ Para que o erro condicional tipo I ocorra, deve-se considerar a distribuição uniforme como verdadeira. A PEC I fica:

$$\alpha_x = \frac{f_0(H(x)) \cdot H'(x)}{f_0(H(x)) \cdot |H'(x)| + f_0(x)} = \frac{1}{1 - \left(1 - \frac{1}{\sqrt{x}}\right)} \Rightarrow \boxed{\alpha_x = \sqrt{x}}.$$

➤ Para o cálculo da probabilidade do erro condicional tipo II, deve-se considerar a distribuição Beta (1/2,1).

Como as áreas A_1 e A_0 são iguais, a probabilidade de erro condicional

tipo II fica sendo $\beta_x = \frac{f_1(x)}{f_1(H(x)) \cdot |H'(x)| + f_1(x)} = \frac{\frac{1}{2\sqrt{x}} \Delta x}{\frac{1}{2\sqrt{x}} \Delta x + \frac{1}{2\sqrt{x}} \Delta x} = \frac{1}{2}$.

O teste com taxas de erros condicionais baseado na estatística ancilar como condicionante fica, então, da forma:

$$Y^A = \begin{cases} \text{se } x \leq \frac{1}{4}, \text{ rejeita-se } H_0 \text{ e apresenta-se PEC I } \alpha_x = \sqrt{x}; \\ \text{se } x > \frac{1}{4}, \text{ aceita-se } H_0 \text{ e apresenta-se PEC II } \beta_x = \frac{1}{2}. \end{cases}$$

É importante observar que a PEC II $\beta_x = \frac{1}{2}$ é inadequada, pois não atende a condição essencial dos testes abordados neste trabalho, que é de variar a probabilidade de erro de acordo com os dados obtidos. Note que se $x > \frac{1}{4}$, $B(x)$ varia de 1 a 2.

Não é possível prever com que frequência este tipo de situação ocorre, pois há várias possibilidades de estatísticas ancilares possíveis para o condicionamento, sendo, então, esta desaconselhável para a presente finalidade quando o problema se tratar de razão de verossimilhanças não-simétrica.

2ª Estatística Condicionante: Significância Intrínseca

A abordagem de Birnbaum (1961) é equivalente a escolher a estatística condicionante $Z = \max \left\{ B(x), \frac{1}{B(x)} \right\} \Rightarrow Z = \max \left\{ 2\sqrt{x}, \frac{1}{2\sqrt{x}} \right\}$.

Na Figura 2.24 pode-se observar o comportamento da estatística condicionante Z baseada nos níveis de significância intrínseca para este caso.

Tem-se que $z = \frac{1}{2\sqrt{x}}$ e $z = 2\sqrt{x}$; então, $x_1 = \frac{1}{4z^2}$ e $x_2 = \frac{z^2}{4}$. A

função particionante é dada por $H(x) \Rightarrow x_1 = \frac{1}{4(4x_2)} = \frac{1}{16x_2}$. O valor mínimo

de Z ocorre para $x = 0,25$. Como $0 < x \leq 1$, então $B(x) \leq 2$ e, portanto, observe que para valores menores que $\frac{1}{16}$, a estatística condicionante vai para infinito.

Assim, para valores de $Z > 2$, apenas um valor de x dado $(2\sqrt{x})^{-1} = z$; então, no intervalo de $[0, \frac{1}{16}]$ não há condicionamento em $B(x)$, pois o valor deve ser limitado a $B = 2$, portanto assumindo-se, para esse intervalo, que $PEC I = 1$.

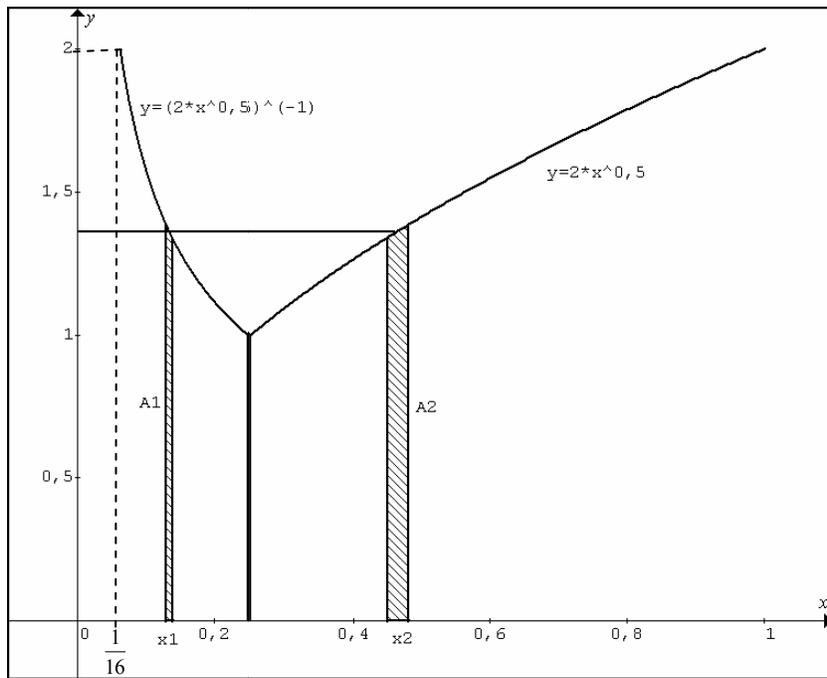


FIGURA 2.24 – Comportamento da estatística condicionante baseada nos níveis de significância intrínseca $Z = \max \left\{ 2\sqrt{x}, \frac{1}{2\sqrt{x}} \right\}$ em função dos dados.

Têm-se as PEC's dadas por:

Sob H_0 :

$$\alpha_x = \frac{f_0(H(x)) \cdot H'(x)}{f_0(H(x)) \cdot |H'(x)| + f_0(x)} = \frac{1}{1 + H'(x)} = \frac{1}{1 + \frac{1}{16x^2}} = \left(1 + (16x^2)^{-1}\right)^{-1}$$

$$\beta_x = \frac{f_1(H(x)) \cdot H'(x)}{f_1(H(x)) \cdot |H'(x)| + f_1(x)} = \frac{\frac{\sqrt{x}}{8x^2}}{\frac{1}{2\sqrt{x}} + \frac{\sqrt{x}}{8x^2}} = \frac{x}{x + 4x^2} \Rightarrow \beta_x = (1 + 4x)^{-1}.$$

O teste fica da forma:

$$Y' = \begin{cases} \text{se } x \leq \frac{1}{4}, \text{ rejeita-se } H_0 \text{ e apresenta-se PEC I } \alpha_x = \begin{cases} 1 & \text{se } x < \frac{1}{16} \\ \left(1 + \frac{1}{16x^2}\right)^{-1} & \text{se } x > \frac{1}{16}; \end{cases} \\ \text{se } x > \frac{1}{4}, \text{ aceita-se } H_0 \text{ e apresenta-se PEC II } \beta_x = (1 + 4x)^{-1}. \end{cases}$$

Porém, apresentar $\alpha_x = 1$ quando $x < \frac{1}{16}$ é totalmente contraditório ao objetivo principal dos testes freqüentistas condicionais, sendo a estatística intrínseca também inapropriada para este exemplo, pois a PEC I no intervalo $\left[0, \frac{1}{16}\right]$ não depende dos dados observados, lembrando que as PEC's possuem valores elevados mesmo distantes da fronteira da região crítica, pois se trata de amostra de tamanho $n = 1$, contribuindo para alta probabilidade de erro na decisão.

3º Estatística Condicionante: Partição com probabilidade de erros condicionais iguais.

Kiefer (1977) sugere a escolha da estatística condicionante Z tal que $\alpha_z = \beta_z$. Com isso, a função $H(x)$ que define a partição é dada pela solução da equação:

$$\int_{-\infty}^{H(x)} \sqrt{f_0(x) f_1(x)} dx = \int_x^{+\infty} \sqrt{f_0(x) f_1(x)} dx$$

Neste exemplo os limites de integração estão entre 0 e 1:

$$\begin{aligned} \int_0^{H(x)} \sqrt{f_0(x) f_1(x)} dx &= \int_x^1 \sqrt{f_0(x) f_1(x)} dx \Rightarrow \sqrt{\frac{1}{2}} \int_0^{H(x)} x^{-\frac{1}{4}} dx = \sqrt{\frac{1}{2}} \int_x^1 x^{-\frac{1}{4}} dx \Rightarrow \\ \sqrt[4]{H(x)^3} &= 1 - \sqrt[4]{x^3} \Rightarrow \boxed{H(x) = \sqrt[3]{\left(1 - x^{\frac{3}{4}}\right)^4}} \end{aligned}$$

E para o valor crítico do teste, tem-se:

$$\begin{aligned} \int_{-\infty}^{a^*} \sqrt{f_0(x) f_1(x)} dx &= \int_{a^*}^{+\infty} \sqrt{f_0(x) f_1(x)} dx \Rightarrow \int_0^{a^*} \sqrt{\frac{1}{2\sqrt{x}}} dx = \int_{a^*}^1 \sqrt{\frac{1}{2\sqrt{x}}} dx \Rightarrow \\ \sqrt{\frac{1}{2}} \int_0^{a^*} x^{-\frac{1}{4}} dx &= \sqrt{\frac{1}{2}} \int_{a^*}^1 x^{-\frac{1}{4}} dx \Rightarrow \frac{x^{\frac{3}{4}}}{\frac{3}{4}} \Big|_0^{a^*} = \frac{x^{\frac{3}{4}}}{\frac{3}{4}} \Big|_{a^*}^1 \Rightarrow \sqrt[4]{a^{*3}} = 1 - \sqrt[4]{a^{*3}} \Rightarrow \\ 2\sqrt[4]{a^{*3}} &= 1 \Rightarrow a^{*3} = \left(\frac{1}{2}\right)^4 \Rightarrow \boxed{a^* = \sqrt[3]{\frac{1}{16}} = 0,397} \end{aligned}$$

Note que para $x = 0,397$, tem-se

$$H(0,397) = \sqrt[3]{\left(1 - 0,397^{\frac{3}{4}}\right)^4} = \sqrt[3]{(1 - 0,5)^4} = \sqrt[3]{0,06257} = 0,397,$$

ou seja, $a^* = 0,397$ (Figura 2.25).

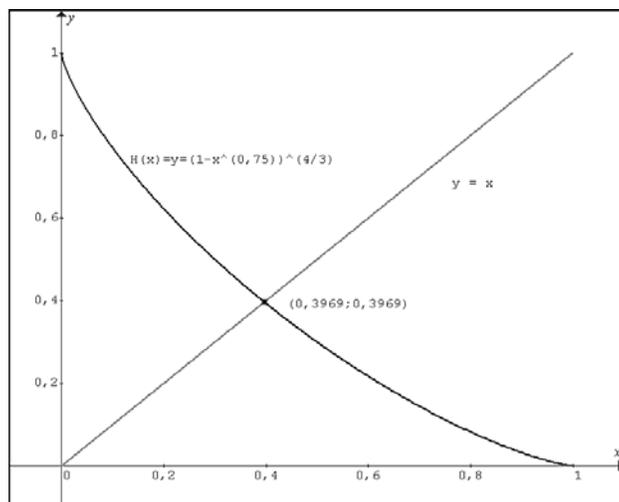


FIGURA 2.25 - Função Particionante H(x) para o critério com probabilidades de erros condicionais iguais.

As probabilidades dos erros condicionais são dadas por:

$$\alpha_x = \frac{f_1(H(x))H'(x)}{f_1(H(x))|H'(x)| + f_1(x)} = \frac{f_1\left(\left(1-x^4\right)^{\frac{4}{3}}\right)\frac{4}{3}\left(1-x^4\right)^{\frac{1}{3}}\frac{3}{4}x^{-\frac{1}{4}}}{f_1\left(\left(1-x^4\right)^{\frac{4}{3}}\right)\frac{4}{3}\left(1-x^4\right)^{\frac{1}{3}}\frac{3}{4}x^{-\frac{1}{4}} + f_1(x)} =$$

$$\frac{\frac{1}{2\sqrt{\left(1-x^4\right)^{\frac{4}{3}}}}\left(1-x^4\right)^{\frac{1}{3}}x^{-\frac{1}{4}}}{\frac{1}{2\sqrt{\left(1-x^4\right)^{\frac{4}{3}}}}\left(1-x^4\right)^{\frac{1}{3}}x^{-\frac{1}{4}} + \frac{1}{2\sqrt{x}}} = \frac{\frac{\left(1-x^4\right)^{\frac{1}{3}}}{2\left(1-x^4\right)^{\frac{2}{3}}}x^{-\frac{1}{4}}}{\frac{\left(1-x^4\right)^{\frac{1}{3}}}{2\left(1-x^4\right)^{\frac{2}{3}}}x^{-\frac{1}{4}} + \frac{1}{2\sqrt{x}}} = \frac{\left(1-x^4\right)^{\frac{1}{3}}x^{-\frac{1}{4}}\sqrt{x}}{\left(1-x^4\right)^{\frac{1}{3}}x^{-\frac{1}{4}}\sqrt{x} + \left(1-x^4\right)^{\frac{2}{3}}}$$

$$\begin{aligned}
&= \frac{\left(1-x^4\right)^{\frac{1}{3}}}{\left(1-x^4\right)^{\frac{1}{3}}+x^{-\frac{1}{4}}\left(1-x^4\right)^{\frac{2}{3}}} \Rightarrow \alpha_x = \left(1+x^{-\frac{1}{4}}\left(1-x^4\right)^{\frac{1}{3}}\right)^{-1} = \left(1+\left(x^{-\frac{1}{12}}-x^{\frac{3}{4}-\frac{1}{12}}\right)^{\frac{1}{3}}\right)^{-1} \\
&= \left(1+\left(x^{-\frac{1}{12}}-x^{\frac{8}{12}}\right)^{\frac{1}{3}}\right)^{-1} = \left(1+\left(x^{-\frac{1}{12}-\frac{8}{12}}-1\right)^{\frac{1}{3}}\right)^{-1} \Rightarrow \alpha_x = \left(1+\left(x^{-\frac{3}{4}}-1\right)^{\frac{1}{3}}\right)^{-1} \\
\beta_x &= \frac{f_1(x)}{f_1(H(x))|H'(x)|+f_1(x)} = \frac{f_1(x)}{f_1\left(\left(1-x^4\right)^{\frac{4}{3}}\right)\frac{4}{3}\left(1-x^4\right)^{\frac{1}{3}}\frac{3}{4}x^{-\frac{1}{4}}+f_1(x)} = \\
&= \frac{\frac{1}{2\sqrt{x}}}{\frac{1}{2\sqrt{\left(1-x^4\right)^{\frac{4}{3}}}}\left(1-x^4\right)^{\frac{1}{3}}x^{-\frac{1}{4}}+\frac{1}{2\sqrt{x}}} = \frac{\frac{1}{2\sqrt{x}}}{\frac{\left(1-x^4\right)^{\frac{1}{3}}}{2\left(1-x^4\right)^{\frac{2}{3}}x^{-\frac{1}{4}}+\frac{1}{2\sqrt{x}}}} = \frac{\left(1-x^4\right)^{\frac{2}{3}}}{\left(1-x^4\right)^{\frac{1}{3}}x^{-\frac{1}{4}}\sqrt{x}+\left(1-x^4\right)^{\frac{2}{3}}} = \\
&= \frac{\left(1-x^4\right)^{\frac{1}{3}}}{x^{-\frac{1}{4}}\sqrt{x}+\left(1-x^4\right)^{\frac{1}{3}}} = \left(1+x^{\frac{1}{4}}\left(1-x^4\right)^{\frac{1}{3}}\right)^{-1} = \left(1+\left(x^{\frac{1}{12}}-x^{\frac{10}{12}}\right)^{\frac{1}{3}}\right)^{-1} = \left(1+\left(x^{-\frac{3}{4}}-1\right)^{\frac{1}{3}}\right)^{-1}
\end{aligned}$$

Para o critério com probabilidades de erros condicionais iguais o teste fica, então, dado por:

$$Y^B = \begin{cases} \text{se } x \leq 0,397, \text{ rejeite } H_0 \text{ e apresente PEC I } \alpha_x = \left(1 + \left(x^{\frac{3}{4}} - 1\right)^{\frac{1}{3}}\right)^{-1} \\ \text{se } x > 0,397, \text{ aceite } H_0 \text{ e apresente PEC II } \beta_x = \left(1 + \left(x^{\frac{3}{4}} - 1\right)^{\frac{1}{3}}\right)^{-1} \end{cases} .$$

4º Estatística Condicionante: valor-p

Considere a estatística condicionante

$$Z = \max\{p_0, p_1\} = \max\{F_0(x), 1 - F_1(x)\},$$

sendo $p_0 = P_0(X \leq x) = x$ e $p_1 = P_1(X \geq x) = 1 - \sqrt{x}$, tem-se

$Z = \max\{x, 1 - \sqrt{x}\}$. Assim, é declarado que, por exemplo, $x = \frac{3}{4}$ na região de

aceitação tem tanta força de evidência quanto $x = \frac{1}{16}$ na região de rejeição, já

que eles levam o mesmo *valor-p* no teste de H_0 e H_1 , respectivamente.

Como $Z = \max\{x, 1 - \sqrt{x}\}$, o valor crítico será dado por

$$x = 1 - \sqrt{x} \Rightarrow (1 - x)^2 = (\sqrt{x})^2 \Rightarrow 1 - 2x + x^2 = x \Rightarrow x^2 - 3x + 1 = 0 \Rightarrow \boxed{x = 0,382}.$$

Este valor crítico pode também ser obtido pela resolução da equação

integral:
$$\int_{-\infty}^{a^*} f_1(H(x)) |H'(x)| dx = \int_{a^*}^{+\infty} f_0(x) dx$$

que nesse exemplo se resume a:

$$\int_0^{a^*} f_1(H(x)) |H'(x)| dx = \int_{a^*}^1 f_0(x) dx \Rightarrow \int_0^{a^*} f_1(u) du = 1 - a^* \Rightarrow \int_0^{a^*} \frac{1}{2\sqrt{u}} du = 1 - a^*$$

$$\Rightarrow \sqrt{u} \Big|_0^{a^*} = 1 - a^* \Rightarrow \sqrt{a^*} = 1 - a^* \Rightarrow a^* = (1 - a^*)^2 \Rightarrow$$

$$a^* = 1 - 2a^* + (a^*)^2 \Rightarrow a^{*2} - 3a^* + 1 = 0 \Rightarrow \boxed{a^* = 0,382}$$

Então, $\int_0^{0,382} 1 dx = x \Big|_0^{0,382} = 0,382$ e $\int_{0,382}^1 \frac{1}{2\sqrt{x}} dx = \frac{1}{2} \frac{x^{1/2}}{1/2} \Big|_{0,382}^1 = 1 - \sqrt{0,382} = 0,382$

(Figura 2.26).

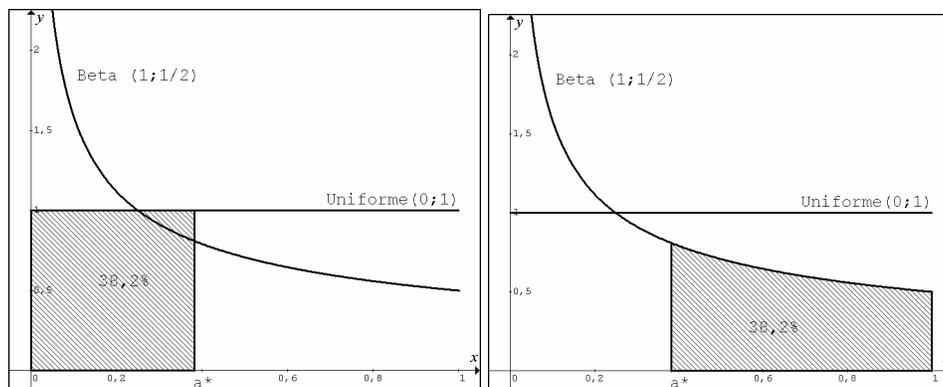


FIGURA 2.26 – Representação das áreas com condição do *valor-p* como condicionante.

Para encontrar o valor da função $H(x)$ que define a partição, basta solucionar a equação integral:

$$\int_{-\infty}^{H(x)} f_1(H(x)) |H'(x)| dx = \int_x^{+\infty} f_0(x) dx$$

Nesse exemplo, têm-se os limites de integração entre 0 e 1, que fornece:

$$\int_0^{H(x)} f_1(H(x)) |H'(x)| dx = \int_x^1 f_0(x) dx \Rightarrow \sqrt{u} \Big|_0^{H(x)} = 1 - x \Rightarrow$$

$$\sqrt{H(x)} = 1 - x \Rightarrow \boxed{H(x) = (1 - x)^2}$$

A função $H(x)$ também pode ser obtida diretamente da forma: dado um valor da estatística $Z = z$, têm-se dois possíveis valores de x : $x = z$ e ; com isso, fica definida a função $H(x)$ dada por $H(x) = (1-x)^2$.

A intersecção entre a reta $y = x$ e a função $H(x)$ ocorre no ponto 0,382, que é o valor associado ao critério de decisão.

As probabilidades de erros condicionais ficam da forma:

$$\alpha_x = \frac{f_1(H(x))H'(x)}{f_1(H(x))|H'(x)| + f_1(x)} = \frac{f_1((1-x)^2)(2(1-x))}{f_1((1-x)^2)(2(1-x)) + f_1(x)} =$$

$$\frac{\frac{1}{2\sqrt{(1-x)^2}}(2(1-x))}{\frac{1}{2\sqrt{(1-x)^2}}(2(1-x)) + f_1(x)} = \frac{\frac{1}{2\sqrt{(1-x)^2}}(2(1-x))}{\frac{1}{2\sqrt{(1-x)^2}}(2(1-x)) + f_1(x)} = \frac{1}{1 + \frac{1}{2\sqrt{x}}} =$$

$$\alpha_x = \frac{f_0(x)}{f_0(x) + f_0(H(x))|H'(x)|}$$

$$\beta_x = \frac{f_1(x)}{f_1(H(x))|H'(x)| + f_1(x)} = \frac{f_1(x)}{f_1((1-x)^2)(2x) + f_1(x)} =$$

$$\frac{\frac{1}{2\sqrt{x}}}{\frac{1}{2\sqrt{(1-x)^2}}(2(1-x)) + \frac{1}{2\sqrt{x}}} = \frac{\frac{1}{2\sqrt{x}}}{1 + \frac{1}{2\sqrt{x}}} = \frac{1}{1 + 2\sqrt{x}}$$

O teste com taxas de erros condicionais com respeito ao *valor-p* é, então,

$$Y^p = \begin{cases} \text{se } x \leq 0,382, & \text{rejeita-se } H_0 \text{ e apresenta-se PEC I } \alpha_x = \left(1 + \frac{1}{2}x^{-\frac{1}{2}}\right)^{-1} \\ \text{se } x > 0,382, & \text{aceita-se } H_0 \text{ e apresenta-se PEC II } \beta_x = \left(1 + 2x^{\frac{1}{2}}\right)^{-1} \end{cases}$$

O condicionamento baseado em *valores-p* apresenta probabilidade de erro condicional superior a 50% no intervalo de $x \in]0,25;0,382[$ e isto é inadequado dependendo do critério estabelecido para o teste, pois, com uma probabilidade de erro em rejeitar superior a 0,5, o intuitivo seria aceitar a hipótese nula.

São apresentados, na Tabela 2.3 a seguir, para diferentes valores de amostra observados, os erros condicionais resultantes quando utilizadas as estatísticas condicionantes discutidas nos exemplos anteriores, e na Figura 2.27, o gráfico obtido de acordo com cada condicionamento.

TABELA 2.3 – Probabilidades de erros condicionais (PEC), para alguns valores de x , de acordo com o valor crítico da decisão do respectivo teste, quando utilizadas diferentes estatísticas condicionantes.

x observado	Estatística Condicionante			
	Ancilar	Intrínseca	Erros Iguais	valor-p
0,05	0,22	1,00	0,33	0,31
0,1	0,32	0,14	0,38	0,39
0,2	0,45	0,39	0,43	0,47
0,3	0,50	0,45	0,47	0,52
0,4	0,50	0,38	0,50	0,44
0,5	0,50	0,33	0,47	0,41
0,6	0,50	0,29	0,44	0,39
0,7	0,50	0,26	0,40	0,37
0,8	0,50	0,24	0,36	0,36
0,9	0,50	0,22	0,30	0,35
0,99	0,50	0,20	0,16	0,33

* Os valores em negrito correspondem à PEC I, e os outros, à PEC II.

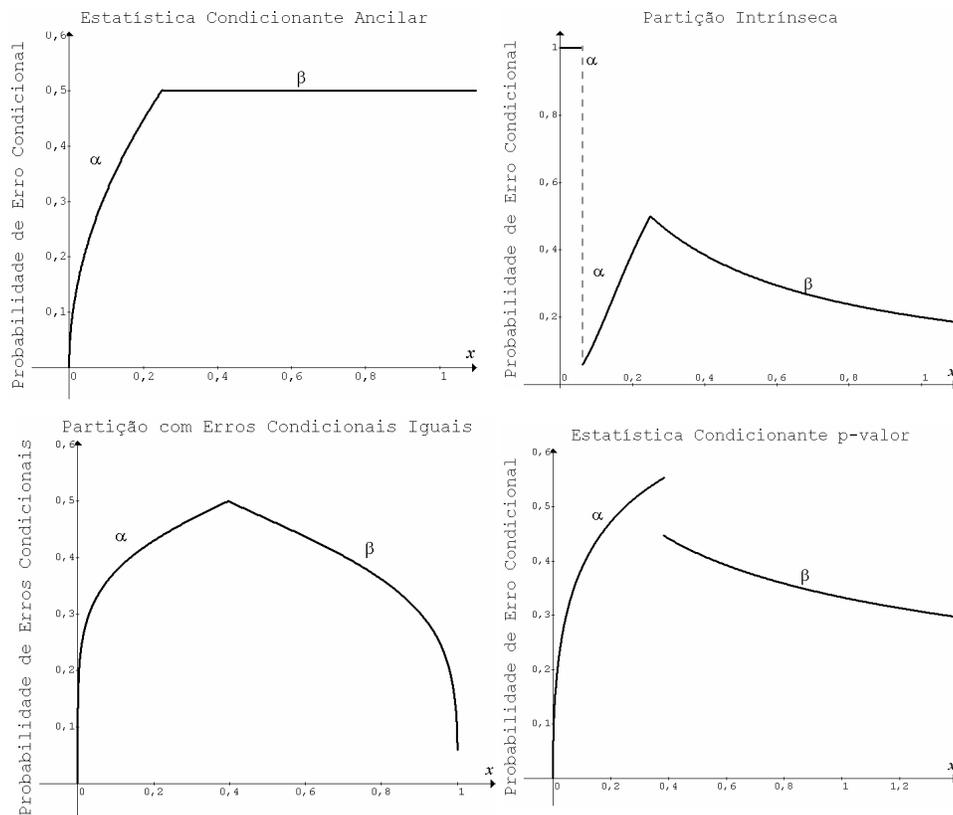


FIGURA 2.27 – Probabilidades de Erro Condicional Tipo I e Tipo II para diferentes condicionamentos.

Comentários Gerais

Para problemas “simétricos”, todos os métodos propostos para selecionar uma função particionante $H(x)$ concordaram, como pode ser visto no exemplo 1. Todavia, para problemas não-simétricos, como no exemplo 2, notou-se que o condicionamento:

- Baseado em uma estatística ancilar pode resultar em probabilidades de erro condicionais não naturais para este tipo de teste, em que se desejam probabilidades de erro dependentes dos dados e estes foram constantes em determinadas regiões.
- A partir dos níveis de significância intrínseca propostos por Birnbaum (1961) (definindo a estatística condicionante Z por meio da razão de verossimilhanças) não é adequado, pois, neste exemplo, possui valores em que a probabilidades de erros resultantes são constantes, sendo este condicionamento apropriado apenas para o caso simétrico.
- Com o critério de probabilidades de erros condicionais iguais proposto por Kiefer (1977) parece funcionar adequadamente, mas, neste exemplo, possui erro condicional tipo II tendendo a zero muito rápido para valores ainda pouco significativos da variável aleatória $B(x)$, mas isto pode ser consequência do pequeno tamanho de amostra.
- Utilizando a estatística condicionante baseada em *valores-p* funciona bem, mas possui valores de probabilidade de erro condicional superiores a 0,5 na região de em que $x \in]0,25; 0,382[$.

Então, no caso de razão de verossimilhanças simétrica, as quatro estatísticas condicionantes utilizadas concordam entre si e estão também em concordância com o valor crítico do teste de hipótese de Neyman-Pearson. Mas, para o caso de razão de verossimilhanças não-simétrica, os valores críticos para os testes com taxas de erros condicionais variam com a estatística condicionante utilizada. Tal fato implica que, para valores próximos aos valores críticos, as decisões tomadas poderão também ser diferentes.

4.3 Região de Não Decisão para Unificação de Decisões em Testes Frequentistas Condicionais

O que se propõe é utilizar um critério para unificação das decisões tomadas para as várias estatísticas condicionantes utilizadas, através do uso de uma denominada “região de não-decisão” (RND) (Berger, Brown & Wolpert, 1994 - ver comentários no capítulo 3 deste trabalho). A RND deve ser um recurso apenas para o caso em que se deseje comparar testes baseados em estatísticas condicionantes diferentes, caso em que normalmente não há concordância do valor crítico utilizado. Ou seja, definida uma única estatística condicionante, não é necessária a utilização da RND.

A justificativa de tal procedimento é que as probabilidades de erros condicionais apresentadas dependem da estatística utilizada no condicionamento e são diferentes; mas, pelo menos as decisões tomadas devem ser as mesmas. Este problema ocorre apenas no caso de razão de verossimilhança não-simétrica.

Para os condicionamentos ancilar e intrínseco o valor crítico é 0,25, para erros condicionais, é 0,397, e para *valor-p*, é 0,382.

O valor crítico do teste não-condicional com $\alpha = \beta$ é dado por:

$$F_0(x) = 1 - F_1(x) \Rightarrow \int_0^x dy = 1 - \int_0^x \frac{(y)^{\frac{1}{2}-1} (1-y)^{1-1}}{B(1/2, 1)} dy \Rightarrow x = 1 - \frac{1}{2} \int_0^x \frac{1}{\sqrt{y}} dy \Rightarrow ,$$
$$x = 1 - \sqrt{x} \Rightarrow x = 0,382$$

correspondendo, então, ao valor crítico obtido no teste utilizando a estatística condicionante baseada no *valor-p*.

As probabilidades de erro não-condicionais seriam, então, iguais a:

$$\alpha = P_{\theta_0} [\text{rejeitar } H_0] = P_{\theta_0} [x \leq 0,382] = \int_0^{0,382} 1 dx = 0,382$$

$$\beta = P_{\theta_1} [\text{aceitar } H_0] = P_{\theta_1} [x > 0,382] = 1 - \int_0^{0,382} \frac{(x)^{\frac{1}{2}-1} (1-x)^{1-1}}{B(1/2,1)} dx =$$

$$= 1 - \frac{1}{2} \int_0^{0,382} \frac{1}{\sqrt{x}} dx = 1 - \sqrt{0,382} \Rightarrow \beta = 0,382$$

Para o valor crítico 0,397, as probabilidades de erros não-condicionais são dadas por:

$$\alpha = P_{\theta_0} [\text{rejeitar } H_0] = P_{\theta_0} [x \leq 0,397] = \int_0^{0,397} 1 dx = 0,397$$

$$\beta = P_{\theta_1} [x > 0,397] = 1 - \int_0^{0,397} \frac{(x)^{\frac{1}{2}-1} (1-x)^{1-1}}{B(1/2,1)} dx = 1 - \frac{1}{2} \int_0^{0,397} \frac{1}{\sqrt{x}} dx \Rightarrow \beta = 0,3699,$$

fornecendo, portanto, probabilidades de erros não-condicionais diferentes.

Com o valor crítico 0,25 obtido pelas estatísticas condicionantes ancilar e intrínseca, têm-se as probabilidades de erros dadas por:

$$\alpha = P_{\theta_0} [\text{rejeitar } H_0] = P_{\theta_0} [x \leq 0,25] = \int_0^{0,25} 1 dx = 0,25$$

$$\beta = P_{\theta_1} [x > 0,25] = 1 - \int_0^{0,25} \frac{(x)^{\frac{1}{2}-1} (1-x)^{1-1}}{B(1/2,1)} dx = 1 - \frac{1}{2} \int_0^{0,25} \frac{1}{\sqrt{x}} dx \Rightarrow \beta = 0,5.$$

Observa-se, nas discussões anteriores, que o condicionamento em uma estatística ancilar ou intrínseca fornece resultados não adequados em problemas com razão de verossimilhanças não-simétrica. No entanto, para problemas deste tipo, as estatísticas baseadas em erros condicionais iguais e *valor-p* mostraram-se adequadas, sendo que, no exemplo desenvolvido, o problema mais grave detectado está no intervalo [0,382; 0,397]. Por exemplo, na obtenção de uma amostra com valor 0,39, um teste decidiria por rejeitar e outro, por aceitar. A

sugestão para resolução desta dificuldade é a extensão do conceito de região de não-decisão proposto por Berger, Brown e Wolpert (1994), que seria um intervalo em que nenhuma decisão seria tomada, promovendo certa unificação entre os testes com taxas de erros condicionais propostos por Kiefer (1977) e Berger (2003).

O procedimento proposto é o seguinte:

Dado que o valor crítico c_1 do condicionamento com $\alpha_x = \beta_x$ é dado pela solução da equação integral dada pela equação (1.8):

$$\int_{-\infty}^{c_1} \sqrt{f_0(x)f_1(x)}dx = \int_{c_1}^{+\infty} \sqrt{f_0(x)f_1(x)}dx$$

E o valor crítico c_2 , do condicionamento baseado no *valor-p* fornecendo $\alpha = \beta$, é dado pela solução de:

$$\int_{-\infty}^{c_2} f_0(x)dx = \int_{c_2}^{+\infty} f_1(x)dx = c_2 \text{ ou } F_0(c_2) = 1 - F_1(c_2) \text{ (Figura 2.28).}$$

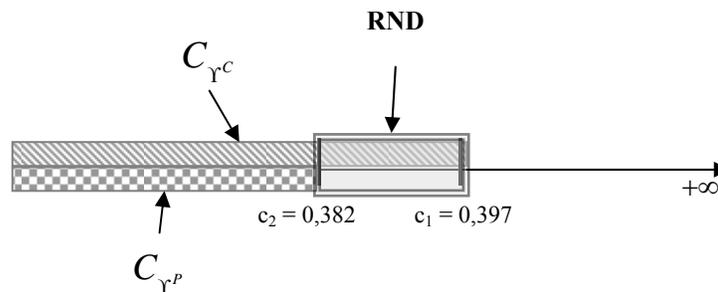


FIGURA 2.28 - Representação dos valores críticos resultantes das estatísticas condicionantes baseadas nos critérios de probabilidade de erros condicionais iguais e *valor-p*.

Tem-se que uma região de não-decisão que evitaria discordância entre estas duas metodologias seria dada pela região estabelecida da seguinte maneira:

$$RND = (\min(c_1, c_2); \max(c_1, c_2))$$

A probabilidade de erro apresentada em cada uma das metodologias continua a ser diferente, mas as decisões fora deste intervalo seriam as mesmas. No exemplo da Uniforme (0,1) versus a Beta (1/2, 1), o teste fica da seguinte maneira:

$$Y^* = \begin{cases} \text{se } x \leq 0,382, \text{ ou } B(x) \leq 1,236 & \text{rejeita-se } H_0 \text{ e apresenta-se PEC I;} \\ \text{se } 0,382 < x < 0,397 \text{ ou } 1,236 < B(x) < 1,26, & \\ & \text{n\~{a}o se toma nenhuma decis\~{a}o;} \\ \text{se } x \geq 0,397 \text{ ou } B(x) \geq 1,26, & \text{aceita-se } H_0 \text{ e apresenta-se PEC II.} \end{cases}$$

As probabilidades de erros condicionais I e II ser\~{a}o determinadas de acordo com o condicionamento optado para o teste.

Caso se inclua a RND, considerando a hip\~{o}tese nula como verdadeira, a probabilidade de ocorr\~{e}ncia de um valor pertencente \~{a} RND \u00e9 de 0,015, e sob a hip\~{o}tese alternativa, de 0,012 (Figura 2.29).

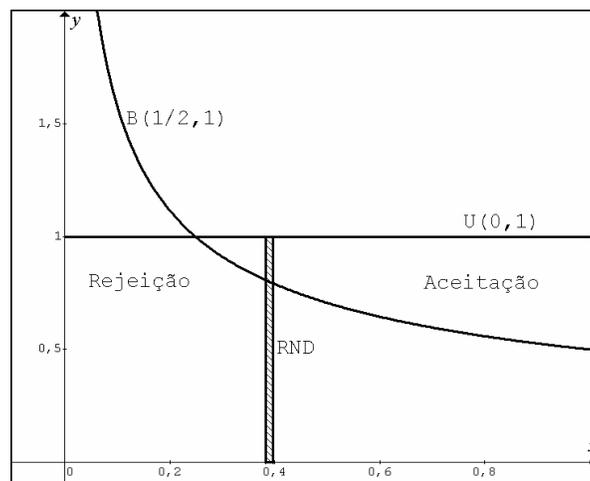


FIGURA 2.29 – Regiões de Aceitação, Rejeição e Não decisão para o teste da Uniforme (0,1) versus a Beta (1/2, 1).

Para comprovar como a inclusão da região de não-decisão afeta a decisão final em teste de hipóteses com razão de verossimilhanças não-simétrica, foi utilizado um programa no SAS (1990) (Anexo A).

As amostras foram obtidas da distribuição Uniforme (0,1) versus Beta (1/2,1) e avaliou-se a taxa de erro Tipo I e Tipo II para o procedimento freqüentista não condicional com a inclusão da região de não-decisão. Verificou-se o número de vezes em que não houve tomada de decisão e os erros cometidos.

O teste com região de não decisão forneceu taxa de erro tipo I igual a 37,86% e taxa de erro tipo II igual a 37,18% e não houve decisão em 1,34% dos casos simulados. O número de casos em que não se tomaram decisões é razoavelmente pequeno. Estes valores estão de acordo com os calculados pela respectiva integração associada às hipóteses nula e alternativa.

Se o condicionamento com uma estatística ancilar ou intrínseca é incorporado, o teste pode ser construído considerando ainda o valor c_3 como

solução da equação $B(x) = \frac{f_0(x)}{f_1(x)} = 1$, e a RND é dada, então, por:

$$RND = (\min(c_1, c_2, c_3); \max(c_1, c_2, c_3)) \text{ (Figura 2.30)}$$

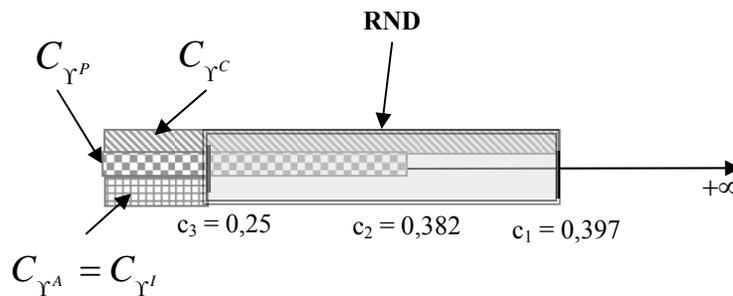


FIGURA 2.30 - Representação dos valores críticos resultantes das estatísticas condicionantes com erros condicionais iguais e *valores-p*.

E o teste é dado por:

$$Y^* = \begin{cases} \text{se } x \leq 0,25, \text{ ou } B(x) \leq 1 \text{ rejeita-se } H_0 \text{ e apresenta-se PEC I;} \\ \text{se } 0,25 < x < 0,397 \text{ ou } 1 < B(x) < 1,26, \text{ não se toma nenhuma decisão;} \\ \text{se } x \geq 0,397 \text{ ou } B(x) \geq 1,26, \text{ aceita-se } H_0 \text{ e apresenta-se PEC II.} \end{cases}$$

Considerando a hipótese nula sendo verdadeira, a probabilidade de ocorrência de um valor na RND é de 0,147, e na hipótese alternativa, é de 0,1301 (Figura 2.31).

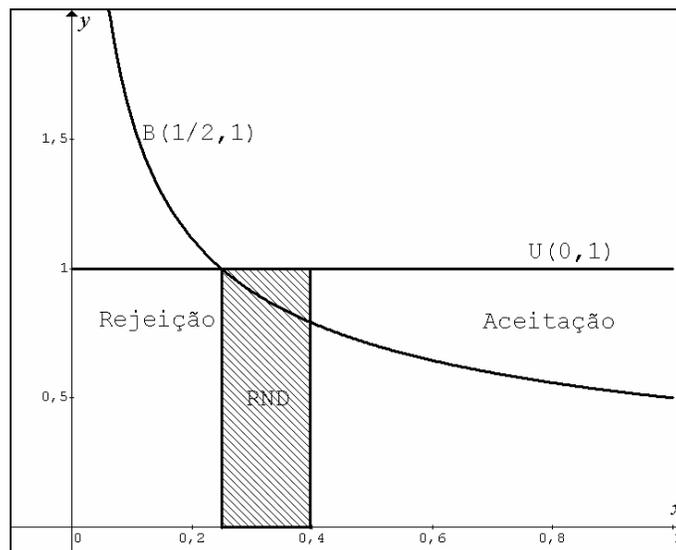


FIGURA 2.31 – Regiões de Aceitação, Rejeição e Não decisão para o teste da Uniforme (0,1) versus a Beta (1/2,1).

Novamente foi utilizado um programa no SAS (1990) (Anexo A) para verificar como a inclusão da RND afeta a decisão final em teste de razão de verossimilhanças não-simétrica. As amostras foram obtidas da distribuição

Uniforme (0,1) versus Beta (1/2,1). Verificou-se o número de vezes em que não houve tomada de decisão e as taxas de erros Tipo I e Tipo II.

O teste com região de não decisão forneceu taxa de erro tipo I igual a 25,14 % e taxa de erro tipo II igual a 37,12% e não houve decisão em 13,95 % dos casos simulados. Estes valores estão também de acordo com os obtidos pela integração das respectivas funções correspondentes às hipóteses nula e alternativa.

A taxa de erro tipo I foi consideravelmente diminuída, porém o número de casos em que não se tomaram decisões foi igualmente elevado.

É importante lembrar que, no caso de razão de verossimilhanças simétrica, a região de não decisão se torna vazia pois $c_1 = c_2 = c_3$.

4.4 Estudo da distribuição da estatística condicionante baseada em *valores-p*

Para avaliar a distribuição da estatística condicionante baseada nos *valores-p*, $Z = \max \{F_T^0(T), 1 - F_T^1(T)\}$, deve-se estudar o seu comportamento, tanto sob a hipótese nula quanto sob a hipótese alternativa, para verificar os valores que têm maior probabilidade de ocorrência, sendo assim:

a) Sob a Hipótese Nula H_0

$$F_Z^0(z) = P_{\theta_0} [Z \leq z].$$

Pela Figura 2.32, observa-se que

$$F_Z^0(z) = P_{\theta_0} [t' \leq T \leq t] = \int_{t'}^t f_T^0(u) du = F_T^0(t) - F_T^0(t'). \quad (1)$$

Mas t' pode ser expresso da seguinte maneira:

$$\text{Sabendo que } F_T^0(t) = z \Rightarrow t = [F_T^0]^{-1}(z) \quad (2)$$

$$E \quad 1 - F_T^1(t') = z \quad (3)$$

Então, substituindo (3) em (2), tem-se:

$$F_T^0(t) = 1 - F_T^1(t') \Rightarrow F_T^1(t') = 1 - F_T^0(t) \Rightarrow t' = [F_T^1]^{-1}(1 - F_T^0(t)) = H(t) \quad (4)$$

Substituindo (4) e (2) em (1), obtém-se:

$$\begin{aligned} F_Z^0(z) &= F_T^0(t) - F_T^0(t') \Rightarrow F_Z^0(z) = F_T^0(t) - F_T^0([F_T^1]^{-1}(1 - F_T^0(t))), \\ F_Z^0(z) &= F_T^0([F_T^0]^{-1}(z)) - F_T^0([F_T^1]^{-1}(1 - F_T^0([F_T^0]^{-1}(z)))) \\ F_Z^0(z) &= z - F_T^0([F_T^1]^{-1}(1 - z)) \quad (5) \end{aligned}$$

Derivando, tem-se:

$$\begin{aligned} f_Z^0(z) &= F_Z^0'(z) = 1 - f_T^0([F_T^1]^{-1}(1 - z)) \cdot ([F_T^1]^{-1})'(1 - z)(-1) \\ f_Z^0(z) &= 1 + f_T^0([F_T^1]^{-1}(1 - z)) \cdot ([F_T^1]^{-1})'(1 - z) \quad (6) \end{aligned}$$

Observação: Derivando-se a identidade $F_X^1 \circ [F_X^1]^{-1}(x) = x$, obtém-se

$$\begin{aligned} F_X^1 \circ [F_X^1]^{-1}(x) &= x \Rightarrow (F_X^1)'([F_X^1]^{-1}(x)) \cdot ([F_X^1]^{-1})'(x) = 1 \\ \Rightarrow f_X^1([F_X^1]^{-1}(x)) \cdot ([F_X^1]^{-1})'(x) &= 1 \Rightarrow ([F_X^1]^{-1})'(x) = \frac{1}{f_X^1([F_X^1]^{-1}(x))} \quad (7) \end{aligned}$$

Então, utilizando (7) em (6), tem-se:

$$f_Z^0(z) = 1 + f_T^0([F_T^1]^{-1}(1 - z)) \frac{1}{f_T^1([F_T^1]^{-1}(1 - z))} = 1 + \frac{f_T^0([F_T^1]^{-1}(1 - F_T^0(t)))}{f_T^1([F_T^1]^{-1}(1 - F_T^0(t)))}$$

$$\boxed{f_Z^0(z) = 1 + \frac{f_T^0(H(t))}{f_T^1(H(t))} = 1 + B(H(t)) = \frac{1}{\beta_z}}$$

Ou seja, a distribuição da estatística condicionante baseada em *valores-p* possui, sob a hipótese nula, distribuição inversamente proporcional à PEC II.

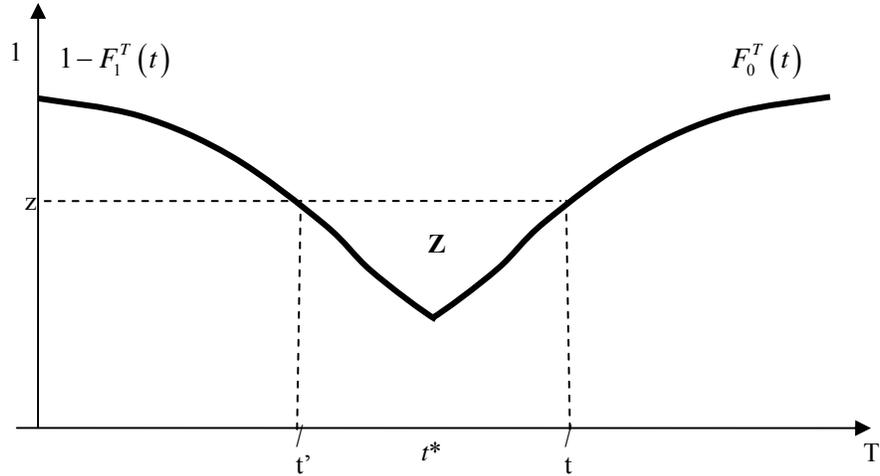


FIGURA 2.32 - Estatística Condicionante baseada em *valores-p*

$$Z = \max \{ F_0^T(T), 1 - F_1^T(T) \}.$$

b) Sob a hipótese alternativa H_1

$$F_Z^1(z) = P_{\theta} [Z \leq z] = P_{\theta} [t' \leq T \leq t] = F_T^1(t) - F_T^1\left(\left[F_T^1\right]^{-1}(1 - F_T^0(t))\right) \quad (8)$$

Substituindo (2) em (8), tem-se:

$$F_Z^1(z) = F_T^1(t) - (1 - F_T^0(t)) = F_T^1\left(\left[F_T^0\right]^{-1}(z)\right) - 1 + F_T^0\left(\left[F_T^0\right]^{-1}(z)\right)$$

$$F_Z^1(z) = F_T^1\left(\left[F_T^0\right]^{-1}(z)\right) - 1 + z$$

Derivando, tem-se:

$$f_Z^1(z) = (F_T^1)'\left(\left[F_T^0\right]^{-1}(z)\right)\left(\left[F_T^0\right]^{-1}\right)'(z) + 1$$

$$f_z^1(z) = (F_T^1)' \left([F_T^0]^{-1}(z) \right) \left([F_T^0]^{-1} \right)'(z) + 1$$

$$f_z^1(z) = f_T^1(t) \left([F_T^0]^{-1} \right)'(z) + 1.$$

Observação: Derivando-se $F_X^0 \circ [F_X^0]^{-1}(x) = x$, obtém-se:

$$\begin{aligned} (F_X^0)' [F_X^0]^{-1}(x) \left([F_X^0]^{-1} \right)'(x) &= 1 \\ \Rightarrow f_X^0 \left([F_X^0]^{-1}(x) \right) \cdot \left([F_X^0]^{-1} \right)'(x) &= 1 \Rightarrow \left([F_X^0]^{-1} \right)'(x) = \frac{1}{f_X^0 \left([F_X^0]^{-1}(x) \right)}. \end{aligned}$$

Então

$$f_z^1(z) = f_T^1(t) \frac{1}{f_T^0 \left([F_T^0]^{-1}(z) \right)} + 1 = \frac{f_T^1(t)}{f_T^0(t)} + 1 = \frac{1}{B(t)} + 1 \Rightarrow$$

$$\boxed{f_z^1(z) = \frac{1+B(t)}{B(t)} = \frac{1}{\alpha_z}}$$

Ou seja, em caso de hipótese alternativa, a distribuição é inversamente proporcional a PEC I.

Note que para $Z = \max \{ F_0^T(T), 1 - F_1^T(T) \}$, tem-se dois comportamentos para a estatística Z, pois $f_z^0(z) = \frac{1}{\beta_z}$ e $f_z^1(z) = \frac{1}{\alpha_z}$. Assim

sendo, pode-se afirmar que:

1ª) Para valores grandes da variável T, pelo Teorema da Probabilidade Integral, a tendência é de que Z tenha distribuição uniforme sob H_0 e seja não-uniforme sob H_1 , pois o máximo será obtido em $F_0^T(T)$, que é uniforme sob H_0 ($Z = \max \{ \text{uniforme}, 1 - F_1^T(T) \}$).

2ª) Para valores pequenos da variável T, a tendência é de que Z tenha distribuição uniforme sob H_1 e seja não-uniforme sob H_0 . Seguindo o mesmo

comentário feito anteriormente, tem-se que para t pequenos, o máximo será obtido em $1 - F_1^T(T)$ e, portanto, será uniforme sob H_1 e não uniforme sob H_0 .

→ Resumindo:

$$\text{Distribuição de } Z \left\{ \begin{array}{l} \text{Uniforme} \left\{ \begin{array}{l} \text{sob } H_0 \rightarrow \text{valores grandes de } t \\ \text{sob } H_1 \rightarrow \text{valores pequenos de } t \end{array} \right. \\ \text{Não-Uniforme} \left\{ \begin{array}{l} \text{sob } H_0 \rightarrow \text{valores pequenos de } t \\ \text{sob } H_1 \rightarrow \text{valores grandes de } t \end{array} \right. \end{array} \right.$$

Observação:

As PEC's obtidas quando a estatística condicionante está baseada em *valores-p* são numericamente iguais às probabilidades *a posteriori* de um teste de hipóteses no contexto bayesiano dadas por:

$$\alpha_x = \frac{B(x)}{1+B(x)} = \frac{2\sqrt{x}}{1+2\sqrt{x}} = \left(1 + \frac{1}{2}x^{-\frac{1}{2}} \right)^{-1}$$

e

$$\beta_x = \frac{1}{1+B(x)} = \frac{1}{1+2\sqrt{x}} = \left(1 + 2x^{\frac{1}{2}} \right)^{-1} \quad (\text{Ver capítulo 3}).$$

4.5 Exemplo de aplicação

A teoria do teste freqüentista condicional foi aplicada a um exemplo com dados reais de duas variedades de café (Icatu e Catuai).

Uma das possíveis análises efetuadas em variedades de café é o teor de amido existente na raiz da planta. Sabe-se que a quantidade de amido presente na raiz pode ser obtida por uma avaliação bioquímica e que esta quantidade

apresenta-se normalmente distribuída, sendo que, para a variedade Icatu 3686, a média obtida por experimentos anteriores é de $m_I = 1,05$, e para a variedade Catuaí 2SL, $m_C = 1,46$. Considerou-se o mesmo desvio padrão para as duas populações de $(0,1)^{1/2}$ (Figura 2.33).

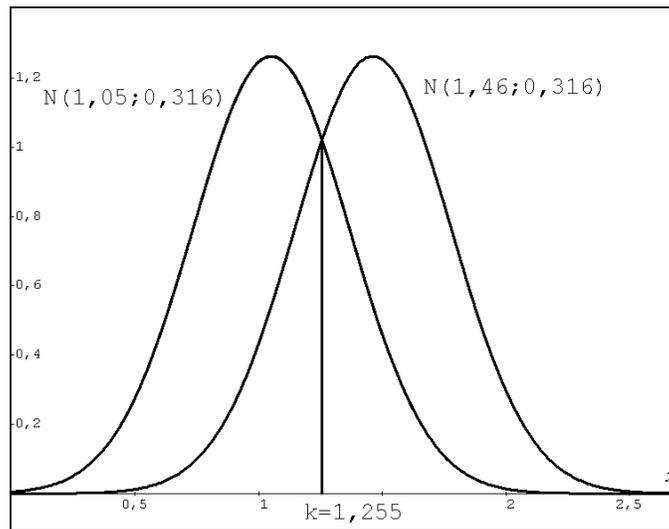


FIGURA 2.33 - Funções densidade de probabilidade de populações normais de média 1,05 e de 1,46.

A análise de amido na raiz de uma amostra de tamanho três de certa variedade revelou média $m_t = 1,248$. Deseja-se saber se esta variedade avaliada é do Icatu ou do Catuaí.

As hipóteses seriam:

H_0 : a média da população a que pertencem as amostras de café é 1,05, ou $m_t = m_I = 1,05$;

H_1 : a média da população a que pertencem as amostras de café testada é 1,46, ou $m_t = m_C = 1,46$.

Como tais distribuições se sobrepõem, deverá ser estabelecido algum ponto de separação entre as populações. No caso de um teste de hipóteses tradicional, este é determinado tendo por base a probabilidade de erro tipo I. Os pesquisadores em geral desejam uma probabilidade pequena de rejeitar incorretamente H_0 . Suponha que a escolha seja 5%. A área de rejeição de H_0 , nessa situação de teste, é unilateral e situa-se na cauda direita da curva, já que a hipótese alternativa considera somente uma média maior que 1,05. As duas são distribuições normais, centradas em 1,46 e 1,05, então o erro padrão será dado

$$\text{por: } \sigma(\bar{x}) = \frac{\sqrt{0,1}}{\sqrt{3}} = 0,182.$$

Como o teste é unilateral, a média que limita a região crítica na distribuição em 1,05 é dada por:

$$\bar{x}_{\text{limite}} = m_a + 1,64\sigma_{(\bar{x})} \Rightarrow \bar{x}_{\text{limite}} = 1,05 + (1,64 \times 0,182) = 1,35.$$

Para obter a probabilidade do erro tipo II, procura-se inicialmente o valor correspondente na normal padrão para $\bar{x} = 1,35$. Ora, considerando a hipótese alternativa, tem-se:

$$z = \frac{\bar{x} - m_b}{\sigma_{\bar{x}}} = \frac{\bar{x} - 1,46}{0,182} \Rightarrow z = \frac{1,35 - 1,46}{0,182} = -0,604.$$

Isto é, na população da variedade Catuaí, a média 1,35 está a 0,604 unidades de erro padrão abaixo da média, que corresponde à probabilidade de erro tipo II igual a 0,269.

Como a média amostral foi de 1,248, a decisão a ser tomada é a de que as amostras são da variedade Icatu, com probabilidade de erro $\alpha = 0,05$ e $\beta = 0,269$ (Figura 2.34). O poder do teste, no exemplo, é $1 - 0,269 = 0,731$.

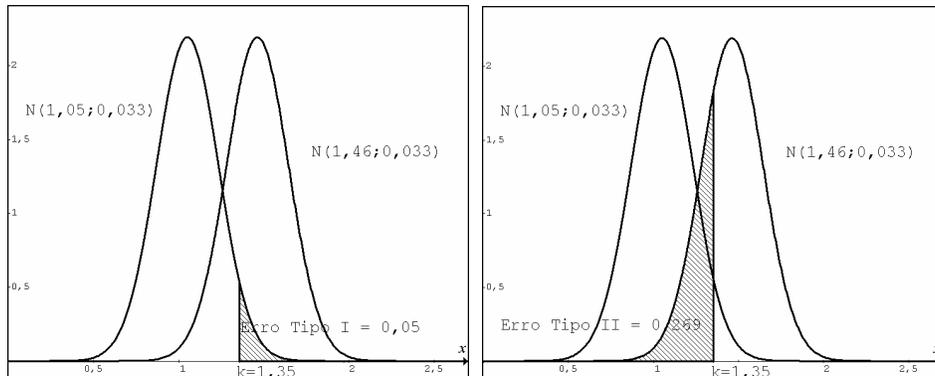


FIGURA 2.34 – Áreas correspondentes aos erros Tipo I e II sob as hipóteses nula e alternativa, respectivamente.

Para este teste, sendo a média observada igual a 1,248, há então uma aceitação da hipótese nula, podendo-se, nesse caso, estar cometendo o erro tipo II, cuja probabilidade seria de 26,9%.

Se fosse considerado um teste de significância, o *valor-p* calculado seria de 0,138.

Se as hipóteses fossem invertidas, o teste seria dado pelo valor crítico $k = 1,162$, invertendo-se também as probabilidades de erro tipo I e II. O *valor-p* correspondente seria, agora, igual a 0,119.

No caso de teste freqüentista condicional sob a condição de erros não-condicionais iguais ($\alpha = \beta$), o valor crítico utilizado para o teste seria 1,255, o que levaria também a uma aceitação de H_0 , mas com erro tipo II igual a 13%. A razão de verossimilhança fica sendo:

$$B = \prod_{i=1}^n \frac{(0,2\pi)^{-1/2} \exp\left\{-\frac{1}{2}\left(\frac{x_i - 1,05}{\sqrt{0,1}}\right)^2\right\}}{(0,2\pi)^{-1/2} \exp\left\{-\frac{1}{2}\left(\frac{x_i - 1,46}{\sqrt{0,1}}\right)^2\right\}} = \prod_{i=1}^n \exp\left\{-\frac{1}{2}\left(\frac{x_i - 1,05}{0,316}\right)^2 + \frac{1}{2}\left(\frac{x_i - 1,46}{0,316}\right)^2\right\} =$$

$$\prod_{i=1}^n \exp\left\{-\frac{1}{0,2}(0,82x_i - 1,029)\right\} = \exp\left\{\sum_1^n (-4,1x_i + 5,145)\right\} = \exp\{-4,1n\bar{x} + 5,145n\}$$

e o teste fornece:

$$Y = \begin{cases} \text{se } B(x) \leq 1 \text{ ou } \bar{x}_3 \geq 1,255, \text{ rejeita-se } H_0 \text{ com PEC I} \\ \alpha_x = \frac{B(x)}{1+B(x)} = \frac{\exp\{-12,3\bar{x} + 15,435\}}{1 + \exp\{-12,3\bar{x} + 15,435\}} \\ \text{se } B(x) > 1 \text{ ou } \bar{x}_3 < 1,255, \text{ aceita-se } H_0 \text{ com PEC I} \\ \beta_x = \frac{1}{1+B(x)} = \frac{1}{1 + \exp\{-12,3\bar{x} + 15,435\}} \end{cases}$$

No caso da amostra observado com média igual a 1,248, aceita-se a hipótese de que a variedade testada seja proveniente da variedade Icatu, com probabilidade de erro condicional tipo II igual a 0,479. Ou seja, apesar de proporcionar aceitação da amostra, como no teste tradicional, percebe-se que o valor da PEC é muito alto, mostrando pouca evidência a favor de H_0 .

Nesse caso, todas as estatísticas condicionantes discutidas neste trabalho concordariam, pois se trata de razão de verossimilhança simétrica.

CONCLUSÕES

Os aspectos matemáticos subjacentes à teoria de testes de hipóteses com erros freqüentistas condicionais, sistematizada por Kiefer, são acessíveis e foram detalhados neste trabalho.

O uso sistemático da definição de função particionante simplifica de maneira significativa os conceitos da teoria de testes de hipóteses com erros freqüentistas condicionais.

Para o problema na presença de razão de verossimilhança simétrica, independentemente das estatísticas condicionantes utilizadas, os testes coincidem.

No caso em que a razão de verossimilhança é não-simétrica, os testes com erros condicionais iguais e os definidos pela estatística condicionante *valores-p* apresentaram uma maior coerência.

A inclusão de uma região de não decisão pode trazer propriedades desejáveis aos testes com erros freqüentistas condicionais.

Apesar de a estatística condicionante *valores-p* não ser ancilar, sua distribuição apresenta comportamento similar sob a hipótese nula e sob a hipótese alternativa, mostrando-se uma boa alternativa para o condicionamento.

REFERÊNCIAS BIBLIOGRÁFICAS

BERGER, J. O. Could Fisher, Jeffreys and Neyman have agreed on testing? **Statistical Science**, Beachwood, v. 18, n. 1, p. 1-12, Feb. 2003. (Com discussão: p. 13-32)

BERGER, J. O.; BROWN, L. D.; WOLPERT, R. L. A unified conditional frequentist and Bayesian test for fixed and sequential simple hypothesis testing. **The Annals of Statistics**, Hayward, v. 22, n. 4, p.1787-1807, Dec. 1994.

BERGER, J.; BOUKAI, B.; WONG, Y. Unified frequentist and Bayesian testing of a precise hypothesis. **Statistics Science**, Hayward, v. 12, n. 3, p. 133-148, 1997. (Com discussão: p. 149-160)

BIRNBAUM, A. On the foundations of statistical inference: binary experiments. **Annals of Mathematical Statistics**, v. 32, p. 414-435, 1961.

KIEFER, J. Conditional confidence statements and confidence estimators **Journal of the American Statistical Association**, Alexandria, v. 72, n. 360, p. 789-807, Dec. 1977. (Com discussão: p. 808-827).

LARSON, R. E.; HOSTETLER, R. P.; EDWARDS, B. H. **Cálculo com aplicações** 4. ed. Rio de Janeiro: LTC, 2001. 711 p.

PAULO, R. M. B. **Problems on the Bayesian/Frequentist Interface**. 2002. 90 p. Tese (Doctor of Philosophy) – Institute of Statistics and Decision Sciences, Duke University, USA.

SAS INSTITUTE SAS/STAT. **SAS/IML Software**: usage and reference, version 6. Cary, 1990. 501p.

CAPÍTULO 3

TESTES COM INTERPRETAÇÃO FREQUENTISTA CONDICIONAL E BAYESIANA

RESUMO

BRIGHENTI, CARLA REGINA GUIMARÃES. **TESTES COM INTERPRETAÇÃO FREQUENTISTA CONDICIONAL E BAYESIANA.** LAVRAS: UFLA, 2007. 64 p. (Tese – Doutorado em Estatística e Experimentação Agropecuária)*

Todo Teste Bayesiano é um Teste de Razão de Verossimilhança com valor crítico dado por $b_c = \rho l$, em que $\rho = \pi_1/\pi_0$ é a razão entre as densidades *a priori* e $l = l_1/l_0$ a razão entre as perdas. Sendo assim, para determinados valores de ρ e l , têm-se regiões críticas coincidentes para um teste bayesiano e para um teste com erros frequentistas condicionais. Mas a medida de evidência apresentada no teste bayesiano é a probabilidade *a posteriori*, enquanto no frequentista condicional, é a probabilidade de erro condicional (PEC). Dependendo da estatística condicionante utilizada e dos valores de ρ e l , esses valores podem ser numericamente iguais, definindo um teste “unificado”. Mas quando a probabilidade *a posteriori* ou a PEC são numericamente iguais e superiores a $1/2$, deve ser incorporada uma “região de não-decisão”, RND. A construção da RND depende da função particionante $H(x)$ do teste condicional e dos valores de ρ e l . Dessa forma, a RND pode apresentar comportamentos variados e o seu tamanho pode ser um fator complicador do teste unificado. Neste capítulo, utilizou-se o conceito de função particionante $H(x)$ para construir detalhadamente a RND e estudou-se o seu tamanho em três exemplos. No primeiro, calcularam-se as probabilidades de erro tipo I e II para o procedimento frequentista não condicional com e sem a inclusão da RND para o teste da Uniforme (0,1) versus Beta (1/2,1). No segundo, estudou-se a RND para o teste da distribuição Exponencial (1) versus Exponencial (2), com amostras de diferentes tamanhos. No terceiro, obteve-se explicitamente o teste unificado estudando o tamanho da RND em função da razão entre as perdas $l = l_1/l_0$. Concluiu-se que o tamanho da RND é dependente do tamanho da amostra e que determinados valores de l podem tornar o teste inadequado.

* Orientador: LUCAS MONTEIRO CHAVES - UFLA.

ABSTRACT

BRIGHENTI, CARLA REGINA GUIMARÃES. **TESTS WITH FREQUENTISTS CONDITIONAL AND BAYESIAN INTERPRETATION.** LAVRAS: UFLA, 2007. 64 p. (Thesis - Doctorate in Statistics and Agricultural Experimentation)*

Every Bayesian Test is a Likelihood Ratio Test with critical values given by $b_c = \rho l$, where $\rho = \pi_1/\pi_0$ it's the ratio between prior densities and $l = l_1/l_0$ the ratio between losses. In this way, for some determined values of ρ and l there are coincident critics regions for a bayesian test and for a conditional frequentist test. The evidence measure reported in the bayesian test is the posterior probability while the conditional test is the conditional error probability (CEP). Depending on the conditioning statistic used and the ρ and l values, these values can be numerically the same, defining a "unified" test. But when the prior probability or the CEP are numerically the same and superior to $1/2$, it must be incorporated a "no-decision region", NDR. The NDR construction depends on the partitioning function $H(x)$ of the conditional test and the p and l values. Its size can be a complicated factor for the unified test. In this chapter, the concept of the partitioning function was used to construct in details the NDR, and it was studied its size in three examples. In the first, Uniform (0,1) versus Beta (1/2,1), it was calculated the probabilities of error type I and II for the frequentists unconditional proceeding with and without the inclusion of the NDR. In the second, Exponential distribution (1) versus Exponential (2), it was done the same but with samples of different sizes. In the third, it was explicitly obtained the unified test studying the size of the NDR due to the ratio for the losses $l = l_1/l_0$. It was concluded that the size of the NDR depends on the size of the sample and that some l values can make the test inadequate.

* Advisor: Lucas Monteiro Chaves -UFLA.

1 INTRODUÇÃO

O problema de testar hipóteses estatísticas é um dos pontos de discordância entre Bayesianos e freqüentistas.

O método freqüentista tradicional constrói uma região crítica e descreve probabilidades de erros associadas. O uso desta abordagem freqüentista tradicional é criticado por descrever probabilidades de erros que não refletem informações proporcionadas pelos dados apresentados. Uma alternativa comum é um teste de significância usando o *valor-p* como medida de evidência contra a hipótese nula H_0 . Porém, o *valor-p* não é uma medida freqüentista verdadeira e tem suas próprias limitações como medida de evidência.

Para solucionar estes problemas, uma tentativa possível é modificar o método freqüentista tradicional incorporando procedimentos baseados em condicionamento que sejam dependentes dos dados. Em Kiefer (1977), o método freqüentista condicional foi formalizado. A idéia básica por detrás desse método é condicionar em uma estatística que mede a força de evidência dos dados e, então, fornecer probabilidades de erro condicionais (PEC's), que podem ser definidas como medidas das probabilidades dos erros tipo I e II, calculadas condicionalmente aos dados estarem em regiões determinadas a partir da partição do espaço amostral.

Uma outra abordagem alternativa é a Bayesiana, que é baseada na forma mais extrema de condicionamento, nos próprios dados observados, discutida no capítulo 1 desse trabalho.

Berger et al. (1994) sugeriram um procedimento que promovesse o acordo entre os métodos Bayesianos e freqüentistas e consideraram, então, o teste de hipótese simples versus hipótese simples, mostrando que o método freqüentista condicional pode ser equivalente ao método Bayesiano, dependendo

da estatística condicionante utilizada. O aspecto mais importante não é o fato de que o Bayesiano e o freqüentista condicional poderiam ter a mesma regra de decisão para rejeitar ou aceitar a hipótese nula, mas sim que apresentarão as mesmas medidas de evidência ao rejeitar ou aceitar. O teste proposto e a idéia de condicionamento indicam que há uma maneira de testar que apresenta medidas de evidências freqüentista condicional (Probabilidades de erros condicionais) e Bayesiana (probabilidades *a posteriori*) que são numericamente iguais. Uma vantagem é que um procedimento de teste unificado que tenha simultaneamente justificativas Bayesiana e freqüentista pode ter maior aceitação.

Alguns problemas com o chamado teste unificado podem ocorrer por causa da possibilidade de diferentes valores de perdas e densidades *a priori*; para solucionar esse problema, Berger et al. (1994) sugeriram a introdução de uma região de não decisão (RND). A RND corresponderia a uma região, $]r,a[$, na qual potencialmente nenhum estatístico sentiria que as evidências são fortes o suficiente para uma decisão conclusiva.

Em virtude do que foi mencionado, este capítulo objetivou utilizar o conceito de função particionante, desenvolvido para teste freqüentistas condicionais no capítulo 2, para auxiliar na construção do teste unificado durante o desenvolvimento da teoria de testes unificados da revisão bibliográfica, assim como verificar o comportamento da região de não decisão em três exemplos de testes de hipóteses simples com razão de verossimilhanças não-simétricas e o comportamento desta na presença de perdas assimétricas.

2 REFERENCIAL TEÓRICO

2.1 TESTE UNIFICADO PARA HIPÓTESES SIMPLES

Berger et al. (1994), resgatando o trabalho desenvolvido por Kiefer (1977) em testes de hipóteses com probabilidades de erros condicionais, utilizando a função particionante $H(z) = F_0^{-1}(1 - F_1(z))$ discutida no item 4.1.2 do capítulo anterior, observaram as semelhanças entre o teste freqüentista condicional Υ^* dado, na página 90, por:

$$\Upsilon^* = \begin{cases} \text{se } B(x) \leq 1, & \text{rejeita-se } H_0 \text{ com PEC I } \alpha_x = \frac{B(x)}{1 + B(x)} \\ \text{se } B(x) > 1, & \text{aceita-se } H_0 \text{ com PEC II } \beta_x = \frac{1}{1 + B(x)} \end{cases}$$

E o teste Bayesiano Υ_π , considerando perdas iguais e densidades *a priori* também iguais dado, na página 44, por:

$$\Upsilon_\pi = \begin{cases} \text{se } B(x) \leq 1, & \text{rejeita-se } H_0 \text{ com probabilidade a posteriori, } \alpha_x = \frac{B(x)}{1 + B(x)} \\ \text{se } B(x) > 1, & \text{aceita-se } H_0 \text{ com probabilidade a posteriori, } \beta_x = \frac{1}{1 + B(x)} \end{cases}$$

Ou seja, para o valor crítico b_c igual a 1, os valores das probabilidades de erro condicional tipo I e tipo II, apresentados como medida de evidência no teste freqüentista condicional, são numericamente iguais às probabilidades *a posteriori* encontradas no teste Bayesiano..

O teste Υ^* apresenta probabilidades de erro do Tipo I e Tipo II iguais, $\alpha = \beta$, e probabilidades de erros condicionais dadas por α_x e β_x , sendo que o valor máximo de α_x ou de β_x é $1/2$, uma vez que a região crítica é dada pelo valor crítico $b_c = 1$.

Note, então, que tanto um frequentista quanto um Bayesiano, observados os dados $x = (x_1, \dots, x_n)$, tomariam a mesma decisão e apresentariam a mesma medida de evidência. É claro que a interpretação desta medida de evidência é diferente; ou seja, para o frequentista é uma medida de probabilidade de erro condicional, e para o Bayesiano, é a probabilidade *a posteriori* (que é numericamente igual à perda esperada *a posteriori* quando se considera perda “0-1”, ou seja, $l_0 = l_1 = 1$).

Para os casos em que as perdas não são simétricas, $l_0 \neq l_1$, o teste Bayesiano apresenta, após a tomada de decisão, a perda esperada *a posteriori* da forma:

$$Y_1 = \begin{cases} \text{se } B(x) \leq b_c = \frac{l_1}{l_0}, \text{ rejeita-se } H_0 \\ \text{com perda esperada a posteriori } l_0 \alpha_x = l_0 \left(\frac{B(x)}{1 + B(x)} \right) \\ \text{se } B(x) > b_c = \frac{l_1}{l_0}, \text{ aceita-se } H_0 \\ \text{com perda esperada a posteriori } l_1 \beta_x = l_1 \left(\frac{1}{1 + B(x)} \right) \end{cases}$$

Como apresentado na página 44.

Nesta situação, supondo $l_1 = 3$ e $l_0 = 1$ e sendo observado $B(x) = 2$, o teste leva à decisão de rejeitar H_0 , apresentando perda esperada *a posteriori* de 0,66. Note que, neste caso, em que $l_0 = 1$, a perda esperada *a posteriori* é numericamente igual à probabilidade *a posteriori* de H_0 , para o caso de rejeição da hipótese nula.

A outra opção de medida de evidência em um teste Bayesiano é a probabilidade *a posteriori*, sendo, neste caso, o teste da forma:

$$\Upsilon_{\pi}^* = \begin{cases} \text{se } B(x) \leq b_c = \frac{l_1}{l_0}, \text{ rejeita-se } H_0 \\ \text{com probabilidade a posteriori } \alpha_x = \frac{B(x)}{1+B(x)} \\ \text{se } B(x) > b_c = \frac{l_1}{l_0}, \text{ aceita-se } H_0 \\ \text{com probabilidade a posteriori } \beta_x = \frac{1}{1+B(x)} \end{cases}.$$

Testes Bayesianos que apresentam como medida de evidência a probabilidade *a posteriori* foram anteriormente utilizados, conforme citado por Berger (2003):

“Harold Jeffreys, com sua abordagem “objetiva” na análise Bayesiana, propôs o seguinte teste Bayesiano:

- Definir o fator de Bayes (ou razão de verossimilhanças no caso simples) $B(x) = f(x|\theta_0)/f(x|\theta_1)$.
- Rejeitar H_0 quando $B(x) \leq 1$ e aceitar H_0 quando $B(x) > 1$.
- Apresentar as probabilidades de erro a posteriori das hipóteses dadas por:

$$\Pr(H_0 | x) = \frac{B(x)}{1+B(x)} \quad \left(\text{ou } \Pr(H_1 | x) = \frac{1}{1+B(x)} \right),$$

assumindo probabilidades a priori iguais a $1/2$ para as duas hipóteses e aplicando o teorema de Bayes“.

Tal procedimento justifica a abordagem de probabilidade *a posteriori* utilizada por Berger et al. (1994) para definir o teste unificado.

Os testes Υ_1 e Υ_{π}^* possuem o mesmo valor crítico l_1/l_0 . A única diferença entre os dois testes está na medida de evidência que acompanha a decisão, pois, em Υ_1 , a perda influencia também no cálculo da medida de evidência (perda esperada *a posteriori*), enquanto, em Υ_{π}^* , a probabilidade *a posteriori* não depende da perda.

O teste Υ_{π}^* é, então, um *teste unificado* Bayesiano e freqüentista condicional, mas apesar de a família de testes assim definida apresentar uma interpretação legitimamente bayesiana e freqüentista condicional, pode levar a situações de difícil explicação nas duas abordagens para os casos em que o valor crítico $b_c \neq 1$.

Se, por exemplo, forem considerados novamente os valores para perda $l_1 = 3$ e $l_0 = 1$, e for observado $B(x) = 2$, o teste Υ_{π}^* diz que: como $B(x) \leq 3$, rejeita-se H_0 , apresentando probabilidade de erro condicional ou probabilidade *a posteriori*

$$\alpha_x = \frac{B(x)}{1+B(x)} = \frac{2}{3} = 66,6\% .$$

Para o conceito freqüentista condicional

de probabilidade, este valor é totalmente incoerente com a decisão de rejeitar H_0 , isto é, questiona-se qual a razão de se rejeitar H_0 se esta hipótese tem 2/3 de chance de ser verdadeira. Para o Bayesiano, tal decisão poderia ser até justificada, já que leva em conta perdas assimétricas, ou seja, $l_0 \neq l_1$.

Berger et al. (1994) propuseram, então, uma modificação do teste Υ_{π}^* de forma a ampliar os casos em que o teste teria uma adequada interpretação genuinamente freqüentista e genuinamente bayesiana, considerando outros valores de perdas, além de “0-1” e densidades *a priori* não-uniforme.

2.1.1 UMA ALTERAÇÃO COM MOTIVAÇÃO BAYESIANA: A REGIÃO DE NÃO-DECISÃO

Uma modificação do teste Υ_{π}^* que resolve a dificuldade, comentada no item anterior, de a PEC ser maior que 1/2, é dada pela incorporação de uma “região de não-decisão”, RND.

2.1.1.1. CONSTRUÇÃO DA REGIÃO DE NÃO-DECISÃO

A incorporação de uma “região de não-decisão” tem o objetivo de impedir que se apresentem probabilidades de erros condicionais superiores a 0,5, o que poderia ocorrer em casos em que $l_0 \neq l_1$ e, assim, o valor crítico b_c é diferente de 1.

Considerando os testes em que se seleciona um b_c de tal forma que, no caso freqüentista, se obtenham erros tipo I e tipo II iguais, isto é, $F_B^0(b_c) = 1 - F_B^1(b_c)$, e utilizando a partição definida pela função particionante $H(z) = [F_B^0]^{-1}(1 - F_B^0(z))$, tem-se que esta fornece probabilidades de erro condicionais iguais (capítulo 2, item 4.1.1). Assim, a partir do valor crítico b_c , a cada valor z é associado um $H(z)$, que fornece a mesma probabilidade de erro condicional, sendo a PEC I para um dos pontos do par e a PEC II para o outro. Mas as PEC's possuem valores numericamente iguais, pois $F_B^0(H(z)) = 1 - F_B^1(z)$ ou $F_B^0(z) = 1 - F_B^1(H(z))$ dependendo da posição do valor crítico b_c em relação ao ponto $b=1$.

A RND exclui os pares $(z, H(z))$ que estão entre o valor crítico e o valor 1, pois, pelo condicionamento utilizado, estes são os pares que forneceriam probabilidades de erros condicionais superiores a 0,5. Avaliando as duas situações possíveis para b_c , tem-se:

1º Caso) O valor crítico é superior a 1, ou seja, $b_c > 1$.

Como em b_c tem-se $F_B^0(b_c) = 1 - F_B^1(b_c)$, para $b = 1$ as probabilidades de erro tipo I e tipo II serão diferentes, pois a função acumulada $F_B^0(b)$ é crescente, enquanto $1 - F_B^1(b)$ é decrescente e, como $b_c > 1$, conseqüentemente

$F_B^0(1) \leq 1 - F_B^1(1)$. Mas o problema é que as PEC's continuam assumindo os mesmos valores e, portanto, no valor de b exatamente igual a 1, as probabilidades dos erros condicionais dados por α_x e β_x são iguais a $\frac{1}{2}$; com isso, entre b_c e 1, estas probabilidades serão superiores a $\frac{1}{2}$ (Figura 3.1).

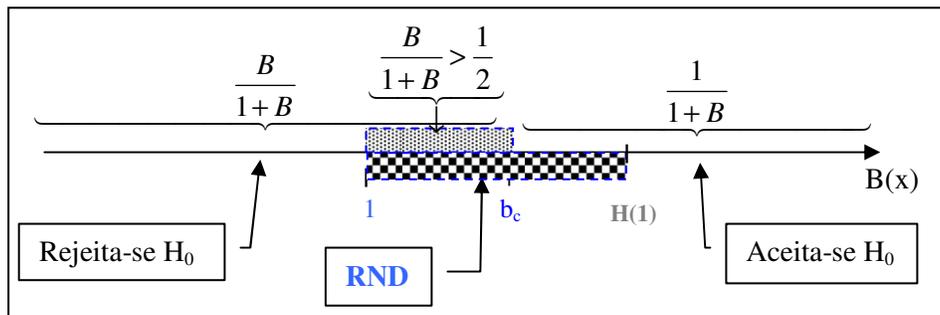


FIGURA 3.1 – Desenho esquemático da Região de Não Decisão para $b_c > 1$.

Dessa forma, a RND fica limitada inferiormente pelo valor 1, em que a probabilidade é o valor máximo admissível 0,5, e determina-se o limite superior através da função particionante $H(z) = [F_B^0]^{-1}(1 - F_B^1(z))$ aplicada no ponto $z = 1$, obtendo-se o par $(z, H(z))$. Dessa forma, tem-se o limite superior $a = H(1) = [F_B^0]^{-1}(1 - F_B^1(1))$, que é um valor à direita de b (Figura 3.2).

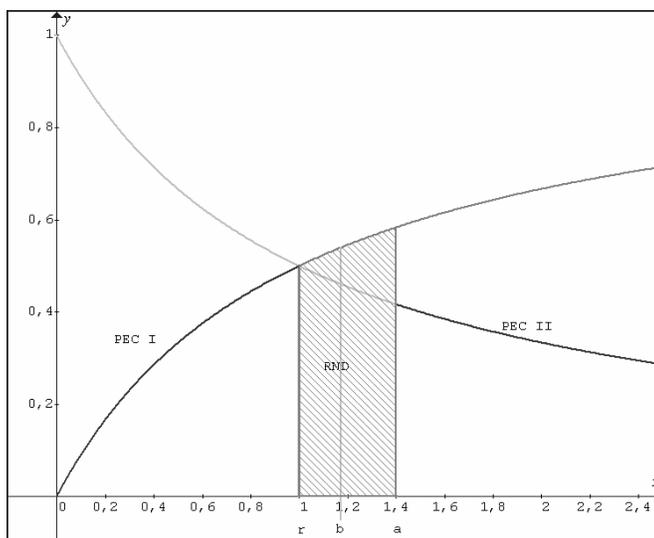


FIGURA 3.2 – Exemplo da região de não-decisão (RND) quando $b_c > 1$.

Tem-se então que $F_B^0(a) = P_{\theta_0}[B(X) \leq a] = 1 - F_B^1(1)$ (Figura 3.3).

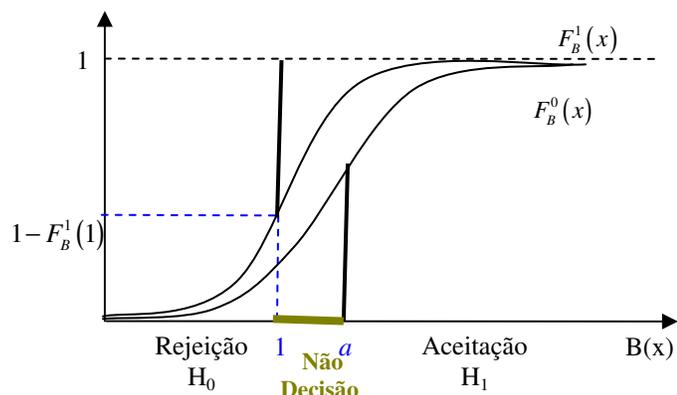


FIGURA 3.3 – Esquema da construção da região de não-decisão para $b_c > 1$ em que $F_B^0(1) \leq 1 - F_B^1(1)$

Como a derivada de $H(z) = [F_B^0]^{-1}(1 - F_B^1(z))$ é negativa, a função é decrescente e o valor de $H(1)$ é maior que $H(b_c) = b_c$ (Figura 3.4).

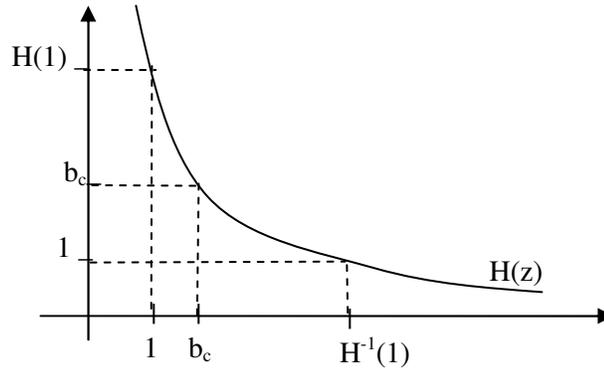


FIGURA 3.4 – Gráfico de $H(z)$ para o caso (i) em que se tem $1 < b_c < H(1)$.

2º caso) O valor crítico é inferior a 1, ou seja, $b_c < 1$.

Agora, como $b_c < 1$, tem-se $F_B^0(1) > 1 - F_B^1(1)$. Em $b = 1$ tem-se $\alpha_x = \beta_x = 1/2$; com isso, no ponto $b_c < 1$, uma destas probabilidades será superior a $1/2$. Esse fato ocorre porque as probabilidades de erro tipo I e tipo II serão diferentes em $b = 1$ (Figura 3.5).

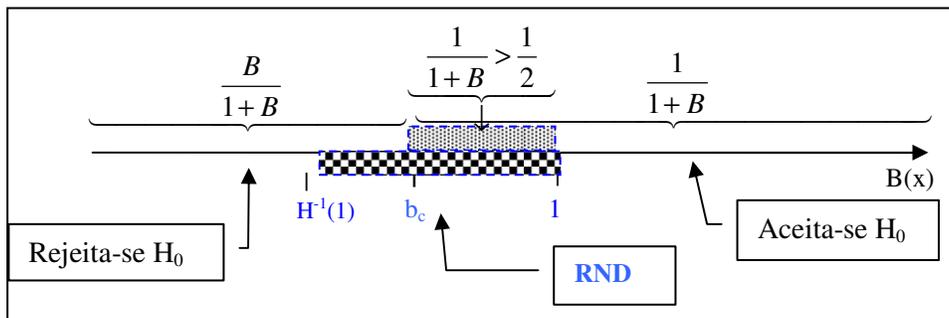


FIGURA 3.5 – Esquema da região de não-decisão para $b_c < 1$.

Logo, a RND fica limitada superiormente pelo ponto $b_c = 1$ e determina-se o limite inferior através da função particionante $r = H^{-1}(1) = [F_B^1]^{-1}(1 - F_B^0(1))$ (Figura 3.6).

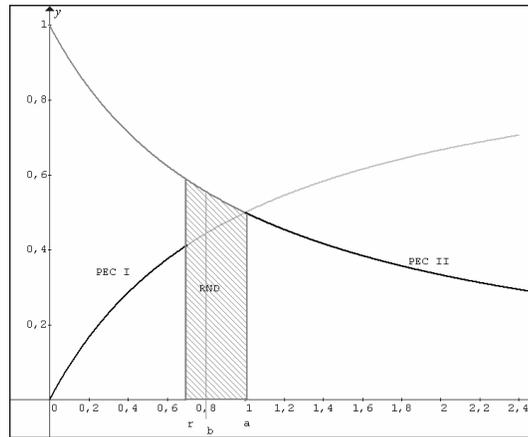


FIGURA 3.6 – Representação da região de não-decisão (RND) quando $b_c < 1$.

Tem-se, então, que $F_B^1(r) = P_{\theta_1}[B(X) \leq r] = 1 - F_B^0(1)$ (Figura 3.7).

Portanto, o teste Y_{π}^* , com a inclusão da região de não decisão definida da forma:

$$r=1 \quad \text{e} \quad a = [F_B^0]^{-1}(1 - F_B^1(1)) \quad \text{se} \quad F_B^0(1) \leq 1 - F_B^1(1)$$

$$r = [F_B^1]^{-1}(1 - F_B^0(1)) \quad \text{e} \quad a = 1 \quad \text{se} \quad F_B^0(1) > 1 - F_B^1(1)$$

Ou, considerando $H(b) = [F_B^0]^{-1}(1 - F_B^1(b))$ e

$H^{-1}(b) = [F_B^1]^{-1}(1 - F_B^0(b))$, a RND é dada por:

$$r=1 \quad \text{e} \quad a = H(1), \quad \text{se} \quad H(1) \geq 1,$$

$$r = H^{-1}(1) \quad \text{e} \quad a = 1, \quad \text{se} \quad H(1) < 1.$$

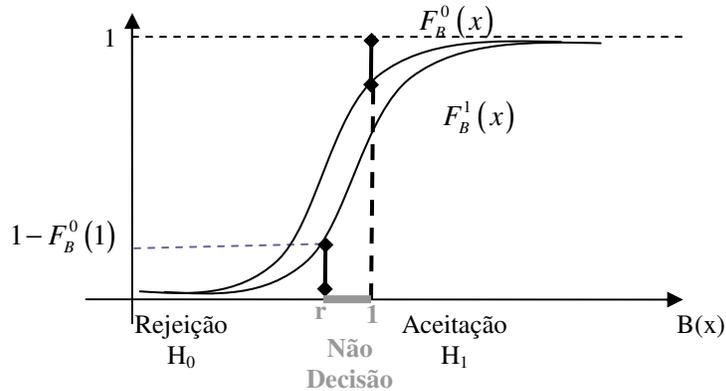


FIGURA 3.7 – Esquema da construção da região de não-decisão para $b < 1$.

E o teste fica:

$$\Upsilon_{\pi}^* = \begin{cases} \text{se } B(x) \leq r, & \text{rejeita-se } H_0 \text{ e apresenta-se PEC I} \\ & \text{ou probabilidade a posteriori } \alpha_x = \frac{B}{1+B}; \\ \text{se } r < B(x) < a, & \text{não se toma nenhuma decisão;} \\ \text{se } B(x) \geq a, & \text{aceita-se } H_0 \text{ e apresenta-se PEC II} \\ & \text{ou probabilidade a posteriori } \beta_x = \frac{1}{1+B}. \end{cases} \quad (3.1)$$

As discussões para construção da RND aqui apresentadas contemplam os testes de hipóteses que envolvem os modelos não-encaixados, isto é, modelos que apresentam formas funcionais completamente diferentes, e os modelos encaixados, sendo que, neste último caso, apenas o 2º caso em que $b_c < 1$ ocorrerá, a variável $B(X)$ assume apenas valores menores que 1 (considerando o modelo mais simples na hipótese nula).

A região de não-decisão neste teste é uma fonte de críticas. Mas a região de não-decisão é o preço que se deve pagar a fim de fazer com que as probabilidades de erro Bayesiano descritas também sejam probabilidades de

erros freqüentistas condicionais limitadas a $\frac{1}{2}$. Assim, o tamanho da região de não-decisão é uma característica particularmente importante para estudar. Deve, também, ficar claro que a região de não-decisão desaparece sempre que $F_B^0(1) = 1 - F_B^1(1)$, pois, nesse caso, $r = a = 1$. Isto ocorre no caso de razão de verossimilhanças simétrica (Berger et al., 1994).

A região de não decisão pode também evitar que um comportamento contra-intuitivo ocorra. Suponha, por exemplo, que $b = 5$. Então, Y_π rejeitaria H_0 quando $B(x) = 4$, embora $B(x) = 4$ seja interpretado (por um Bayesiano) como evidências 4 para 1 a favor de H_0 . Para um Bayesiano, não é sensato aceitar ou rejeitar quando *a posteriori* favorece a ação oposta (pelo menos quando se pressupõem probabilidades *a priori* iguais e perdas iguais).

2.1.2 Região de não decisão em testes freqüentistas não-condicionais

No contexto freqüentista tradicional, a inclusão de uma região de não-decisão ajuda a diminuir algum comportamento paradoxal de testes não-condicionais. Para verificar isto, considere dois estatísticos tradicionais (não-condicionais), A e B, que pretendem, baseados nas mesmas observações x 's, construir um teste mais poderoso de tamanho α para testar duas hipóteses simples, $X \sim f_0(x)$ versus $X \sim f_1(x)$. Além disso, suponha que ambos os estatísticos são indiferentes quanto à escolha de qual f. d. p. deve ser considerada na hipótese nula.

O estatístico A escolhe que as hipóteses sejam

$$H_0^A: X \sim f_0(x) \quad \text{versus} \quad H_1^A: X \sim f_1(x)$$

e constrói um teste mais poderoso de tamanho α como

$$\begin{aligned} \text{se } B(x) \leq c_0, & \quad \text{rejeita-se } H_0^A, \\ \text{se } B(x) > c_0, & \quad \text{aceita-se } H_0^A, \end{aligned}$$

em que o valor crítico c_0 é determinado pela equação $F_B^0(c_0) = \alpha$.

O estatístico B escolhe que as hipóteses sejam

$$H_0^B: X \sim f_1(x) \quad \text{versus} \quad H_1^B: X \sim f_0(x)$$

e constrói um teste mais poderoso de tamanho α como

$$\begin{aligned} \text{se } B(x) \geq c_1, & \quad \text{rejeita-se } H_0^B, \\ \text{se } B(x) < c_1, & \quad \text{aceita-se } H_0^B, \end{aligned}$$

em que, nesse caso, o valor crítico c_1 é determinado pela equação $1 - F_B^1(c_1) = \alpha$. Em ambos os casos, $B(x) = f_0(x)/f_1(x)$.

A dificuldade surge sempre que $c_0 \neq c_1$, ou seja, nesse caso o conjunto $\{x: \min(c_0, c_1) < B(x) < \max(c_0, c_1)\}$ não é vazio. Este é o conjunto de discordância entre os dois estatísticos, os quais chegarão a diferentes conclusões. Este conflito pode ser solucionado. Contudo, deve-se estar disposto a modificar o teste tradicional para incorporar a possibilidade de não-decisão. Com isso em mente, seja $r_0 = \min(c_0, c_1)$ e $a_0 = \max(c_0, c_1)$; então a modificação do teste tradicional para as hipóteses simples, que inclui uma região de não-decisão, é:

$$Y': \begin{cases} \text{se } B(x) \leq r_0, & \text{rejeita-se } H_0 \\ \text{se } r_0 < B(x) < a_0, & \text{não se toma decisão alguma,} \\ \text{se } B(x) \geq a_0, & \text{aceita-se } H_0. \end{cases}$$

Outra maneira de explicitar esse fato é a de que, se o desejo tratar f_0 e f_1 com mesma importância, com probabilidades de erro tipo I e Tipo II iguais a α especificado, então a introdução de uma região de não-decisão é necessária.

Serão apresentados dois exemplos com a construção da RND para o caso de um teste frequentista tradicional, o primeiro exemplo tratando de teste de hipótese para a média de uma população normal e o segundo exemplo testando o parâmetro de uma distribuição exponencial.

EXEMPLO 5: Caso de Razão de Verossimilhanças Simétrica

Suponha que X_1, X_2, \dots, X_4 , são i.i.d. $N(0,1)$, e que se deseja testar:

Estatístico A: $H_0: X \sim N(0,1)$ versus $H_1: X \sim N(1,1)$

A média k_A que limita a região crítica considerando $\alpha = 0,05$ é dada por:

$$\frac{k_A - \mu_A}{\sigma_{\bar{x}}} = 1,64 \Rightarrow k_A = \mu_A + 1,64\sigma_{\bar{x}} \Rightarrow k_A = 0 + \left(1,64 \times \frac{1}{\sqrt{4}}\right) = 0,82.$$

O teste mais poderoso de tamanho α é dado por

$$\begin{cases} \text{se } \bar{X}_4 \geq 0,82, & \text{rejeita-se } H_0^A, \\ \text{se } \bar{X}_4 < 0,82, & \text{aceita-se } H_0^A, \end{cases}$$

Em termos de razão de verossimilhanças, tem-se que:

$$B = \prod_{i=1}^n \frac{f_0}{f_1} = \prod_{i=1}^n \frac{(2\pi)^{-1/2} \exp\left\{-\frac{1}{2}(x_i - 0)^2\right\}}{(2\pi)^{-1/2} \exp\left\{-\frac{1}{2}(x_i - 1)^2\right\}} = \exp\left\{-n\bar{x} + \frac{n}{2}\right\}$$

Utilizando o valor crítico $k_A = 0,82$:

$$B = \exp\left\{-nk_A + \frac{n}{2}\right\} = \exp\{-4k_A + 2\} = \exp\{-4(0,82) + 2\} \Rightarrow c_0 = 0,278$$

Estatístico B: $H_0: X \sim N(1,1)$ versus $H_1: X \sim N(0,1)$

$$k_B = \mu_B - 1,64\sigma_{\bar{x}} \Rightarrow k = 1 - \left(1,64 \times \frac{1}{\sqrt{4}}\right) = 0,18.$$

Um teste mais poderoso de tamanho α é dado por:

$$\begin{cases} \text{se } \bar{X}_4 \leq 0,18, & \text{rejeita-se } H_0^B \\ \text{se } \bar{X}_4 > 0,18, & \text{aceita-se } H_0^B \end{cases}$$

Em termos de razão de verossimilhanças, tem-se que:

$$B = \exp\{-4k_B + 2\} = \exp\{-4(0,18) + 2\} \Rightarrow c_1 = 3,597$$

Note que $c_0 \neq c_1$, ou seja, o conjunto $\{x: \min(c_0, c_1) < B(x) < \max(c_0, c_1)\}$ não é vazio (Figura 3.8 e 3.9).

$$Y_x^* = \begin{cases} \text{se } x \leq 0,18 \text{ ou } B(x) \leq 0,278, & \text{rejeita-se } H_0; \\ \text{se } 0,18 < x < 0,82 \text{ ou } 0,278 < B(x) < 3,597, & \text{não se toma nenhuma decisão;} \\ \text{se } x \geq 0,82 \text{ ou } B(x) \geq 3,597, & \text{aceita-se } H_0. \end{cases}$$

É importante ressaltar que a região de não-decisão neste caso é muito grande, pois o tamanho da amostra é pequeno ($n = 4$) e os erros tipo I e II são iguais e pequenos.

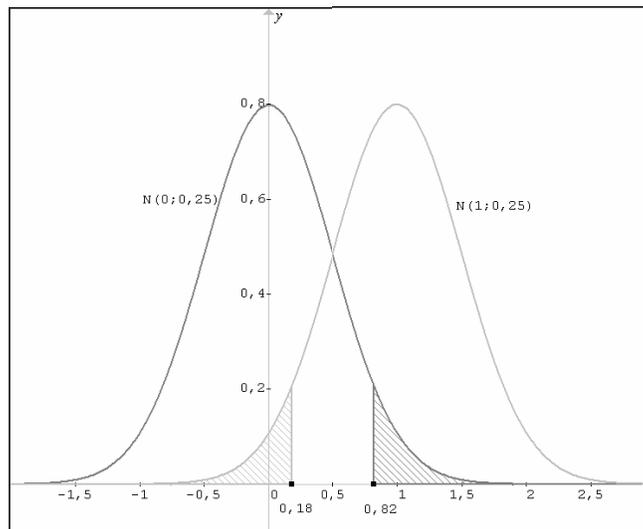


FIGURA 3.8 - Limites da região de não-decisão em função da densidade de probabilidade.

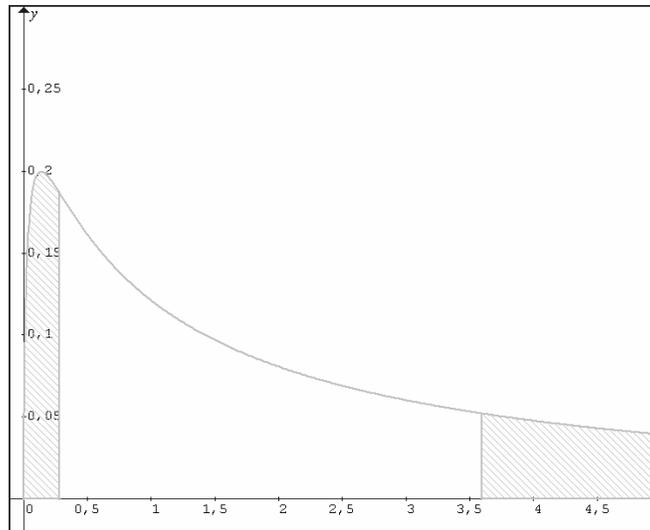


FIGURA 3.9 - Limites da região de não-decisão em função da razão de verossimilhanças b.

EXEMPLO 6: Caso de Razão de Verossimilhanças Não-Simétrica

Serão feitos os cálculos para incluir a RND, no exemplo 1(d) do capítulo anterior, de se testar $H_0: \theta = 1$ versus $H_1: \theta = 2$, da exponencial $f(x; \theta) = \theta e^{-\theta x}$, com tamanho de amostra $n = 4$. Fixando $\alpha = 0,05$, o valor crítico para a região de rejeição é de $k_1 = 1,367$ e a probabilidade de erro tipo II é igual a 0,707 (Figura 3.10).

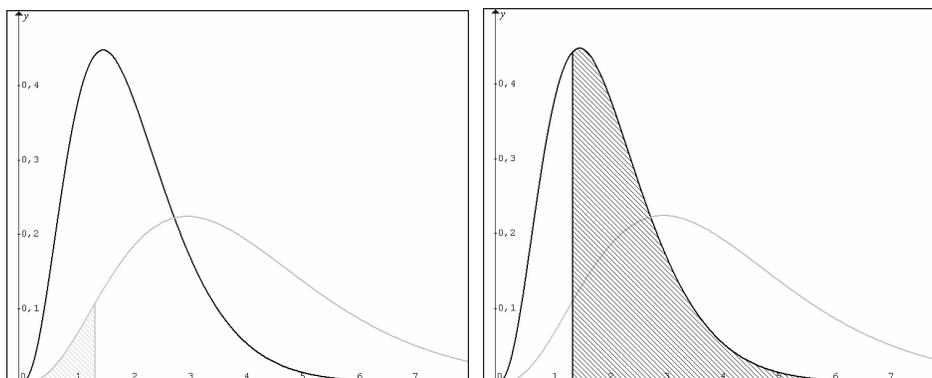


FIGURA 3.10 – Probabilidades de erros Tipo I e II considerando o teste de $H_0: \theta = 1$ versus $H_1: \theta = 2$ da exponencial com valor crítico $k_1=1,367$.

Sob a desigualdade equivalente

$$B(x_1, \dots, x_n) = \left(\frac{\theta_0}{\theta_1}\right)^n \exp\left\{-\left(\theta_0 - \theta_1\right) \sum_{i=1}^n x_i\right\} =$$

$$= \left(\frac{1}{2}\right)^4 \exp\{-(1-2)1,367\} = 0,0625 \exp(1,367) = 0,245$$

Se as hipóteses fossem invertidas, ou seja, $\theta_0 = 2$ e $\theta_1 = 1$.

Sob H_0 , $\sum_{i=1}^4 X_i \sim \text{Gama}(4, 2)$, então

$$\alpha = 1 - P_{\theta_0} \left[\sum_{i=1}^4 X_i \leq k \right] = 1 - \int_0^k \frac{1}{\Gamma(4)} y^3 2^4 e^{-2y} dy = 0,05 \Rightarrow k = 3,877.$$

Que corresponde a uma probabilidade de erro tipo II dada por:

$$\beta = P_{\theta_1} \left[\sum_{i=1}^4 X_i < 3,877 \right] = \int_0^{0,3877} \frac{1}{\Gamma(4)} y^3 e^{-y} dy \Rightarrow \boxed{\beta = 54,21} \text{ (Figura 3.11).}$$

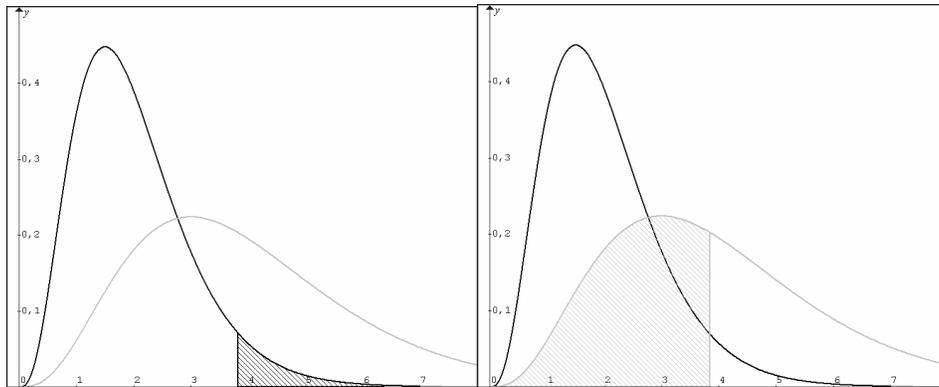


FIGURA 3.11 – Probabilidades de erros Tipo I e II considerando o teste de $H_0: \theta = 2$ versus $H_1: \theta = 1$ da exponencial com valor crítico $k_1=3,877$.

Sob a desigualdade equivalente e utilizando o valor crítico k_1 , tem-se (Figura 3.12):

$$B(x_1, \dots, x_n) = \left(\frac{\theta_0}{\theta_1} \right)^n \exp \left\{ -(\theta_0 - \theta_1) \sum_{i=1}^n x_i \right\} = (2)^{-4} \exp \{3,877\} = 3,017$$

$$\sum_{i=1}^4 x_i \leq \ln [0,245(2)^4] \Rightarrow \sum_{i=1}^4 x_i \leq 1,366 \text{ e } \sum_{i=1}^4 x_i \geq \ln [3,017(2)^4] \Rightarrow \sum_{i=1}^4 x_i \geq 3,877$$

$$\Upsilon': \begin{cases} \text{se } \sum_{i=1}^4 X_i \leq 1,367 \text{ ou } B(x) \leq 0,245, \text{ rejeita-se } H_0 \\ \text{se } 1,367 < \sum_{i=1}^4 X_i < 3,877 \text{ ou } 0,245 < B(x) < 3,017 \\ \quad \quad \quad \text{n\~{a}o se toma decis\~{a}o alguma,} \\ \text{se } \sum_{i=1}^4 X_i \geq 3,877 \text{ ou } B(x) \geq 3,017, \text{ aceita-se } H_0 . \end{cases}$$

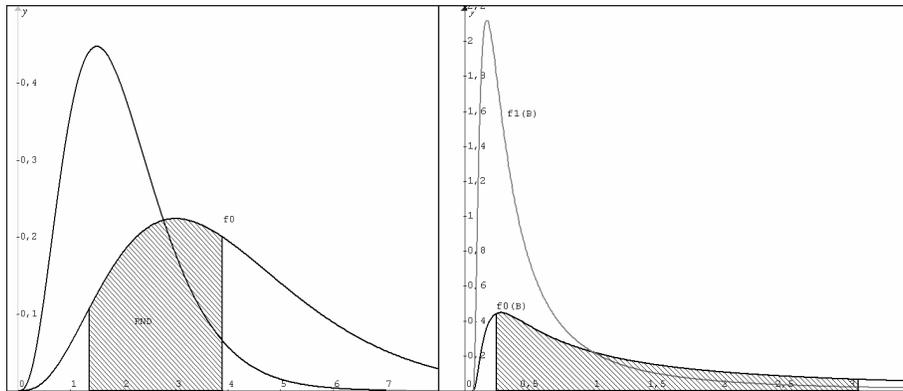


FIGURA 3.12 –Região de não decisão para o teste de $H_0: \theta = 1$ versus $H_1: \theta = 2$.

2.2. GENERALIZAÇÃO DO TESTE UNIFICADO

Considerando a razão entre as perdas $l = \frac{l_1}{l_0}$ e a razão entre as densidades

a priori $\rho = \frac{\pi_1}{\pi_0}$, pode-se generalizar o teste Bayesiano Υ_π da seguinte forma:

$$\Upsilon_{l,\rho} = \begin{cases} \text{se } B(x) \leq \rho l, & \text{rejeita-se } H_0 \text{ com probabilidade a posteriori} \\ & f(\theta_0 | x) = \frac{B(x)}{B(x) + \rho}; \\ \text{se } B(x) > \rho l, & \text{aceita-se } H_0 \text{ com probabilidade a posteriori} \\ & f(\theta_1 | x) = \frac{\rho}{B(x) + \rho}. \end{cases}$$

Pode-se, ainda, generalizar o teste unificado Υ_{π}^* , considerando uma região de não-decisão, com limites $r_{l,\rho}$ e $a_{l,\rho}$, para $0 < l < \infty$ e $0 < \rho < \infty$, definidos como se segue

$$r_{l,\rho} = l\rho \quad \text{e} \quad a_{l,\rho} = [F_B^0]^{-1}(1 - \rho F_B^1(l\rho)) \quad \text{se} \quad F_B^0(l\rho) \leq \rho - F_B^1(l\rho);$$

$$r_{l,\rho} = [F_B^1]^{-1}((1/\rho)[1 - F_B^0(l\rho)]) \quad \text{e} \quad a_{l,\rho} = l\rho \quad \text{se} \quad F_B^0(l\rho) > \rho - F_B^1(l\rho)$$

Então, a generalização do teste Υ_{π}^* , dado pela expressão (3.1), denominado agora de $\Upsilon_{l,\rho}$, fica:

$$\Upsilon_{l,\rho} = \begin{cases} \text{se } B(x) \leq r_{l,\rho}, \text{ rejeita-se } H_0 \text{ com probabilidade a posteriori} \\ \qquad \qquad \qquad f(\theta_0 | x) = \alpha_x = \frac{B(x)}{B(x) + \rho}; \\ \text{se } r_{l,\rho} < B(x) < a_{l,\rho}, \text{ não se toma nenhuma decisão;} \\ \text{se } B(x) \geq a_{l,\rho}, \text{ aceita-se } H_0 \text{ com probabilidade a posteriori} \\ \qquad \qquad \qquad f(\theta_1 | x) = \beta_x = \frac{\rho}{B(x) + \rho}. \end{cases} \quad (3.2)$$

Os valores para os limites da RND seguem o mesmo padrão de construção do teste Υ_{π}^* , sendo que, se $F_B^0(l\rho) \leq \rho - F_B^1(l\rho)$, então $r = l\rho$ e $a_{l,\rho} = [F_B^0]^{-1}(1 - \rho F_B^1(l\rho))$. Então, $F_B^0(a) = P_{\theta_0}[B(X) \leq a] = 1 - \rho F_B^1(l\rho)$ (Figura 3.13 e 3.14).

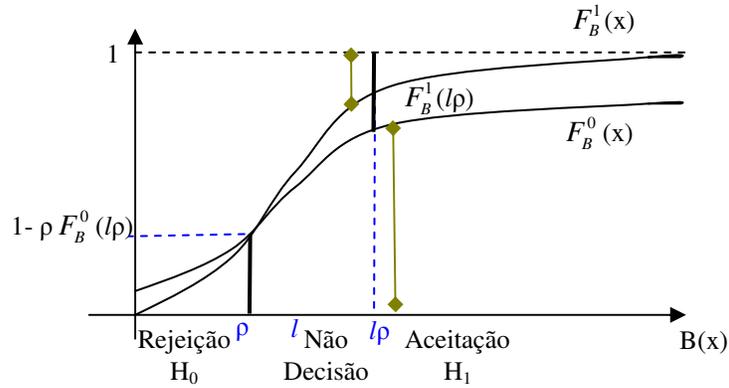


FIGURA 3.13 – Esquema da construção da região de não-decisão no teste $\Upsilon_{l,\rho}$ para $b_c > l\rho$.

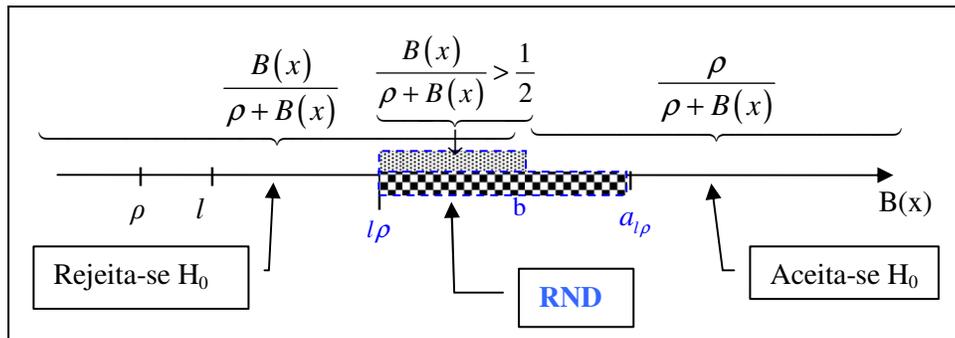


FIGURA 3.14 - Esquema da Construção da região de não-decisão para $b_c > l\rho$.

Para $b_c < l\rho$ a RND, ou seja, quando $F_B^0(l\rho) > \rho - F_B^1(l\rho)$, têm-se os limites $a = l\rho$ e $r_{l,\rho} = [F_B^1]^{-1}\left(\frac{1}{\rho}(1 - F_B^0(l\rho))\right)$, de forma que $F_B^1(r) = P_{\theta_1}[B(X) \leq r] = \frac{1}{\rho}(1 - F_B^0(l\rho))$ (Figura 3.15).

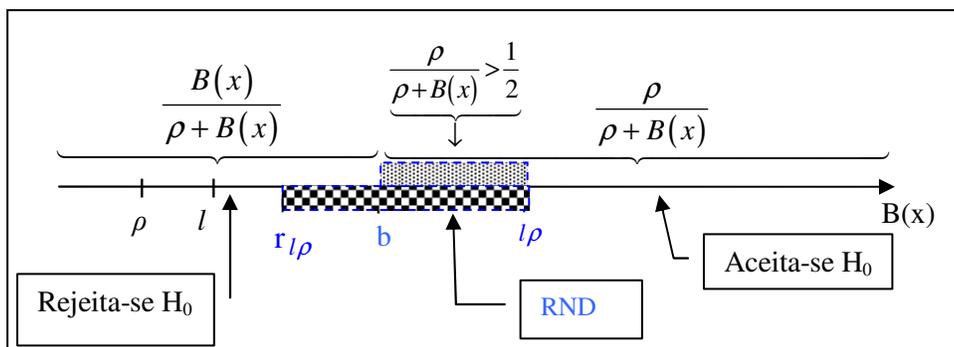


FIGURA 3.15 – Esquema da Região de não-decisão para $b_c > l\rho$.

O teste $\Upsilon_{l\rho}$ é o próprio teste Bayesiano Υ_π , exceto pela presença da região de “não-decisão” $(r_{l,\rho}, a_{l,\rho})$. (Claramente, multiplicando α_ρ por l_0 e β_ρ por l_1 , converte-se a probabilidade de erro condicional em $\Upsilon_{l\rho}$ para perda esperada *a posteriori*).

Segundo Berger et al. (1994), um freqüentista rigoroso pode também usar Υ_π efetivamente. Suponha que um freqüentista tenha planejado usar o teste Neyman-Pearson não-condicional com valor crítico b_c e probabilidade de erro $\alpha = F_B^0(b_c)$ e $\beta = (1 - F_B^1(b_c))$. Este pode ser substituído pelo teste $\Upsilon_{\ell^*\rho^*}$, em que

$$\rho^* = \frac{(1-\alpha)}{(1-\beta)} \quad \text{e} \quad \ell^* = \frac{b_c}{\rho^*}.$$

Verifica-se diretamente que $\Upsilon_{\ell^*\rho^*}$ ainda tem b_c como um valor crítico, e não possui uma região de não-decisão. Com $\rho^* = \frac{(1-\alpha)}{(1-\beta)}$ tem-se, para os limites

da RND:

$$a_{\ell^*\rho^*} = [F_B^0]^{-1} [1 - \rho^* F_B^1(b)] = [F_B^0]^{-1} \left[1 - \frac{(1-\alpha)}{(1-\beta)} F_B^1(b) \right] =$$

$$\begin{aligned}
&= [F_B^0]^{-1} \left[1 - \frac{(1 - F_B^0(b))}{F_B^1(b)} F_B^1(b) \right] = b = b_c \\
r_{l^*, \rho^*} &= [F_B^1]^{-1} \left[\frac{1}{\rho^*} (1 - F_B^0(b)) \right] = [F_B^1]^{-1} \left[\frac{(1 - \beta)}{(1 - \alpha)} (1 - F_B^0(b)) \right] = \\
&= [F_B^1]^{-1} \left[\frac{F_B^1(b)}{(1 - F_B^0(b))} (1 - F_B^0(b)) \right] = b = b_c
\end{aligned}$$

Note que a escolha de l^* e ρ^* é vista aqui justamente como um formalismo; suas interpretações em termos de perdas e densidades *a priori* não são necessárias (Berger et al., 1994).

2.2.1 INTERPRETAÇÃO FREQUËNTISTA CONDICIONAL DO TESTE UNIFICADO GENERALIZADO

Para validar o teste unificado $\Upsilon_{l,\rho}$, no contexto de testes freqüentistas condicionais (discutido em detalhes no capítulo 2 deste trabalho) deve-se estabelecer uma função particionante $H(z)$ com boas propriedades e que forneça como probabilidades de erros condicionais tipo I e II, os mesmos valores apresentados pelo teste $\Upsilon_{l,\rho}$ como probabilidades *a posteriori*.

Uma opção para a função $H(z)$ é dada por $H(z) = [F_B^0]^{-1} (1 - \rho F_B^1(z))$ para $z \geq a_{l,\rho}$ e $H(z) = [F_B^1]^{-1} \left(\frac{1}{\rho} (1 - F_B^0(z)) \right)$ para $z \leq r_{l,\rho}$ (Figura 3.16).

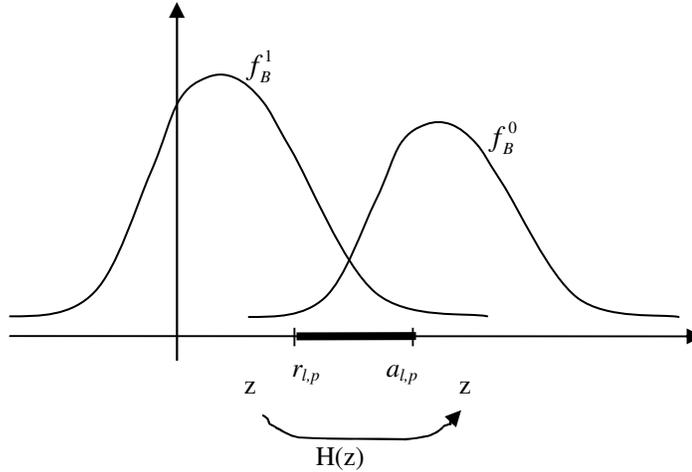


FIGURA 3.16 – Comportamento da função $H(z)$.

É importante ressaltar que a função particionante $H(z) = F_0^{-1}(1 - \rho F_1(z))$ possui como inversa a função particionante

$H^{-1}(z) = F_1^{-1}\left(\frac{1}{\rho}(1 - F_0(z))\right)$, pois:

$$\begin{aligned} H(H^{-1}(z)) &= F_0^{-1}\left(1 - \rho F_1\left(F_1^{-1}\left(\frac{1}{\rho}(1 - F_0(z))\right)\right)\right) = \\ &= F_0^{-1}\left(1 - \rho\left(\frac{1}{\rho}(1 - F_0(z))\right)\right) = F_0^{-1}(1 - (1 - F_0(z))) = z \end{aligned}$$

e

$$\begin{aligned} H^{-1}(H(z)) &= F_1^{-1}\left(\frac{1}{\rho}(1 - F_0(F_0^{-1}(1 - \rho F_1(z))))\right) = \\ &= F_1^{-1}\left(\frac{1}{\rho}(1 - (1 - \rho F_1(z)))\right) = F_1^{-1}\left(\frac{1}{\rho}(\rho F_1(z))\right) = z \end{aligned}$$

TEOREMA

Para o teste $\Upsilon_{l,\rho}$ com partição definida por $H(z)$, se $z > 0$,

$$\alpha(z) = P_0(\text{rejeita-se } H_0 \mid (z \text{ ou } H(z) \text{ foi observado})) = \alpha_x = \frac{B(x)}{\rho + B(x)},$$

$$\beta(z) = P_1(\text{aceita-se } H_0 \mid (z \text{ ou } H(z) \text{ foi observado})) = \beta_x = \frac{\rho}{\rho + B(x)}.$$

PROVA

Seja $f_B^i(\cdot)$ a densidade da estatística suficiente $B(X)$ sob f_i , $i = 1, 2$.

Tem-se que, pelo teorema apresentado no capítulo 2, página 91,

$$f_B^0(b) = b f_B^1(b).$$

Com a função particionante $H(z)$ e utilizando a composição

$$F_B^0(H(z)) = F_B^0 \circ [F_B^0]^{-1}(1 - \rho F_B^1(z)) = 1 - \rho F_B^1(z)$$

Derivando em relação a z , tem-se:

$$F_B^{0'}(H(z)) \cdot H'(z) = -\rho F_B^{1'}(z)$$

$$f_B^0(H(z)) \cdot H'(z) = -\rho f_B^1(z)$$

$$\Rightarrow \boxed{H'(z) = \frac{-\rho f_B^1(z)}{f_B^0(H(z))}}$$

Substituindo na expressão freqüentista para o erro condicional tipo I, obtém-se:

$$\alpha(z) = P_0(\text{rejeitar } H_0 \mid z \text{ ou } H(z) \text{ foi observado}) = \frac{f_B^0(z)}{f_B^0(z) + f_B^0(H(z)) \left| \frac{d}{dz} H(z) \right|} \Rightarrow$$

$$\alpha(z) = \frac{f_B^0(z)}{f_B^0(z) + \rho f_B^1(z)} = \frac{\frac{f_B^0(z)}{f_B^1(z)}}{\frac{f_B^0(z)}{f_B^1(z)} + \rho} = \frac{B(x)}{B(x) + \rho}$$

Para a ocorrência de $\beta(z)$ tem-se $a_{l,\rho} < z \leq \infty$ e, agora, $r_{l,\rho} = H(a_{l,\rho})$, e a função particionante $H(z)$ para a construção da partição dada por:

$$H(z) = F_1^{-1}\left(\frac{1}{\rho}(1 - F_0(z))\right)$$

Pois, partindo da composição:

$$F_B^1(H(z)) = F_B^1 \circ [F_B^1]^{-1}\left(\frac{1}{\rho}(1 - F_B^0(z))\right) = \frac{1}{\rho}(1 - F_B^0(z))$$

e derivando em relação a z , tem-se:

$$F_B^{1'}(H(z))H'(z) = -\frac{1}{\rho}F_B^{0'}(z)$$

$$f_B^1(H(z))H'(z) = -\frac{1}{\rho}f_B^0(z) \Rightarrow \boxed{H'(z) = -\frac{f_B^0(z)}{\rho f_B^1(H(z))}}$$

Substituindo na expressão frequentista para o erro condicional tipo II, obtém-se:

$$\beta(z) = P_1(\text{aceita-se } H_0 \mid z \text{ ou } H(z) \text{ foi observado}) = \frac{f_B^1(z)}{f_B^1(z) + f_B^1(H(z)) \left| \frac{d}{dz} H(z) \right|} \Rightarrow$$

$$\beta(z) = \frac{f_B^1(z)}{f_B^1(z) + \frac{f_B^0(z)}{\rho}} = \frac{\frac{f_B^1(z)}{f_B^1(z)}}{\frac{f_B^1(z)}{f_B^1(z)} + \frac{f_B^0(z)}{\rho f_B^1(z)}} = \frac{1}{1 + \frac{B}{\rho}} = \frac{\rho}{\rho + B}$$

■

2.3 COMENTÁRIOS SOBRE O TESTE UNIFICADO

Os testes $\Upsilon_{l,\rho}$ possuem, simultaneamente, uma justificativa Bayesiana e uma freqüentista condicional, com a decisão e a medida de evidência (probabilidade de erro condicional) apresentada sendo numericamente a mesma sob ambos os paradigmas. É claro que a interpretação da medida de evidência dada pelo teste diferirá para Bayesianos e freqüentistas, mas tal fato será irrelevante na prática. Aqueles que vêm mérito nas filosofias Bayesiana e freqüentista certamente ficarão satisfeitos com esta interpretação dual.

Freqüentistas rigorosos poderiam reclamar que, na prática, $\Upsilon_{l,\rho}$ não fornece toda a informação necessária. O dogma Freqüentista afirma que não se pode apenas, dizer, mas sim olhar para a probabilidade de erro tipo I ao rejeitar, mas também diz que se deve olhar para erro tipo II. No contexto freqüentista condicional, este dogma implicaria que devem ser informados ambos os erros $\alpha(z)$ e $\beta(z)$, tanto ao rejeitar quanto ao aceitar. No entanto, o procedimento dado por $\Upsilon_{l,\rho}$ só pode apresentar um dos tipos de erros, uma vez que estes ocorrem em pontos distintos (Berger et al., 1994).

Há várias possíveis soluções para este dilema. Primeiro porque a forma do teste está definida com os dois tipos de erros e, caso fosse necessário, as duas probabilidades de erro poderiam ser calculadas. Na prática, porém, calcular, por exemplo, a PEC II, $\beta(z)$, ao rejeitar H_0 não teria utilidade. Segundo, pode ser argumentado que a origem do dogma freqüentista sobre informar ambas as probabilidades de erro se apóia em uma tentativa para intuitivamente compensar o fato de não condicionar; tal compensação é claramente não necessária para $\Upsilon_{l,\rho}$. Finalmente, na “situação simétrica”, $\alpha_z = \beta_z$ para todo z ; assim, o problema desaparece (Berger et al., 1994).

O principal problema possível no uso de $\Upsilon_{l,\rho}$ é a presença da "região de não-decisão". De certo modo, esta é a região na qual não se pode obter concordância entre Bayesianos e freqüentistas. Se esta região é muito grande, então a utilidade do teste $\Upsilon_{l,\rho}$ é reduzida (lembrando ainda que um freqüentista que está menos interessado sobre um acordo com Bayesianos pode apenas usar Υ_{π} ou Υ_{l^*,ρ^*} nenhum dos quais possui uma "região de não-decisão").

O tamanho da RND é uma característica particularmente importante para ser estudada, pois desaparece sempre que $F_B^0(b_c) = 1 - F_B^1(b_c)$, em que F_B^0 e F_B^1 são as funções de densidade acumuladas da razão de verossimilhanças $B(x)$ sob f_0 e f_1 e b_c é o valor crítico do teste. Nesse caso, $r = a = 1$. Isso ocorre no caso de razão de verossimilhança simétrica.

Um fato encorajador é que se pode mostrar, em geral, que para qualquer l ou qualquer ρ , uma solução na outra variável pode ser encontrada de modo que $r_{l,\rho} = a_{l,\rho} = l\rho$, situação em que a "região de não-decisão" é vazia. Para isso pode-se mostrar, também, que as soluções em l e ρ estão inversamente relacionadas, pois, neste caso, a condição de igualdade entre os "erros" apresentados nos testes Bayesianos e freqüentistas condicionais deve ser satisfeita. Então:

$$\left\{ \begin{array}{l} \frac{B(X)}{1+B(X)} = l_0 \left(\frac{B(X)}{\rho+B(X)} \right) \\ \frac{1}{1+B(X)} = l_1 \left(\frac{\rho}{\rho+B(X)} \right) \end{array} \right. \Rightarrow$$

$$\frac{l_0}{B(X)} \left(\frac{B(X)}{\rho+B(X)} \right) = l_1 \left(\frac{\rho}{\rho+B(X)} \right) \Rightarrow \frac{l_0}{l_1} = \rho \Rightarrow \boxed{\frac{1}{l} = \rho}$$

■

Berger et al. (1997) comentam que o procedimento proposto de teste unificado é muito diferente do que é tipicamente usado na prática e que uma séria mudança na metodologia estatística prática deveria ocorrer com base nos princípios de condicionamento. O teste unificado propõe um método que evita que se interpretem erroneamente *valores-p* como probabilidades *a posteriori*.

2.2.1 CONSIDERAÇÕES GERAIS

O artigo publicado por Berger et al. em 1994 teve uma forte repercussão na comunidade estatística, mas os comentários sobre o mesmo só foram publicados juntamente com o segundo artigo de Berger et al., em 1997. Nesse artigo foram publicados comentários e críticas de Dennis V. Lindley, Thomas A. Louis e David Hinkley, principalmente sobre a inclusão da RND. A proposta de Berger et al. (1994) de um teste unificado é ainda polêmica e traz à tona várias discussões, mas o próprio Berger deu continuidade a este trabalho juntamente com outros profissionais da área e em seus trabalhos de orientação, entre os quais aqui alguns trabalhos, não utilizados na íntegra neste capítulo, merecem destaque:

- Berger, J. O., Boukai, B. e Wang, Y. **Properties of unified bayesian-frequentist tests**. In: *Advances in Statistical Decision Theory and Applications*. Birkhauser, Boston, 207-223. 1997a.
- Berger, J.O., Boukai, B. and Wang, W. (1997). Unified frequentist and Bayesian testing of a precise hypothesis. *Statistical Sciences*, 12, 133-160 (com discussão).

Fazendo a mesma abordagem do artigo de 1994 (fundamentação do capítulo 2 deste trabalho), mas com hipótese precisa.

- Dass, S.C. and Berger, J.O. (1998). Unified Bayesian and conditional frequentist testing of composite hypotheses. *Institute of Statistics and Decision Sciences Discussion Paper, Duke University, USA*, 43-98.

Continuidade do trabalho de teste unificado com discussões sobre o problema da hipótese precisa.

- Berger, J.O., Boukai, B. and Wang. Y. (1999). Simultaneous Bayesian-frequentist sequential testing of nested hypotheses. *Biometrika*, 86, 79-92.

Trabalha novamente com hipótese precisa, mas agora no contexto de análise seqüencial, mostrando como a região de não-decisão pode ser interpretada naturalmente nesta situação.

- Sarat C. Dass. Unified bayesian and conditional frequentista testing for discrete distributions. *Sankhya : The Indian Journal of Statistics*. 2001, Volume 63, Series B, 3.

Esse artigo também é parte de trabalho de orientação de doutorado com a mesma situação de unificação considerando distribuições discretas, ressaltando, com alguns exemplos, qual seria o comportamento da região de não-decisão em casos de “fronteira”.

- BERGER James O.; GUGLIELMI Alessandra. Bayesian and conditional frequentist testing of a parametric model versus non parametric alternatives. *Journal of the American Statistical Association (J. Am. Stat. Assoc.)*, 2001, vol. 96, n°453, pp. 174-184.

Outra orientação envolvendo o teste unificado.

- Paulo, R. M. B. Problems on the Bayesian/Frequentist Interface. (Tese de Doutorado - orientação de James Berger.). 2002, Duke University.
- Paulo, R. M. B. Conditional Frequentist Sequential Tests for the Drift of Brownian Motion. 2002, Duke University.

Esse artigo faz parte da Tese de Ph. D. na Duke University, em 2002, com orientação do Prof. James Berger.

- James O. Berger. Could Fisher, Jeffreys and Neyman Have Agreed on Testing? *Statistical Science*. 2003, Vol. 18, No. 1, 1–32. Institute of Mathematical Statistics, 2003.

Artigo em que Berger expõe de maneira mais “filosófica” as vantagens do teste unificado, comentando um pouco das diferentes escolas. Resultado de uma palestra ministrada pelo autor durante o VIII LATIN AMERICAN CONGRESS OF PROBABILITY AND MATHEMATICAL STATISTICS (C.L.A.P.E.M.) La Habana, Cuba, em Novembro de 2001.

- *SARAT C. DASS & JAMES O. BERGER*. Unified Conditional Frequentist and Bayesian Testing of Composite Hypotheses. *Scandinavian Journal of Statistics*. Volume 30 Issue 1, Page 193 - March 2003.

Novamente outro artigo, fruto de orientação envolvendo o teste unificado.

- M. J. Bayarri and J. O. Berger. The Interplay of Bayesian and Frequentist Analysis. *Statist. Sci.* 19, no. 1 (2004), 58–80.

3 METODOLOGIA

Para ilustrar a construção da "região de não-decisão", foram considerados três exemplos. O primeiro exemplo mostra como é a construção da RND considerando o teste de hipóteses entre as distribuições Beta (1/2, 1) e Uniforme (0,1) com amostra de tamanho 1, considerando a função particionante $H(z) = [F_B^0]^{-1}(1 - F_B^1(1))$. Verificou-se também qual foi o comportamento da esperança da probabilidade de erros condicionais com a condição de igualdade entre as probabilidades de erro Tipo I e Tipo II não-condicionais.

No segundo exemplo, considerou-se um teste de hipóteses entre distribuições exponenciais considerando os tamanhos amostrais: $n = 1$; $n = 4$; $n = 10$; $n = 30$ e $n = 100$. Para encontrar o valor crítico e os limites da RND em cada caso utilizou-se o software Maple 10. Nesse caso considerou-se, para o condicionante, a função particionante $H^{-1}(z) = [F_B^1]^{-1}(1 - F_B^0(1))$. Elaborou-se uma rotina no programa SAS, versão 8 (Anexo B), verificando o número de vezes em que não houve tomada de decisão. Os resultados foram também realizados pela integração da região envolvida em cada situação. Inicialmente um número aleatório uniforme $u \sim U(0,1)$ foi gerado com a função ranuni. Caso u fosse menor que $1/2$, gerava-se uma amostra de tamanho n da distribuição Exponencial (1); se u fosse maior que $1/2$, a amostra era gerada da distribuição Exponencial (2). Utilizavam-se os limitantes da região de aceitação, rejeição e não-decisão para realizar a contagem dos erros tipo I e tipo II. O processo foi repetido 100.000 vezes e a proporção de vezes em que não houve tomada de decisão foi computada.

No terceiro exemplo avaliou-se o comportamento da RND em função da razão entre as perdas $l = l_1/l_0$.

O teste com a inclusão da região de não decisão foi aplicado ao exemplo com dados reais para o teste de hipóteses de duas variedades de café, utilizando os valores obtidos no capítulo 2.

4 RESULTADOS E DISCUSSÃO

4.1 O teste unificado para as hipóteses $H_0: X \sim \text{Uniforme}(0,1)$ versus $H_1: X \sim \text{Beta}(1/2,1)$

Considerando o teste de hipóteses simples $H_0: X \sim \text{Uniforme}(0,1)$ versus $H_1: X \sim \text{Beta}(1/2,1)$, a razão de verossimilhanças de H_0 para H_1 é $B(x) = 2\sqrt{x}$, considerando $n=1$.

Do ponto de vista frequentista, o problema seria resolvido primeiramente estabelecendo-se o valor crítico do teste não-condicional quando $\alpha = \beta$, que é dado por:

$$F_0(x) = 1 - F_1(x) \Rightarrow \int_0^x dy = 1 - \int_0^x \frac{(y)^{\frac{1}{2}-1} (1-y)^{1-1}}{B(1/2,1)} dy \Rightarrow x = 1 - \frac{1}{2} \int_0^x \frac{1}{\sqrt{y}} dy \Rightarrow \\ x = 1 - \sqrt{x} \Rightarrow x = 0,382$$

Para que o teste frequentista seja unificado com o teste Bayesiano, o valor crítico $x = 0,382$ (ou $B(x) = 1,236$) deve ser também o mesmo. Isto significa, do ponto de vista Bayesiano, que o valor $l\rho$ é diferente de 1. Uma maneira de garantir a unificação do teste seria considerar ρ igual a 1 para se obter a probabilidade *a posteriori* igual à PEC, e considerar uma razão entre as perdas igual a 1,236, ou seja, $l = \frac{l_1}{l_0} = 1,236$.

O teste unificado frequentista condicional e Bayesiano é dado por (Figura 3.17):

$$Y_{\pi}^* = \begin{cases} \text{se } x \leq 0,382 \text{ ou } B(x) \leq 1,236, \text{ rejeita-se } H_0 \text{ e apresenta-se PEC I} \\ \text{ou probabilidade a posteriori } \alpha_x = \frac{2\sqrt{x}}{1+2\sqrt{x}}; \\ \text{se } x > 0,382 \text{ ou } B(x) > 1,236, \text{ aceita-se } H_0 \text{ e apresenta-se PEC II} \\ \text{ou probabilidade a posteriori } \beta_x = \frac{1}{1+2\sqrt{x}}. \end{cases}$$

Este é o teste freqüentista condicional com estatística condicionante baseada em *valores-p*.

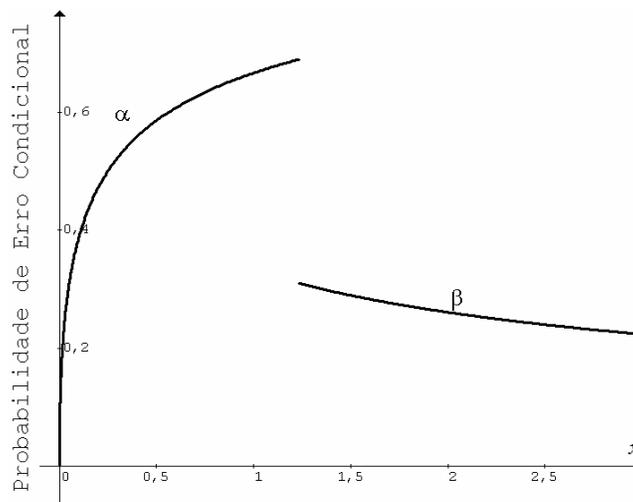


FIGURA 3.17 – Probabilidades de erro condicional associadas ao valor crítico $b_c = \rho l = 1,236$.

Note que para valores de razão de verossimilhanças no intervalo $[1; 1,236]$ (correspondente a $0,25 < x \leq 0,382$) a PEC I (e a probabilidade *a posteriori*) é superior a 0,5, mas mesmo assim ocorre a rejeição de H_0 . Então,

deve-se incluir uma RND, avaliando primeiramente se $F_B^0(1) < 1 - F_B^1(1)$ ou se $F_B^0(1) > 1 - F_B^1(1)$.

Utilizando a transformação de variáveis

$$f_B(b) = \left| \frac{d}{db} g^{-1}(b) \right| f_Y(g^{-1}(b)), \text{ tem-se:}$$

Sob H_0 :

$$g^{-1}(b) = \frac{b^2}{4} \Rightarrow \frac{d}{db} g^{-1}(b) = \frac{b}{2} \quad f_Y(g^{-1}(b)) = 1$$

$$f_B^0(b) = \left| \frac{d}{db} g^{-1}(b) \right| f_Y(g^{-1}(b)) = \frac{b}{2} \quad \text{e} \quad F_B^0(b) = P_0(x \leq b) = \int_0^b \frac{x}{2} dx = \frac{b^2}{4}$$

Sob H_1 :

$$f_B^1(b) = \left| \frac{d}{db} g^{-1}(b) \right| f_Y(g^{-1}(b)) = \frac{1}{2 \sqrt{\left(\frac{b^2}{4}\right)}} \frac{b}{2} = \frac{1}{2}$$

$$F_B^1(b) = P_1(x \leq b) = \int_0^b \frac{1}{2} dx = \frac{b}{2}$$

Daí tem-se $F_B^0(1) = 1/4$ e $F_B^1(1) = 1/2$, e como $F_B^0(1) < 1 - F_B^1(1)$, a RND é dada por:

$$r=1 \text{ e } a = (F_B^0)^{-1}(1 - F_B^1(1)) = (F_B^0)^{-1}(1/2) = 1,414.$$

O teste fica, então (Figura 3.18):

$$Y_{\pi}^* = \begin{cases} \text{se } x \leq 0,25 \text{ ou } B(x) \leq 1, \text{ rejeita-se } H_0 \text{ e apresenta-se PEC I} \\ \text{ou probabilidade a posteriori } \alpha_x = \frac{2\sqrt{x}}{1+2\sqrt{x}}; \\ \text{se } 0,25 < x < 0,5 \text{ ou } 1 < B(x) < 1,41, \text{ não se toma nenhuma decisão;} \\ \text{se } x \geq 0,5 \text{ ou } B(x) \geq 1,41, \text{ aceita-se } H_0 \text{ e apresenta-se PEC II} \\ \text{ou probabilidade a posteriori } \beta_x = \frac{1}{1+2\sqrt{x}}. \end{cases}$$

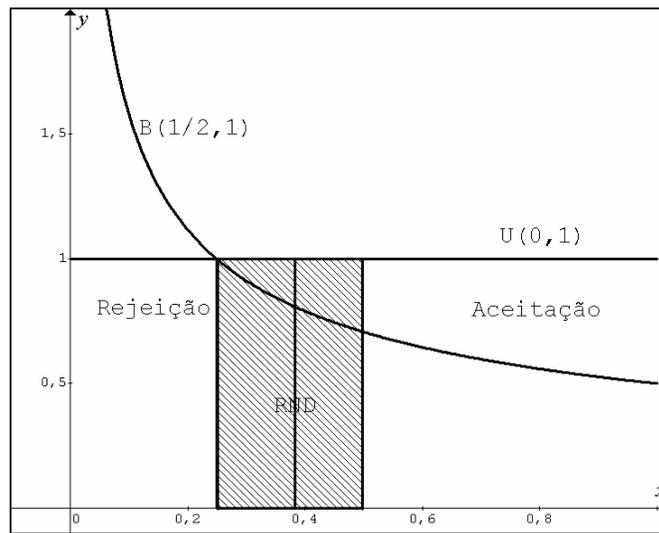


FIGURA 3.18 – Região de não-decisão para o teste unificado para as hipóteses $H_0: X \sim \text{Uniforme}(0,1)$ versus $H_1: X \sim \text{Beta}(1/2,1)$.

Os valores incluídos na RND vão além do intervalo em que a PEC é superior a 0,5, pois devem ser retirados os pontos em que ocorre este problema e, também, os pontos que formam os pares correspondentes a estes no condicionamento (Figura 3.19).

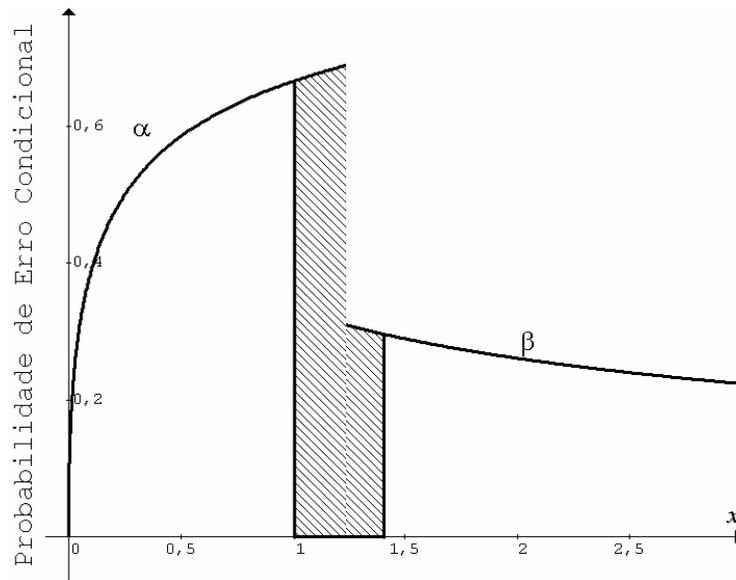


FIGURA 3.19 - Representação da Região de não-decisão em função da razão de verossimilhanças $B(x) = 2\sqrt{x}$.

Para fins de comparação foram obtidas, por simulação, 100000 amostras da distribuição Uniforme (0,1) versus Beta (1/2,1), utilizando uma rotina no programa SAS (anexo B), e avaliou-se a taxa de erro Tipo I e Tipo II para o procedimento freqüentista não condicional com e sem a inclusão da região de não-decisão (Tabela 3.1).

TABELA 3.1 – Probabilidades de erro Tipo I e II com e sem a região de não-decisão para o teste unificado para as hipóteses H_0 : $X \sim$ Uniforme (0,1) versus H_1 : $X \sim$ Beta(1/2,1)

Teste usual		Teste com região de não decisão		
[Erro Tipo I]	[Erro Tipo II]	[Erro Tipo I]	Não Decidiu	[Erro Tipo II]
38,44 %	38,28 %	25,44 %	22,64 %	29,42 %

As taxas de erros são consideravelmente diminuídas, porém o número de casos em que não se tomam decisões é igualmente elevado. Este resultado está de acordo com os valores obtidos pela integração das regiões de aceitação e rejeição quando se leva em conta a RND (Figura 3.20).

Note que as esperanças das probabilidades de erros condicionais não diminuem igualmente, apesar de inicialmente se ter $\alpha = \beta$. Como não é um caso de razão de verossimilhanças simétrica e as esperanças são relativas, ora a hipótese é nula (PEC I) ora a hipótese é alternativa (PEC II), então as novas esperanças podem variar independentemente uma da outra. Nesse caso, por exemplo, tem-se:

$$E_0[\alpha_x] = \int_0^{0,25} dx = 0,25 \quad \text{e} \quad E_1[\beta_x] = \int_{0,5}^1 \frac{1}{2\sqrt{x}} dx = \sqrt{x} \Big|_{0,5}^1 = 0,2929$$

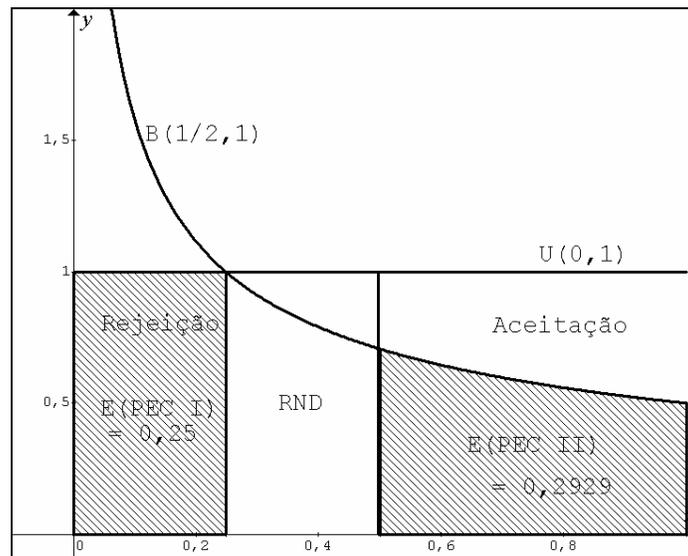


FIGURA 3.20 – Regiões de Rejeição e Aceitação modificadas pela inclusão da RND no teste unificado para as hipóteses $H_0: X \sim \text{Uniforme}(0,1)$ versus $H_1: X \sim \text{Beta}(1/2,1)$

A probabilidade de não decidir é dada por:

$$P_0 [0,25 < x < 0,5] = \int_{0,25}^{0,5} dx = 0,25$$

e

$$P_1 [0,25 < x < 0,5] = \int_{0,25}^{0,5} \frac{1}{2\sqrt{x}} dx = \sqrt{x} \Big|_{0,25}^{0,5} = 0,2071$$

Este é o tamanho do risco de não decidir que o pesquisador corre ao utilizar o teste unificado, ganhando, em troca, uma interpretação tanto freqüentista condicional quanto Bayesiana do teste, lembrando, ainda, que o teste freqüentista condicional neste caso foi elaborado a partir de uma estatística condicionante baseada em *valores-p*. Ou seja, a região crítica do teste foi definida a partir do critério estabelecido pelo teste de Neyman-Pearson; após esta definição efetuou-se uma partição do espaço amostral baseada em *valores-p*, como estabelecido por Fisher. E, ao se apresentar uma probabilidade de erro condicional de acordo com esta partição, estes valores são numericamente iguais às probabilidades *a posteriori*, fornecendo uma interpretação Bayesiana como proposta por Jeffreys.

4.2 O teste unificado para hipóteses sobre o parâmetro da Exponencial

Neste segundo exemplo será abordado o problema para diferentes tamanhos de amostras, considerando o teste de hipóteses $H_0: \theta = 1$ versus $H_1: \theta = 2$, em que $f(x; \theta) = \theta e^{-\theta x}$.

A razão de verossimilhanças é: $B(x_1, \dots, x_n) = \left(\frac{1}{2}\right)^n \exp\left\{\sum_{i=1}^n x_i\right\}$, com distribuição $f_B(b) = \frac{1}{b} \frac{\theta^n}{\Gamma(n)} \{\ln[2^n b]\}^{n-1} e^{-\theta \ln[2^n b]}$.

Então, $F_B^0(b) = P_{\theta=1}(B(X) \leq b) = \Gamma(n, \ln[2^n b])$

$F_B^1(b) = P_{\theta=2}(B(X) \leq b) = \Gamma(n, 2 \ln[2^n b])$.

Na Tabela 3.2 são rerepresentados os valores críticos b_c para se obter $\alpha = \beta$ no teste não-condicional (calculados no capítulo 1 deste trabalho); a partir desse valor, foi calculado o limite inferior da RND para cada caso. Com a função particionante $H(z)$, que define os limites da RND, utilizou-se o software *Maple 10* para a realização dos cálculos.

Vale lembrar que o limite superior neste caso é sempre 1, pois este exemplo trata do caso em que $b_c < 1$.

TABELA 3.2 – Valores críticos para o teste não-condicional de $X \sim \text{Exp}(1)$ versus $X \sim \text{Exp}(2)$ com diferentes tamanhos de amostra n e o limite inferior da RND.

n	$\alpha = \beta$ (%)	Valor Crítico $\left(\sum_{i=1}^n x_i\right)$	Valor Crítico (b_c)	Limite Inferior da RND
1	38,2	0,4812	0,809	0,7071
4	25,2	2,5461	0,7973	0,6717
10	14,1	6,7027	0,7955	0,6655
30	2,98	20,565	0,7947	0,6629
100	0,028	69,085	0,7945	0,6620

Simulou-se o procedimento de obtenção de amostras da distribuição Exponencial (1) versus Exponencial (2) de acordo com uma rotina desenvolvida no programa SAS apresentada no anexo B adaptada para a exponencial.

Avaliou-se a taxa de erro Tipo I e Tipo II para o procedimento tradicional com e sem a inclusão da região de não-decisão (Tabela 3.3).

TABELA 3.3 – Probabilidades do erro Tipo I e II com e sem a inclusão da RND no teste unificado para hipóteses sobre o parâmetro da Exponencial.

Tamanho da amostra	Teste usual		Teste Com Região de Não Decisão		
	Erro Tipo I	Erro Tipo II	Erro Tipo I	Não Decidiu	Erro Tipo II
1	0,3810	0,3806	0,2918	0,2285	0,2494
4	0,2530	0,2522	0,2168	0,0955	0,1968
10	0,1404	0,1404	0,1246	0,0421	0,1156
30	0,0298	0,0296	0,0270	0,0073	0,0254
100	0,0030	0,0025	0,0027	0,0007	0,0022

Pode-se perceber então que, como era esperado, com o aumento do tamanho da amostra, a probabilidade de se obter uma amostra na RND diminui, assim como acontece com os erros tipo I e II. Assim, a inclusão da RND, neste exemplo, apresenta várias qualidades desejáveis como o controle simultâneo das taxas de erros tipo I e tipo II, além do fato de poder ser usada no contexto Bayesiano ou freqüentista.

A inclusão da Região de Não-Decisão pode ser muito útil em testes de hipóteses quando se deseja uma concordância entre os diferentes contextos, pois traz redução na taxa de erro para um freqüentista clássico e apresenta medida de evidência Bayesiana e probabilidade de erro condicional numericamente iguais nas abordagens Bayesiana e Freqüentista Condicional, respectivamente.

É evidente que o fato de não decidir pode trazer incômodo para alguns pesquisadores, mas basta avaliar se isso é realmente menos aconselhável do que tomar uma decisão incorreta.

Como comentado anteriormente, o teste assim definido utiliza a partição baseada em *valores-p*. Ou seja, o teste unificado propõe o uso do *valor-p* de Fisher para condicionar um teste frequentista de Neyman-Pearson e, além disso, fornece uma interpretação Bayesiana de acordo com a escola de Jeffreys.

4.3 Estudo da região de não-decisão no teste unificado com perdas assimétricas

Suponha que X_1, \dots, X_n seja uma amostra aleatória i.i.d. de uma exponencial $f(x; \theta) = \theta e^{-\theta x}$ para $x > 0$, e que se deseje testar as hipóteses simples $H_0: \theta = \theta_0$ x $H_1: \theta = \theta_1$, sendo $\theta_1 > \theta_0$. Definindo $\gamma = \frac{\theta_0}{\theta_1}$, é imediato que

$B(X_1, \dots, X_n) = (\gamma)^n \exp\{(1 - \gamma)n\theta_1 \bar{X}\}$, em que B é monótona crescente em \bar{x} .

Note que $\gamma^n \leq B \leq \infty$.

Há um ponto de intersecção entre todas as curvas com os diferentes tamanhos de amostra m e n , o qual é obtido através da solução da equação:

$$\begin{aligned} \gamma^n e^{n(1-\gamma)\theta_1 \bar{x}} &= \gamma^m e^{m(1-\gamma)\theta_1 \bar{x}} \Rightarrow \gamma^{n-m} e^{(n-m)(1-\gamma)\theta_1 \bar{x}} = 1 \Rightarrow \\ &(n-m)\log \gamma + (n-m)(1-\gamma)\theta_1 \bar{x} = 0 \\ (n-m)\log \gamma &= (m-n)(1-\gamma)\theta_1 \bar{x} \Rightarrow \frac{-(m-n)}{(m-n)} \frac{\log \gamma}{(1-\gamma)} = \theta_1 \bar{x} \Rightarrow \boxed{\bar{x} = -\frac{\log \gamma}{\theta_1 (1-\gamma)}} \end{aligned}$$

Note que \bar{x} não depende de m e n (dois tamanhos quaisquer da amostra). Tem-se o valor de \bar{x} explicitamente em função de γ e θ_0 .

Como demonstrado no capítulo 1, tem-se que a densidade acumulada de B é:

$$\text{Sob } H_0: F_B^0(b) = P_{\theta=\theta_0}(B(X) \leq b) = \Gamma\left(n, \frac{\theta_0}{(\theta_1 - \theta_0)} \ln\left[\left(\frac{\theta_1}{\theta_0}\right)^n b\right]\right)$$

$$\text{Sob } H_1: F_B^1(b) = P_{\theta=\theta_1}(B(X) \leq b) = \Gamma\left(n, \frac{\theta_1}{(\theta_1 - \theta_0)} \ln\left[\left(\frac{\theta_1}{\theta_0}\right)^n b\right]\right).$$

Ou, em termos de γ :

$$\text{Sob } H_0: F_B^0(b) = P_{\theta=\theta_0}(B(X) \leq b) = \Gamma\left(n, \frac{\gamma}{(1-\gamma)} \ln[\gamma^{-n} b]\right)$$

$$\text{Sob } H_1: F_B^1(b) = P_{\theta=\theta_1}(B(X) \leq b) = \Gamma\left(n, \frac{1}{(1-\gamma)} \ln[\gamma^{-n} b]\right), \text{ em que } \Gamma \text{ é a}$$

distribuição gama incompleta.

Para o caso particular $n = 1$, tem-se:

$$F_B^0 = \int_0^{\frac{\gamma}{(1-\gamma)} \ln[\gamma^{-1} b]} e^{-w} dw = -e^{-w} \Big|_0^{\frac{\gamma}{(1-\gamma)} \ln[\gamma^{-1} b]} = 1 - \left[\frac{\gamma}{b}\right]^{\left[\frac{\gamma}{(1-\gamma)}\right]}$$

e

$$F_B^1 = \int_0^{\frac{1}{(1-\gamma)} \ln[\gamma^{-1} b]} e^{-w} dw = -e^{-w} \Big|_0^{\frac{1}{(1-\gamma)} \ln[\gamma^{-1} b]} = 1 - \left[\frac{\gamma}{b}\right]^{\left[\frac{1}{(1-\gamma)}\right]}$$

Com a RND generalizada definida pelo teste $Y_{l,\rho}$ $0 < l < \infty$ e $0 < \rho < \infty$:

$$r_{l,\rho} = l\rho \quad \text{e} \quad a_{l,\rho} = [F_B^0]^{-1}(1 - \rho F_B^1(l\rho)) \quad \text{se} \quad F_0(l\rho) \leq \rho - F_1(l\rho); \quad (1)$$

$$r_{l,\rho} = [F_B^1]^{-1}((1/\rho)(1 - F_B^0(l\rho))) \quad \text{e} \quad a_{l,\rho} = l\rho \quad \text{se} \quad F_0(l\rho) > \rho - F_1(l\rho). \quad (2)$$

Os valores de $r_{l,p}$ e $a_{l,p}$ podem ser obtidos explicitamente. Considerando o efeito de l , a razão entre as perdas l/l_0 , quando $\rho = \frac{\pi_1}{\pi_0} = 1$ e $\gamma = \frac{\theta_0}{\theta_1} = \frac{1}{2}$ (para testar $H_0: \theta = 1$ contra $H_1: \theta = 2$ com probabilidade *a priori* iguais), tem-se:

$B(X) = \left(\frac{1}{2}\right) \exp\{X\}$ em que $B(x)$ é monótona crescente em x e $\frac{1}{2} \leq B \leq \infty$,

com funções acumuladas dadas por $F_B^0(b) = \left(1 - \frac{1}{2b}\right)$ e $F_B^1(b) = 1 - \left(\frac{1}{2b}\right)^2$. Os

limites da RND são:

$$r_{l,l} = l \quad \text{e} \quad a_{l,l} = [F_B^0]^{-1}(1 - F_B^1(l))$$

$$a_{l,l} = [F_B^0]^{-1}\left(1 - \left(1 - \left(\frac{1}{2l}\right)^2\right)\right) = [F_B^0]^{-1}\left(\frac{1}{4l^2}\right), \text{ como } F_B^0(l) = \left(1 - \frac{1}{2l}\right) \Rightarrow$$

$$[F_B^0]^{-1}(l) = (2(1-l))^{-1}, \text{ então } a_{l,l} = [F_B^0]^{-1}\left(\frac{1}{4l^2}\right) = \left(2\left(1 - \frac{1}{4l^2}\right)\right)^{-1}$$

De (2), tem-se:

$$r_{l,l} = [F_B^1]^{-1}(1 - F_B^0(l)) \quad \text{e} \quad a_{l,l} = l$$

$$r_{l,l} = [F_B^1]^{-1}(1 - F_B^0(l)) = [F_B^1]^{-1}\left(\frac{1}{2l}\right), \text{ como } F_B^1(l) = 1 - \left(\frac{1}{2l}\right)^2 \Rightarrow$$

$$[F_B^1]^{-1}(l) = (4(1-l))^{-\frac{1}{2}}, \text{ então } r_{l,l} = [F_B^1]^{-1}\left(\frac{1}{2l}\right) = r_{l,l} = \left(4\left(1 - \frac{1}{2l}\right)\right)^{-\frac{1}{2}}$$

O comportamento de (1) e (2) pode ser resumido na expressão dada seguinte, considerando sempre $r_{l,p}$ como limite inferior da região de não-decisão e $a_{l,p}$, como limite superior (Figura 3.21).

$$r_{l,l} = \min\left\{l, \left(\frac{4l-2}{l}\right)^{-\frac{1}{2}}\right\} \quad \text{e} \quad a_{l,l} = \max\left\{l, \left(2 - \frac{1}{2l^2}\right)^{-1}\right\} \quad \text{se } \frac{1}{2} \leq l \leq \infty.$$

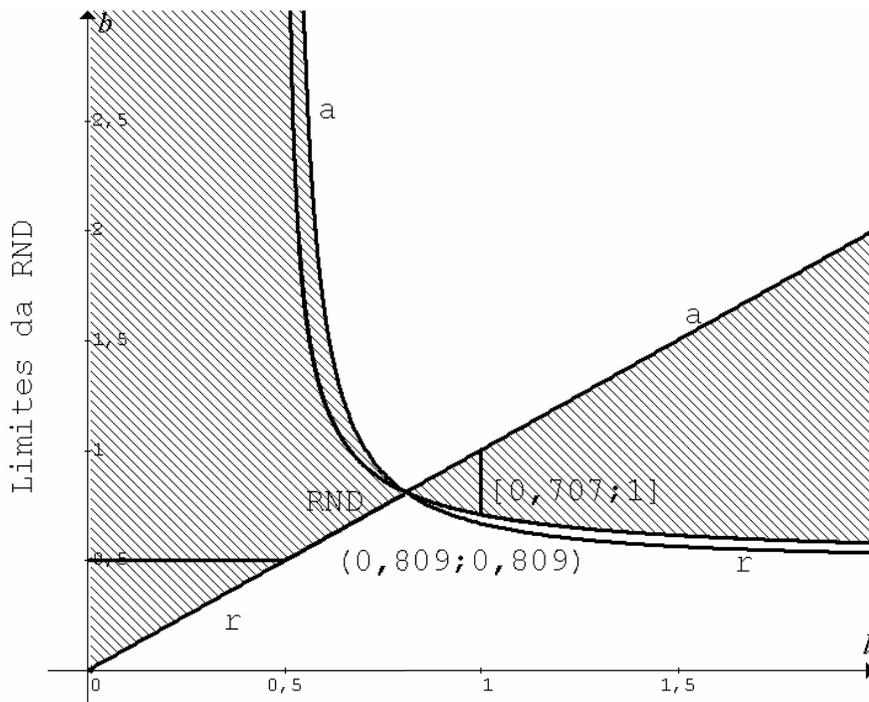


FIGURA 3.21 - A região de não-decisão $]r,a[$ como uma função de l , para testar hipóteses sobre o parâmetro da Exponencial $H_0:\theta = 1$ contra $H_1:\theta = 2$ sob probabilidades *a priori* iguais com $n=1$.

Lembrando que os testes de maior interesse neste trabalho são aqueles em que $\alpha = \beta$, nesse caso, em que $n = 1$, corresponde a ter o valor crítico $b_c = 0,809$ (Tabela 3.2). Então, quando a razão entre as perdas $l = 0,809$, ou seja, perdas assimétricas, o valor crítico no teste Bayesiano, dado por $l\rho$, estaria em completa concordância com o teste freqüentista e, por isso, a RND é vazia (Figura 3.21). O teste $\Upsilon_{l=0,809, \rho=1}$ coincide com o teste freqüentista de Neyman-Pearson com probabilidades de erros iguais, $\alpha = \beta = 0,382$. O valor 0,809 é o

mesmo obtido pela expressão $\rho^* = \frac{(1-\alpha)}{(1-\beta)}$ e $\ell^* = \frac{b_c}{\rho^*}$ fornecida no item 2.5 do referencial teórico deste capítulo.

Quando as perdas são simétricas, ou seja, $l = 1$, a RND não é vazia, pois leva em conta que um Bayesiano discordaria da decisão de um frequentista no intervalo de 0,809 a 1. Isto ocorre porque, para o frequentista, esta região indica rejeição da hipótese nula. Todavia, para o Bayesiano, a medida de evidência probabilidade *a posteriori* associada a esta região é dada por um valor superior a $\frac{1}{2}$, o que indica uma aceitação da hipótese nula. Esta incoerência é solucionada incluindo uma RND. Devido ao critério de condicionamento, deve-se excluir também a região $]0,707; 0,809[$, que contém os valores correspondentes aos pares $(z, H(z))$, ficando, dessa forma, com a RND no intervalo $]0,707; 1[$. Esta região está de acordo com os limites estabelecidos na Tabela 3.2.

Percebe-se também que, para $l < \frac{1}{2}$, tem-se que $r_{l,p} = l$ e $a_{l,p} = \infty$. Como $B(x) = \frac{1}{2}e^x \Rightarrow B(x) > \frac{1}{2}$, uma decisão nunca é tomada nesta situação, pois qualquer valor b obtido pertence à RND.

Um comentário sobre a razão das densidades *a priori* é que, de acordo com Berger et al. (1994), quando $\rho \neq 1$, a região de não decisão cresce enormemente quando $n \rightarrow \infty$.

Para n grande, a RND deveria ser pequena, pois a incerteza deveria diminuir; porém, o fato aqui é a importância que *a priori* tem na decisão. Se há imparcialidade, ou seja, densidades *a priori* não informativas, a razão vai para 1, e a probabilidade da região de não-decisão vai para zero.

4.4 Exemplo de Aplicação

Utilizando os mesmos dados do exemplo de aplicação do capítulo 2, pode-se fazer uma análise do comportamento da RND para este caso.

Para o valor crítico de $B(x) = 1$ (ou $x = 1,255$), os valores de PEC's e probabilidades *a posteriori* são numericamente iguais. Entretanto, se fosse mantido o valor crítico em 1,35 (que representa um valor $B(x) = 0,3104$, para o teste com $\alpha = 0,05$), deveria ser incluída a RND para que o teste tivesse uma interpretação freqüentista condicional e Bayesiana. Nesse caso, o valor $l\rho$ seria dado por 0,3104.

Os valores para os limites da RND seguem o mesmo padrão de construção do teste Y_{π}^* , sendo que, quando $F_B^0(l\rho) \leq \rho - F_B^1(l\rho)$, então $r = l\rho$ e $a_{l,\rho} = [F_B^0]^{-1}(1 - \rho F_B^1(l\rho))$. Tem-se, então, que $F_B^0(a) = P_{\theta_0}[B(X) \leq a] = 1 - \rho F_B^1(l\rho)$. Considerando $l = 0,3104$, tem-se $\rho = 1$.

Mas a distribuição de $B(X)$ é uma lognormal, pois, se $X \sim N$, então $\exp(X) \sim \text{LogNormal}$, sendo, nesse caso: $f_B^0(b) = \log normal(2,52; (3,89)^2)$ e $f_B^1(b) = \log normal(-2,523; (3,89)^2)$ (Figura 3.22).

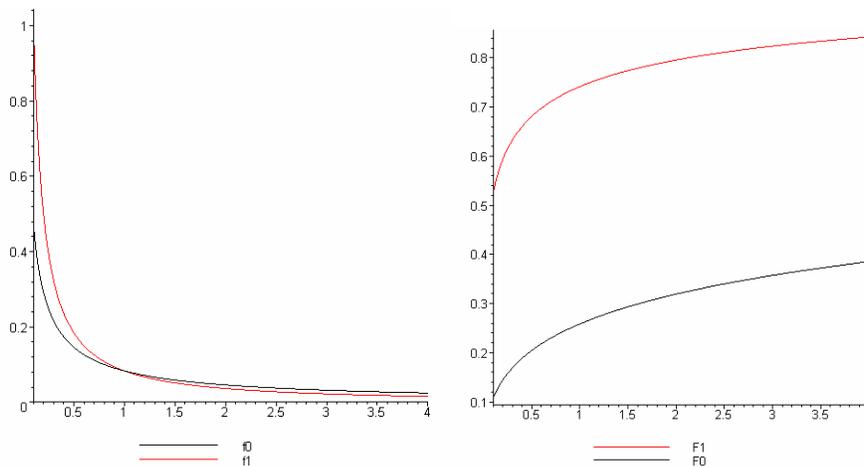


FIGURA 3.22 – Função densidade de probabilidade e distribuição acumulada considerando $f_B^0(b) = \log normal(2,52; (3,89)^2)$ e $f_B^1(b) = \log normal(-2,523; (3,89)^2)$.

Então:

$$a_{l,\rho} = [F_B^0]^{-1}(1 - \rho F_B^1(l\rho)) = [F_B^0]^{-1}(1 - F_B^1(0,3104)) = [F_B^0]^{-1}(1 - 0,636) = [F_B^0]^{-1}(0,364) \Rightarrow a_{l,\rho} = 3,212$$

A RND seria construída da seguinte forma:

$$Y_{l,\rho} = \begin{cases} \text{se } B(x) \leq 0,3104, \text{ ou } \bar{x} \geq 1,35 \text{ rejeita-se } H_0 \\ \text{com probabilidade a posteriori ou PEC I, } \alpha_x = f(\theta_0 | x) = \frac{B(x)}{B(x)+1}; \\ \text{se } 0,3104 < B(x) < 3,212 \text{ ou } 1,16 < \bar{x} < 1,35, \\ \text{não se toma nenhuma decisão;} \\ \text{se } B(x) \geq 3,212 \text{ ou } \bar{x} \leq 1,16 \text{ aceita-se } H_0 \\ \text{com probabilidade a posteriori ou PEC II, } \beta_x = f(\theta_1 | x) = \frac{1}{B(x)+1}. \end{cases}$$

Como se trata de razão de verossimilhança simétrica, a probabilidade da RND quando H_0 é verdadeira é igual à probabilidade desta quando H_1 é verdadeira, sendo dada por 0,1926 (Figura 3.23).

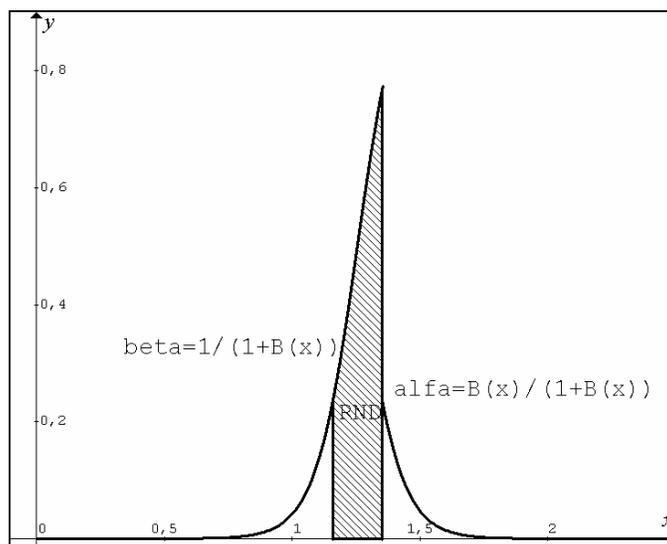


FIGURA 3.23 – Área correspondente à Região de não-decisão.

Na amostra observada neste exemplo, com média igual a 1,248, seria o caso de não tomar decisão alguma. Esta decisão está coerente com a probabilidade de erro condicional tipo II apresentada no capítulo anterior, pois no teste freqüentista condicional não modificado a decisão era de aceitação, mas com PEC II muito alta (igual a 0,479).

CONCLUSÕES

O conceito de função particionante simplifica a construção do teste unificado.

Nos exemplos desenvolvidos, a inclusão da região de não-decisão diminuiu a taxa de um dos erros do teste unificado, mas forneceu uma probabilidade de não-decisão proporcional. As esperanças dos erros condicionais também foram proporcionalmente diminuídas.

Observou-se que o tamanho da região de não-decisão foi influenciado pela razão entre as perdas e pelo tamanho da amostra.

Uma das formas de se manter a região de não-decisão vazia é utilizar a razão entre as perdas igual ao valor crítico do teste com erros não-condicionais iguais, no caso que a razão entre as densidades *a priori* for igual a 1, ou escolher um valor crítico b_c dependente de ρl , o que poderia provocar problemas com relação à esperança dos erros condicionais.

REFERÊNCIAS BIBLIOGRÁFICAS

BERGER, J. O. Could Fisher, Jeffreys and Neyman have agreed on testing? **Statistical Science**, Beachwood, v. 18, n. 1, p. 1-12, Feb. 2003. (Com discussão: p. 13-32)

BERGER, J. O.; BROWN, L. D.; WOLPERT, R. L. A unified conditional frequentist and Bayesian test for fixed and sequential simple hypothesis testing. **The Annals of Statistics**, Hayward, v. 22, n. 4, p.1787-1807, Dec. 1994.

BERGER, J.; BOUKAI, B.; WONG, Y. Unified frequentist and Bayesian testing of a precise hypothesis. **Statistics Science**, Hayward, v. 12, n. 3, p. 133-148, 1997. (Com discussão: p. 149-160)

KIEFER, J. Conditional confidence statements and confidence estimators **Journal of the American Statistical Association**, Alexandria, v. 72, n. 360, p. 789-807, Dec. 1977. (Com discussão: p. 808-827).

SAS INSTITUTE SAS/STAT. **SAS/IML Software**: usage and reference, version 6. Cary, 1990. 501p.

ANEXOS

	Página
ANEXO A Rotina SAS, programa utilizado para a obtenção do número de vezes em que a amostra gerada pertencia à região de não-decisão no teste das hipóteses simples $H_0: X \sim \text{Uniforme}(0,1)$ versus $H_1: X \sim \text{Beta}(1/2,1)$	207
ANEXO B Rotina SAS, programa utilizado para a obtenção do número de vezes em que a amostra pertencia à região de não-decisão no teste de hipóteses sobre o parâmetro de distribuições exponenciais considerando os tamanhos amostrais: $n = 1$; $n = 4$; $n = 10$; $n = 30$ e $n = 100$	209

ANEXO A

```
proc iml;
  /* Teste RND para uniforme (0,1) x beta (1/2,1) */
  n=1; nsim=100000;
  errala=0; acertala=0; rnd1=0; rnd2=0;
  erra2a=0; acert2a=0; erralb=0; acertalb=0; erra2b=0;
  acert2b=0;
  x1=0; y1=0;
  do il=1 to nsim;
    u=ranuni(0);

    if (u<=0.5) then do;
      do i=1 to n;
        x=ranuni(0); /* simula uniforme (0,1)*/
        x1=x1+ x;
      end;
    /* teste Neyman-Pearson */
    mx=x1/n;
    x1=0;
    bx=2*((n*mx)**(1/2));
    if (bx<=1.236)then errala=errala+1;
    if (bx>1.236) then acertala=acertala+1;
  /* teste RND */
    if (bx<=1)then erralb=erralb+1;
    if (bx>=1.414) then acertalb=acertalb+1;
    if (bx<1.414 && bx>1) then rnd1=rnd1+1;
  end;

    if (u> 0.5) then do;
      do j=1 to n;
        y=ranuni(0)**2; /* simula beta (1/2,1)*/
        y1=y1+ y;
      end;
    /* teste classico */
    my=y1/n;
    y1=0;
    by=2*((n*my)**(1/2));
    if (by>1.236)then erra2a=erra2a+1;
    if (by<=1.236) then acert2a=acerta2a+1;
  /* teste RND */
    if (by>=1.414)then erra2b=erra2b+1;
    if (by<=1) then acert2b=acerta2b+1;
  end;
end;
```

```
        if (by<1.414 && by>1) then rnd2=rnd2+1;
end;

end;
total=acertala+acerta2a+acertalb+acerta2b+errala+erralb+err
a2a+erra2b+rnd1+rnd2;
*print acertala;
*print acerta2a;
print errala;
print erra2a;
*print acertalb;
*print acerta2b;
print erralb;
print erra2b;
print rnd1;
print rnd2;
print total;

run;

quit;
```

ANEXO B

```
proc iml;
  /* Teste RND para exponencial (theta0) x
  exponencial(theta1) */
  /* Importante: distribuicao exponencial dada por
  y=(theta)e^(-theta*x) sera gerada por ranexp(0)/theta */

  n=100; nsim=10000; theta0=1; theta1=2;

  /* teste Neyman-Pearson */
  criticoa=0.7945;

  /* teste RND */
  LIb=0.6620;
  LSb=1;

  errala=0; acertala=0; rnd1=0; rnd2=0;
  erra2a=0; acert2a=0; erralb=0; acertalb=0; erra2b=0;
  acert2b=0;
  x1=0; y1=0;

  do il=1 to nsim;
    u=ranuni(0);

    if (u<=0.5) then do;
      do i=1 to n;
        z1=ranexp(0);
        x=z1/theta0; /* simula exponencial (theta0)*/
        x1=x1+ x;
      end;

  /* teste classico */
      mx=x1/n;
      x1=0;
      bx=((theta0/theta1)**n)*exp(((theta0)-(theta1))*(-
1)*mx*n);
      if (bx<=criticoa)then errala=errala+1;
      if (bx>criticoa) then acertala=acertala+1;

  /* teste RND */
      if (bx<=LIb)then erralb=erralb+1;
      if (bx>=LSb) then acertalb=acertalb+1;
      if (bx<LSb && bx>LIb) then rnd1=rnd1+1;
    end;
  end;
```

```

    if (u> 0.5) then do;
        do i=1 to n;
            z2=ranexp(0);
            y=z2/theta1; /* simula exponencial (theta1)*/
            y1=y1+ y;
        end;

/* teste classico */
    my=y1/n;
    y1=0;
    by=((theta0/theta1)**n)*exp(((theta0)-(theta1))*(-
1)*my*n);
    if (by>=criticoa)then erra2a=erra2a+1;
    if (by<criticoa) then acerta2a=acerta2a+1;

/* teste RND */
    if (by>=LSb)then erra2b=erra2b+1;
    if (by<=LIb) then acerta2b=acerta2b+1;
    if (by<LSb && by>LIb) then rnd2=rnd2+1;
end;

end;
total=acertala+acerta2a+acertalb+acerta2b+errala+erralb+err
a2a+erra2b+rnd1+rnd2;
rnd=rnd1+rnd2;
erralA= 2*errala;
erra2A= 2*erra2a;
erralB= 2*erralb;
erra2B= 2*erra2b;
RND=2*rnd;
print erralA;
print erra2A;
print erralB;
print RND;
print erra2B;
print total;

run;

quit;

```