

## REGIÃO DE NÃO-DECISÃO EM TESTES COM INTERPRETAÇÃO FREQUENTISTA CONDICIONAL E BAYESIANA

Carla Regina Guimarães BRIGHENTI<sup>1</sup>  
Lucas Monteiro CHAVES<sup>2</sup>

- RESUMO: O Teste Bayesiano é um Teste de Razão de Verossimilhanças com valor crítico dado pelo produto da razão entre as densidades *a priori*  $\rho$  pela razão entre as perdas  $l$ . Para determinados valores de  $\rho$  e  $l$ , têm-se regiões críticas coincidentes para um teste bayesiano e para um teste com erros frequentistas condicionais. Uma medida de evidência apresentada no teste bayesiano é a probabilidade *a posteriori*, enquanto no teste frequentista condicional, é a chamada probabilidade de erro condicional (PEC). Dependendo da estatística condicionante utilizada e dos valores de  $\rho$  e  $l$ , esses valores podem ser numericamente iguais, definindo então um teste “unificado”. Mas quando a probabilidade *a posteriori* ou a PEC são numericamente iguais e superiores a  $\frac{1}{2}$ , para que os testes sejam coerentes, deve ser incorporada uma “região de não-decisão”, RND. A construção da RND depende da função particionante  $H$  do teste condicional e dos valores de  $\rho$  e  $l$ . Neste trabalho, estudou-se o tamanho da RND para teste de hipóteses simples sobre o parâmetro da distribuição Exponencial sob amostras de diferentes tamanhos. Concluiu-se que o tamanho da RND é dependente do tamanho da amostra e que determinadas situações pode tornar o teste unificado inadequado.
- PALAVRAS-CHAVE: Probabilidade *a posteriori*; probabilidade de erro condicional; razão de verossimilhanças; região de não-decisão.

### 1 Introdução

O problema de testar hipóteses estatísticas é um dos pontos de discordância entre Bayesianos e frequentistas. O método frequentista tradicional constrói uma região crítica e apresenta probabilidades de erros pré-experimentais. O uso desta abordagem frequentista tradicional é criticada porque as probabilidades dos erros não refletem integralmente as informações proporcionadas pelos dados. Uma alternativa comum é a utilização de 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 frequentista verdadeira e tem suas próprias limitações como medida de evidência. Para solucionar estes problemas, uma alternativa possível é modificar o método frequentista tradicional incorporando procedimentos baseados em condicionamento que sejam dependentes dos dados utilizando a abordagem denominada frequentista condicional, formalizada por Kiefer (1977). Uma outra

---

<sup>1</sup> Escola Preparatória de Cadetes do Ar – EPCAR, Comando da Aeronáutica, CEP: 36200-000, Barbacena, MG, Brasil. E-mail: [carla\\_brighenti@hotmail.com](mailto:carla_brighenti@hotmail.com)

<sup>2</sup> Departamento de Ciências Exatas, Universidade Federal de Lavras – UFLA, Caixa Postal 37, CEP:37200-000, Lavras, MG, Brasil. E-mail: [lucas@ufla.br](mailto:lucas@ufla.br)

abordagem é a Bayesiana, que é baseada na forma mais extrema de condicionamento, nos próprios dados observados.

Berger, Brown e Wolpert (1994) sugeriram, para o caso de hipótese simples versus hipótese simples, um procedimento que promovesse uma unificação entre os métodos Bayesianos e freqüentistas condicionais 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 apresentariam as mesmas medidas de evidência ao rejeitar ou aceitar, sendo as medidas de evidências freqüentista condicional (probabilidades de erros condicionais) e Bayesiana (probabilidades *a posteriori*) numericamente iguais.

Alguns problemas com o chamado teste unificado podem ocorrer para valores particulares de perdas e densidades *a priori*; para solucionar esse problema, Berger, Brown e Wolpert (1994) incluíram no teste uma região de não-decisão (RND). A RND corresponderia a uma região na qual potencialmente não haveria consenso, pois nenhum estatístico sentiria que as evidências são fortes o suficiente para uma decisão conclusiva.

Este trabalho verifica o comportamento da região de não-decisão em testes de hipóteses simples versus simples, no caso de razão de verossimilhanças simétrica e não-simétrica. Para tal objetivo, se fez necessário, inicialmente, uma referência aos testes de hipóteses simples na seção 2, que foi subdividida em três itens, sendo que na seção 2.1 foi realizada uma breve descrição dos testes de hipóteses no contexto freqüentista e suas principais diferenças em relação aos testes de significância. Na seção 2.2 o mesmo tipo de teste é discutido no enfoque Bayesiano com a inclusão de uma medida de evidência e na seção 2.3 o teste de hipóteses freqüentista condicional. Na seção 3 apresenta-se o teste unificado, objeto principal deste trabalho, cujo maior problema é a região de não-decisão, trabalhada detalhadamente na seção 4. Na seção 5 exemplificou-se a construção da região de não-decisão no caso de um teste com razão de verossimilhanças simétrica e na seção 5.1 discutiu-se sua generalização. Na seção 6 apresentou-se o caso não-simétrico com uma análise com dados simulados sendo realizada na seção 6.1, seguindo-se as conclusões.

## **2 Teste de hipóteses simples versus simples**

### **2.1 Teste de hipóteses no contexto freqüentista**

A teoria de testes de hipóteses sistematizada por Neyman e Pearson (1933) 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  $B$  para um valor crítico  $b_c$ , a hipótese nula é recusada se  $B \leq b_c$ . 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.

A maior crítica à teoria dos testes de hipóteses de Neyman-Pearson é devida ao fato de os erros Tipo I e Tipo II não refletirem a evidência dos dados observados; 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  $C_T$ , 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. Considerando  $x_1$ ,  $x_2$ ,  $x_3$  e  $x_4$  diferentes observações no espaço amostral  $X^n$ , na Figura 1 pode-se perceber que apesar de as observações  $x_1$  e  $x_3$  estarem muito próximas, corresponderiam a decisões diferentes, enquanto,  $x_1$  e  $x_2$ , levariam à mesma decisão de se rejeitar a hipótese nula. O mesmo ocorrendo com  $x_3$  e  $x_4$ .

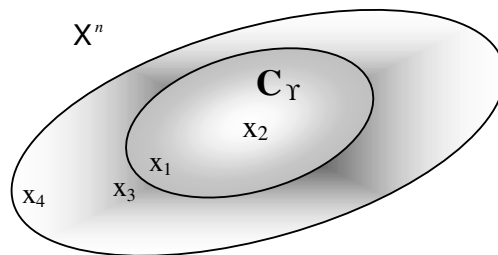


Figura 1 - Representação, no espaço amostral das realizações, de diferentes amostras em relação à região crítica, tanto para rejeição ( $x_1$ ,  $x_2$ ) quanto para aceitação ( $x_3$ ,  $x_4$ ) da hipótese nula.

Tal fato pareceu a Ronald Fisher altamente não científico. Fisher também criticou a necessidade de se considerar uma hipótese alternativa (Berger, 2003). Como o objetivo de um experimento é encontrar a magnitude das diferenças entre as hipóteses, a forma tradicional de se tentar contornar este problema é através de um teste de significância (teste Fisheriano) utilizando um *valor-p*. Observe, entretanto, que *valores-p* não são verdadeiras medidas de erro freqüentistas (Berger, Brown & Wolpert, 1994). 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 erros freqüentistas em testes de hipóteses. 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 razão de verossimilhanças, utilizada nos testes de hipóteses simples, que compara modelos diferentes. Geralmente, tem-se uma hipótese alternativa  $H_1$ , mas a construção do teste de significância não permite incorporar uma hipótese alternativa específica (Berger, 2003).

A principal discussão é que alguns pesquisadores – nas aplicações – utilizam *valores-p*, mas depois se precipitam 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 probabilidades 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 uma significância. A principal diferença entre testes de significância e o teste como regra para decidir entre duas 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  $H_0$ . 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, devem-se focalizar igualmente duas probabilidades, as probabilidades dos dois tipos de erro, fazendo opção por uma das duas hipóteses.

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,  $\alpha$ . Devido a uma idéia comum, 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. 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. Nesse caso, seria adequado pensar em um teste no qual  $\alpha = \beta$ . O problema nesse caso é que há situações em que os valores de  $\alpha$  e  $\beta$  poderiam ser muito grandes, e para diminuir o tamanho dos erros, o tamanho da amostra deve ser aumentado excessivamente.

## 2.2 Teste de hipóteses no contexto Bayesiano

Na teoria de teste de hipóteses é usual ponderar os erros tipo I e tipo II através de uma função denominada função perda,  $l(\delta_i; \theta_j)$ , em que  $\delta_i$  é a decisão de aceitar  $H_i$  e  $\theta_j$  é o valor do parâmetro. Geralmente  $l(\delta_i; \theta_i) = 0$ , enquanto  $l(\delta_1; \theta_0) = l_0$  e  $l(\delta_0; \theta_1) = l_1$  (Garthwaite et al., 1995).

Para testes de hipóteses dado por  $H_0: X \sim f_0(x)$  versus  $H_1: X \sim f_1(x)$ , o Lema de Neyman-Pearson estabelece que os testes mais poderosos de tamanho  $\alpha$  são os testes de razão de verossimilhanças  $B(x) = \frac{L_0}{L_1} = \frac{\prod_{i=1}^n f(x_i | \theta_0)}{\prod_{i=1}^n f(x_i | \theta_1)}$ ,  $x = (x_1, \dots, x_n)$ , isto é,

aqueles que têm as regiões críticas da forma:  $C = \{(x_1, \dots, x_n) : B(x) \leq b_c\}$ , para alguma constante não-negativa  $b_c$ . Mas testes de Bayes e Testes de Razão de Verossimilhanças são equivalentes, isto é, o Teste de Razão de Verossimilhanças é um Teste de Bayes para alguma densidade *a priori*  $\pi = (\pi_0, \pi_1)$  e o Teste de Bayes é um Teste de Razão de Verossimilhanças com  $b_c = \frac{\pi_1 l_1}{\pi_0 l_0}$ . Assim, considerando o seguinte teste Bayesiano

$Y_\pi$  com região crítica dada por:

$$C_\pi = \left\{ (x_1, \dots, x_n) : B(x) = \frac{f(x|\theta_0)}{f(x|\theta_1)} \leq \frac{l_1\pi_1}{l_0\pi_0} \right\}, \quad (1)$$

se a razão entre as densidades *a priori* for  $\rho = \frac{\pi_1}{\pi_0}$  e a razão entre as perdas for  $l = \frac{l_1}{l_0}$ , como a densidade *a posteriori*  $f(\theta_i|x) \approx f(x|\theta_i)\pi_i$ , tem-se

$$C_\pi = \left\{ (x_1, \dots, x_n) : \frac{f(\theta_0|x)}{f(\theta_1|x)} \leq \frac{l_1}{l_0} \right\} \quad (2)$$

(Mood et al., 1974).

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

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

Se for utilizada uma distribuição *a priori* objetiva  $\pi_0 = 1/2$  e 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:

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

Portanto, os testes consistem em simplesmente aceitar a hipótese que possui a maior probabilidade *a posteriori*.

No problema estatístico de Testes de Hipóteses sob a abordagem Bayesiana após a realização do teste, é necessário apresentar uma medida de evidência. Na abordagem freqüentista, como comentado no item 2.1, é muito usado, às vezes inconvenientemente, o *valor-p* como medida de evidência, que se baseia apenas na distribuição da estatística do teste sob  $H_0$ . Na abordagem Bayesiana, as medidas de evidência mais usuais para teste de hipóteses 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).

Se forem consideradas perdas iguais e a densidade *a priori* for particularizada para  $\pi_i = 1/2$ , tem-se que as perdas esperadas *a posteriori* são simplesmente as probabilidades *a posteriori*.

### 2.3 Teste de hipóteses freqüentista condicional

Para superar as dificuldades apresentadas em 2.1, Kiefer (1977) propôs que os freqüentistas tentassem empregar a metodologia denominada “*teste freqüentista condicional*”. A idéia básica por trás desta abordagem freqüentista condicional é construir uma medida estatística de “força de evidência” nos dados. Entenda-se “força de evidência” como uma grandeza que indica, por meio de cálculos de probabilidade, a existência de razões para rejeição ou aceitação de uma hipótese.

A idéia é particionar o espaço amostral,  $X^n$ , utilizando uma estatística condicionante  $Z$ , tal que

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

e então desenvolver medidas de evidência frequentistas condicionais em  $C^z$ .

Um teste estatístico  $\Upsilon$  com região crítica  $C_\Upsilon$  define uma regra de decisão  $\delta(x)$ , que é uma função do espaço amostral  $X^n$  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  $x \in \bar{C}_\Upsilon$ . Considerando agora uma estatística  $Z$ , que será denominada de *estatística condicionante*, tem-se:

- $C^{z_i} = \{x \in X^n; Z(x) = z_i\}$  o conjunto das amostras para as quais a estatística  $Z$  vale  $z_i$ .
- $C_0^{z_i} = C^{z_i} \cap \bar{C}_\Upsilon$ , 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 rejeitar  $H_0$  (Figura 2).

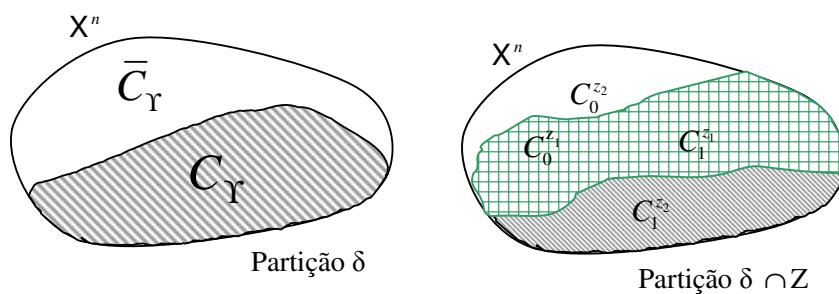


Figura 2 - Desenho esquemático considerando as duas partições definidas por  $\delta$  e por  $Z$  no caso de  $Z = \{z_1, z_2\}$ .

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 taxas de erros condicionais dos tipos I e II são assim definidas:

$\alpha_{z_i}$  = Probabilidade de se cometer um erro Tipo I dado que  $Z = z_i$  foi observada,  $P_{\theta_0}[(\text{Rejeitar } H_0 | C^{z_i})]$ .

$\beta_{z_i}$  = Probabilidade de se cometer um erro Tipo II dado que  $Z = z_i$  foi observada,  $P_{\theta_1}[(\text{Aceitar } H_0 \mid C^{z_i})]$ .

serão chamadas de “probabilidades de erro condicionais” e abreviadas por “PECI” e “PECII”, respectivamente.

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-se os erros condicionais  $\alpha_{z_i}$  pelas probabilidades de  $C^{z_i}$  ocorrerem, tem-se a probabilidade de erro não condicional. De fato, como para cada valor  $z_i$  da variável aleatória  $Z$  tem-se uma probabilidade de erro condicional  $\alpha_{z_i}$ , pode-se considerar estas como variáveis aleatórias  $\alpha_{z_i}$  e  $\beta_{z_i}$ . Tem-se, então:

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

ou seja, a esperança da probabilidade do erro condicional é a própria probabilidade de erro freqüentista estipulada no teste freqüentista de Neyman-Pearson (Berger, Brown & Wolpert, 1994). Assim, a partição da região crítica e da região de aceitação fornece testes com probabilidades de erro condicionais em função dos dados, mas que, em média, ou seja, em termos freqüentistas, apresentam a mesma probabilidade de erro estabelecida inicialmente pelo valor crítico utilizado. Desta forma é possível obter probabilidades de erros que sejam dependentes da distância à fronteira da região crítica. Considerar que a estatística condicionante  $Z$  assume apenas um número finito de valores é muito restritivo. Em razão disso, Kiefer (1977) desenvolve sua teoria considerando uma “partição contínua” definida por estatísticas que assumem infinitos valores reais.

Considere o caso em que  $n = 1$  e o espaço amostral  $X$  tem dimensão 1, isto é,  $X = \mathfrak{R}$ . A idéia é partir  $\mathfrak{R}$  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.

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  $A_0$ ,  $A_0'$ ,  $A_1$ ,  $A_1'$ , em que  $A_1$  corresponde à área sob a curva  $f_1$  no intervalo  $[x, x + \Delta x]$  e  $A_1'$  corresponde à área sob a curva  $f_1$  no intervalo  $[x', x' + \Delta x]$ .

Nesse caso, a Probabilidade de Erro Condicional Tipo I (PECI) é dada por

$$\alpha_{x'} = \frac{A_0'}{A_0' + A_0}; \text{ ou seja, sob a hipótese nula, o erro condicional é dado pela probabilidade}$$

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 (PECII), que acontece se  $x$  é observado e aceita-se  $H_0$  quando esta é falsa, é dada por

$$\beta_x = \frac{A_1}{A_1' + A_1} \text{ (Figura 3).}$$

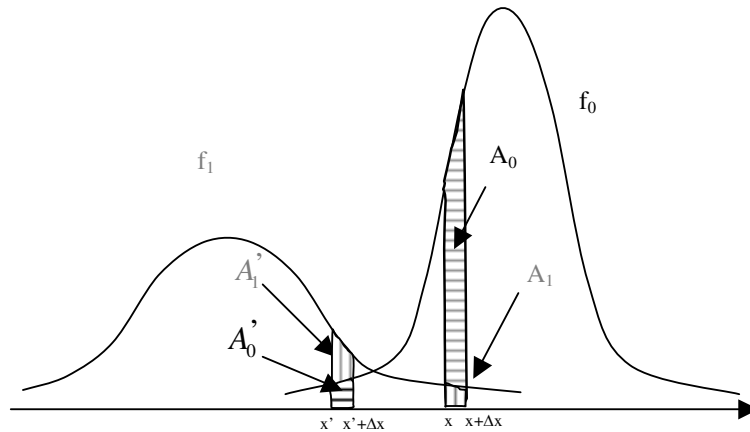


Figura 3 - Desenho esquemático da relação entre as áreas associadas aos pontos  $x$  e  $x'$  em  $f_0$  e  $f_1$  com  $k \in ]x' + \Delta x, x[$ .

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

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

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

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”.

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 (Brighenti, 2007).

Definida uma relação funcional entre  $x$  e  $x'$ , fica também estabelecida uma relação funcional entre as áreas  $A_0$  e  $A_0'$  assim como entre  $A_1$  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



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

Então  $A_0' = |F_0(H(x+\Delta x)) - F_0(H(x))|$  e, pelo Teorema do Valor Médio para Integrais (Larson, Hostetler & Edwards, 2001), existe  $\xi \in [H(x+\Delta x), H(x)]$ , tal que  $A_0' = f_0(\xi) \left[ \frac{|H(x+\Delta x) - H(x)|}{\Delta x} \right] \Delta x$ . O valor absoluto é necessário, pois  $H(x)$  é decrescente e, portanto,  $H'(x)$  é negativa.

Considerando esta partição, definida pela função particionante  $H(x)$ , as probabilidades de erros condicionais são dadas por

$$\alpha_x = \frac{A_0'}{A_0' + A_0} = \lim_{\Delta x \rightarrow 0} \frac{f_0(\xi) \left[ \frac{|H(x+\Delta x) - H(x)|}{\Delta x} \right]}{f_0(\xi) \left[ \frac{|H(x+\Delta x) - H(x)|}{\Delta x} \right] + f_0(x)} \Rightarrow$$

$$\alpha_x = \frac{f_0(H(x)) |H'(x)|}{f_0(H(x)) |H'(x)| + f_0(x)} \quad (3)$$

e

$$\beta_x = \frac{A_1}{A_1 + A_1'} = \lim_{\Delta x \rightarrow 0} \frac{f_1(x)}{f_1(x) + f_1(\xi) \left[ \frac{|H(x+\Delta x) - H(x)|}{\Delta x} \right]} \Rightarrow$$

$$\beta_x = \frac{f_1(x)}{f_1(x) + f_1(H(x)) |H'(x)|} \quad (4)$$

Então, o teste freqüentista condicional  $\Upsilon^C$ , utilizando (3) e (4), levando-se em consideração o valor observado  $z$  da estatística condicionante  $Z$ , e o valor crítico  $b_c$ , é dado pela estrutura geral apresentada a seguir:

$$\Upsilon^c = \begin{cases} \text{rejeita-se } H_0 \text{ e apresenta-se} \\ \text{PEC I } \alpha_z = \left( \frac{f_0(z)}{f_0(z) + f_0(H(z))|H'(z)|} \right), \text{ se } z \leq b_c; \\ \text{aceita-se } H_0 \text{ e apresenta-se} \\ \text{PEC II } \beta_z = \left( \frac{f_1(z)}{f_1(z) + f_1(H(z))|H'(z)|} \right), \text{ se } z > b_c. \end{cases} \quad (5)$$

O valor crítico  $b_c$  e as probabilidades dos erros  $\alpha_z$  e  $\beta_z$  apresentados para o teste podem ser diferentes para diferentes estatísticas condicionantes. A escolha da função particionante é totalmente arbitrária, mas nem todas as funções normalmente utilizadas fornecem bons condicionamentos. As principais formas de condicionamento são baseadas em uma estatística ancilar, nos denominados “níveis de significância intrínsecos” e no procedimento de partição “contínua com probabilidades de erros condicionais iguais” ( $\alpha_z = \beta_z, \forall z$ ) (Brighenti, 2007).

### 3 Teste unificado para hipóteses simples versus simples

Berger, Brown e Wolpert (1994), resgatando o trabalho desenvolvido por Kiefer (1977) em testes de hipóteses com probabilidades de erros condicionais, observaram que o teste freqüentista condicional  $\Upsilon^*$ , derivado de (5), poderia ser escrito como:

$$\Upsilon^* = \begin{cases} \text{rejeita-se } H_0 \text{ com PECI } \alpha_x = \frac{B(x)}{1+B(x)}, & \text{se } B(x) \leq 1; \\ \text{aceita-se } H_0 \text{ com PECII } \beta_x = \frac{1}{1+B(x)}, & \text{se } B(x) > 1, \end{cases} \quad (6)$$

utilizando a função particionante  $H(z) = [F_B^0]^{-1}(1 - F_B^1(z))$ , em que  $F_B^i(\cdot)$ ,  $i = 0, 1$  é a função distribuição acumulada da estatística unidimensional B, em que estas são funções diferenciáveis e inversíveis no domínio de B. Mas desta forma, este se torna semelhante ao teste Bayesiano  $\Upsilon_\pi$ , considerando perdas iguais e densidades *a priori* também iguais dado por:

$$\Upsilon_\pi = \begin{cases} \text{rejeita-se } H_0 \text{ com probabilidade a posteriori, } \alpha_x = \frac{B(x)}{1+B(x)}, & \text{se } B(x) \leq 1; \\ \text{aceita-se } H_0 \text{ com probabilidade a posteriori, } \beta_x = \frac{1}{1+B(x)}, & \text{se } B(x) > 1. \end{cases} \quad (7)$$

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  $\Upsilon^*$ , são numericamente iguais às probabilidades *a posteriori* encontradas no teste Bayesiano  $\Upsilon_\pi$ .

O teste  $\Upsilon^*$  apresenta probabilidades de erro do tipo I e tipo II iguais,  $\alpha = \beta$ , e probabilidades de erros condicionais dadas por  $\alpha_x$  e  $\beta_x$ , com valor máximo de  $\alpha_x$  ou de  $\beta_x$  sendo  $1/2$ , quando a região crítica é dada em função do valor crítico  $b_c = 1$ . Isto ocorre quando a distribuição de  $B(X)$  sob  $H_0$  é igual à distribuição de  $B(X)^{-1}$  sob  $H_1$ , ou seja,  $f_B^0(x) = f_B^1(x)$ ; neste caso, a estatística de razão de verossimilhanças  $B(X)$  é dita simétrica.

Note, então, que tanto um freqüentista quanto um Bayesiano, observados os dados  $x = (x_1, \dots, x_n)$ , tomariam a mesma decisão e apresentariam medidas de evidência numericamente iguais. É claro que a interpretação desta medida de evidência é diferente; ou seja, para o freqüentista 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,  $l_0 = l_1 = 1$ ).

O teste  $\Upsilon_\pi$  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  $b_c \neq 1$ . Tal procedimento justifica a abordagem utilizada por Berger, Brown e Wolpert (1994) para definir o teste unificado.

No contexto Bayesiano, há duas possibilidades de ocorrência para  $b_c \neq 1$ , a primeira quando as perdas não são simétricas,  $l_0 \neq l_1$ , e a segunda quando as densidades *a priori* são diferentes.

Para os casos em que as perdas não são simétricas e as densidades *a priori* são iguais, o teste Bayesiano fica da forma:

$$\Upsilon_\pi^* = \begin{cases} \text{rejeita-se } H_0 \text{ com probabilidade } a \text{ posteriori } \alpha_x = \frac{B(x)}{1+B(x)}, \\ \text{se } B(x) \leq b_c = \frac{l_1}{l_0}; \\ \text{aceita-se } H_0 \text{ com probabilidade } a \text{ posteriori } \beta_x = \frac{1}{1+B(x)}, \\ \text{se } B(x) > b_c = \frac{l_1}{l_0}. \end{cases} \quad (8)$$

Se, por exemplo, forem considerados os valores para perda  $l_1 = 3$  e  $l_0 = 1$ , e for observado  $B(x) = 2$ , o teste  $\Upsilon_\pi^*$  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, 2003).

Berger, Brown e Wolpert (1994) propuseram, então, uma modificação do teste  $\gamma_\pi^*$  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 perdas “0-1”, incorporando ao teste uma região em que não se toma decisão.

#### 4 A região de não-decisão

Uma modificação do teste  $\gamma_x^*$  que resolve a dificuldade, comentada no item anterior, que ocorre quando a PEC é maior que 1/2, é dada pela incorporação de uma “região de não-decisão”, abreviada por RND dada pelo intervalo  $]r, a[$ .

Considere 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,  $F_B^0(b_c) = 1 - F_B^1(b_c)$ , em que  $F_B^0$  e  $F_B^1$  são as funções distribuições acumuladas da razão de verossimilhanças  $B(x)$  sob  $f_0$  e  $f_1$ , respectivamente e  $b_c$  é o valor crítico do teste. 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. Assim, a partir do valor crítico  $b_c$ , a cada valor  $z$ , que fornece uma probabilidade de erro condicional  $\beta_z$ , é associado um  $H(z)$ , que fornece  $\alpha_{H(z)} = \beta_z$ , sendo apresentada 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  $b = 1$ .

Avaliando as duas situações possíveis para  $b_c$ , tem-se:

##### 4.1 O valor crítico é superior a 1 ( $b_c > 1$ ).

Quando o critério freqüentista é  $\alpha = \beta$ , tem-se a escolha de  $b_c$  dada pelo ponto que satisfaz a igualdade  $F_B^0(b_c) = 1 - F_B^1(b_c)$ , então, 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 dados por (3) e (4) e, portanto, no valor de  $b$  exatamente igual a 1, as probabilidades dos erros condicionais dados por  $\alpha_x$  e  $\beta_x$  são iguais a 1/2; com isso, entre 1 e  $b_c$ , estas probabilidades serão superiores a 1/2, devendo ser excluídas na tomada de decisão. Além disso, devido à construção do condicionamento, os pontos entre  $b_c$  e  $H(1)$  não terão o ponto correspondente para o par  $(z, H(z))$  que estão entre o valor 1 e o valor crítico, pois, estes são os pontos que

forneceriam probabilidades de erros condicionais superiores a 0,5, que foram excluídos (Figura 4).

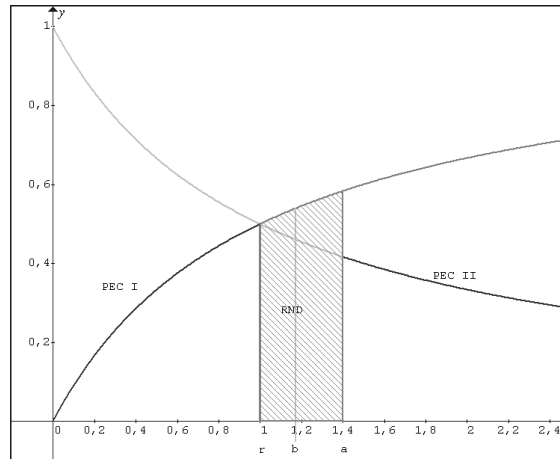


Figura 4 - Exemplo da região de não-decisão (RND) quando  $b_c > 1$ .

Então a RND é o intervalo  $]1, H(1)[$ , sendo que para  $b \leq 1$  apresenta-se PEC I igual a  $\alpha_z = \frac{B(z)}{1+B(z)}$  e para  $b \geq H(1)$  PEC II dada por  $\beta_z = \frac{1}{1+B(z)}$ . 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  $(1, H(1))$ . 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  $b$ . 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 5).

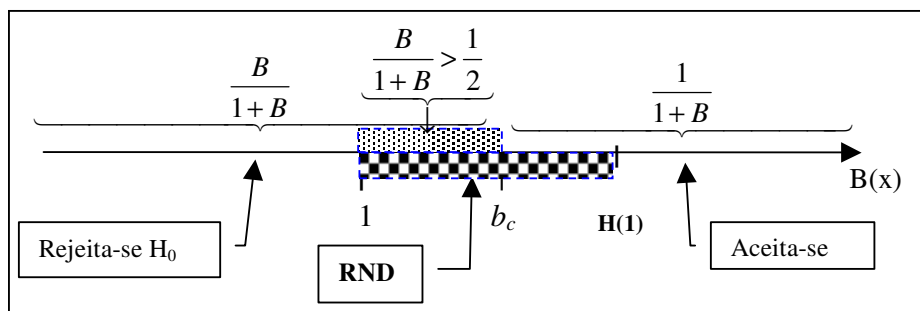


Figura 5 - Esquema da construção da Região de Não-Decisão para  $b_c > 1$ .



Tem-se, então, que  $F_b^1(r) = P_{\theta} [B(X) \leq r] = 1 - F_b^0(1)$ .

Portanto, a região de não-decisão fica definida da forma:

$$r=1 \text{ e } a = [F_b^0]^{-1}(1 - F_b^1(1)) \text{ se } F_b^0(1) \leq 1 - F_b^1(1),$$

$$r = [F_b^1]^{-1}(1 - F_b^0(1)) \text{ e } a = 1 \text{ se } F_b^0(1) > 1 - F_b^1(-1).$$

Ou, considerando a função particionante H, sob a distribuição da variável aleatória

B(X), tem-se  $H(b) = [F_b^0]^{-1}(1 - F_b^1(b))$ ,  $H^{-1}(b) = [F_b^1]^{-1}(1 - F_b^0(b))$ , e a RND é dada,

de forma simplificada, por:

$$r=1 \text{ e } a = H(1), \text{ se } H(1) \geq 1,$$

$$r = H^{-1}(1) \text{ e } a = 1, \text{ se } H(1) < 1.$$

E o teste  $\Upsilon_{\pi}^*$ , com a inclusão da RND fica:

$$\Upsilon_{\pi}^* = \begin{cases} \text{rejeita-se } H_0 \text{ e apresenta-se PEC I ou} \\ \text{probabilidade } a \text{ posteriori } \alpha_x = \frac{B}{1+B}, \text{ se } B(x) \leq r; \\ \text{não se toma nenhuma decisão,} & \text{se } r < B(x) < a; \\ \text{aceita-se } H_0 \text{ e apresenta-se PEC II ou} \\ \text{probabilidade } a \text{ posteriori } \beta_x = \frac{1}{1+B}, \text{ se } B(x) \geq a. \end{cases} \quad (9)$$

As discussões para construção da RND aqui apresentadas contemplam tanto os testes de hipóteses envolvendo modelos não-encaixados, isto é, modelos que apresentam formas funcionais completamente diferentes, quanto os modelos encaixados, sendo que, neste último caso, apenas o 2º caso em que  $b_c < 1$  ocorrerá, pois sendo esta uma variável B(X) que assume apenas valores menores que 1 (considerando o modelo mais simples na hipótese nula).

Berger, Boukai e Wong (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.

A presença da "região de não-decisão" é o problema no uso de  $\Upsilon_{\pi}^*$  é, fonte de críticas do teste. Mas ela é o preço que se deve pagar a fim de que medidas de evidência Bayesiana e freqüentista condicional fiquem limitadas a 1/2. 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  $\Upsilon_{\pi}^*$  é reduzida. Assim, o tamanho da região de não-decisão é uma característica particularmente importante para se 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, Brown e Wolpert, 1994).

A família de testes  $\Upsilon_{\pi}^*$  possui, simultaneamente, uma justificativa Bayesiana e uma freqüentista condicional, com a decisão e a medida de evidência 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 (Berger, 2003).

Seguem duas aplicações da construção da região de não-decisão, uma para o caso de razão de verossimilhanças simétrica e outro para não-simétrica. Lembrando que o teste unificado tem sua principal motivação baseada em testes em que cometer os erros tipo I e II têm a mesma importância, sendo considerada a condição inicial de erros não-condicionais iguais ( $\alpha = \beta$ ).

## 5 Razão de verossimilhanças simétrica: aplicação à distribuição beta

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:  $B(x) = \frac{1-x}{x}$ .

A distribuição de  $B(X)$ , obtida pelo teorema da transformação de variáveis (Mood et al., 1974),

$$f_B^0(b) = \left| \frac{d}{db} g^{-1}(b) \right| f_Y(g^{-1}(b)) \quad (10)$$

sob a hipótese nula é dada por:

$$f_B^0(b) = \frac{1}{(1+b)^2} \frac{[g^{-1}(b)]^1 [1-g^{-1}(b)]^2}{B(2,3)} = \frac{12}{(1+b)^3} \frac{b^2}{(1+b)^2} \Rightarrow \boxed{f_B^0(b) = 12b^2(1+b)^{-5}}.$$

E sob a hipótese alternativa, é:

$$f_B^1(b) = \frac{1}{(1+b)^2} \frac{[g^{-1}(b)]^2 [1-g^{-1}(b)]^1}{B(3,2)} = \frac{1}{(1+b)^2} \frac{4!(1+b)^{-2} [1-(1+b)^{-1}]^1}{1!2!} \Rightarrow$$

$$\boxed{f_B^1(b) = 12b(1+b)^{-5}}.$$

Como a distribuição de  $[B(X)]^1$  sob a hipótese alternativa é dada por:

$$f_{1/B}^1(b) = \left| \frac{d}{db} g^{-1}(b) \right| f_Y(g^{-1}(b)) = \frac{1}{b^2} 12 \left( \frac{1}{b} \right) \left( 1 + \left( \frac{1}{b} \right) \right)^{-5} = \frac{12}{b^3} \left( \frac{b+1}{b} \right)^{-5},$$

então,  $f_{1/B}^1(b) = 12b^2(b+1)^{-5} = f_B^0(b)$ .

Conclui-se que a razão de verossimilhanças é simétrica.



Neste trabalho, o principal enfoque é a 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:

$$F_B^0(b_c) = 1 - F_B^1(b_c). \quad (11)$$

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

Para o caso em que os erros são igualmente importantes, o valor crítico de  $b_c$  pode ser calculado pela igualdade seguinte.

$$\alpha = \beta \Rightarrow 1 - F_x^0(x) = F_x^1(x) \Rightarrow \int_{b_c}^1 \frac{(x)^{2-1} (1-x)^{3-1}}{B(2,3)} dx = \int_0^{b_c} \frac{(x)^{3-1} (1-x)^{2-1}}{B(3,2)} dx \Rightarrow b_c = 1.$$

O valor crítico de  $b_c = 1$  corresponde na desigualdade equivalente  $B(x) = \frac{1-x}{x}$  a  $x = 0,5$ . Então, o erro tipo I e II são dados por:

$$F_B^1(1) = 1 - F_B^0(1) = 31,25\% \text{ ou } \alpha = \beta = 0,3125 \text{ (Figura 8).}$$

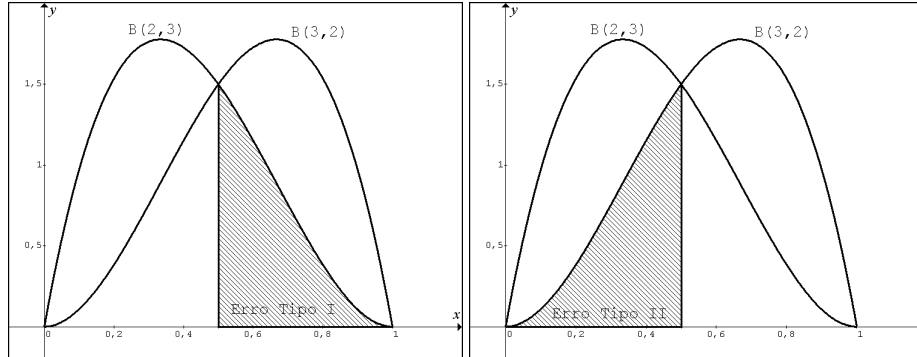


Figura 8 - Gráficos exibindo as áreas correspondentes às probabilidades de erros tipos I e II iguais a 31,25% para o caso de  $n = 1$ .

Assim, para o caso de razão de verossimilhanças simétrica, o valor crítico  $b_c = 1$  garante taxa de erro tipo I igual à taxa de erro tipo II, pois quando  $F_B^0(x) = F_{1/B}^1(x)$  tem-se  $F_B^0(1) = 1 - F_B^1(1)$  e neste caso a região de não-decisão é vazia. O teste fica então:

$$Y_{\pi}^* = \begin{cases} \text{rejeita-se } H_0 \text{ e apresenta-se PECl ou} \\ \text{probabilidade a posteriori } \alpha_x = \frac{B(x)}{1+B(x)}, & \text{se } B(x) \geq 1; \\ \text{aceita-se } H_0 \text{ e apresenta-se PEClI ou} \\ \text{probabilidade a posteriori } \beta_x = \frac{1}{1+B(x)}, & \text{se } B(x) < 1. \end{cases}$$

### 5.1 Generalização da RND para o caso de razão de verossimilhanças simétrica

Para o caso de razão de verossimilhanças simétricas, como o caso Beta feito anteriormente, vale a seguinte proposição:

Se a razão de verossimilhanças é simétrica, então a função particionante  $H(b) = \frac{1}{b}$  é solução de  $f_B^1(H(b))|H'(b)| = f_B^0(b)$ , ou seja, quando tem-se o critério  $\alpha = \beta$ .

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

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

se trata de razão de verossimilhanças simétrica,

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

**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,  $F_B^0(b) = 1 - F_B^1(b)$ , então,  $b = 1$ , isto é, o teste de razão de verossimilhanças é dado por: Rejeita-se  $H_0$  se  $b \leq 1$ .

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

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

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

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

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

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. E como  $b = 1$ , tem-se  $H(1) = 1$  o que torna a RND sempre vazia, independente do tamanho da amostra, pois note que esta prova é para a variável aleatória  $B$ , razão de verossimilhanças, e ela pode ser obtida a partir de qualquer tamanho de amostra.

## 6 Razão de verossimilhanças não-simétrica: aplicação à distribuição exponencial

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  $\alpha$  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:

$$B(x_1, \dots, x_n) = \frac{L_0}{L_1} = \frac{\theta_0^n \exp\left\{-\theta_0 \sum_{i=1}^n x_i\right\}}{\theta_1^n \exp\left\{-\theta_1 \sum_{i=1}^n x_i\right\}} = \left(\frac{\theta_0}{\theta_1}\right)^n \exp\left\{-\left(\theta_0 - \theta_1\right) \sum_{i=1}^n x_i\right\}. \quad (12)$$

A família de testes dada pela razão de verossimilhanças consiste em rejeitar  $H_0$  se

$$B(x_1, \dots, x_n) \leq b_c \Rightarrow \left(\frac{\theta_0}{\theta_1}\right)^n \exp\left\{-\left(\theta_0 - \theta_1\right) \sum_{i=1}^n x_i\right\} \leq b_c. \quad (13)$$

Para os objetivos deste trabalho, será conveniente calcular diretamente a distribuição da variável  $B(X_1, \dots, X_n) = \left(\frac{\theta_0}{\theta_1}\right)^n \exp\left\{-\left(\theta_0 - \theta_1\right) \sum_{i=1}^n X_i\right\}$ . Para o cálculo da densidade de  $B(X)$  utiliza-se o teorema de transformação de variáveis dado em (10).

Note que  $B = g(Y) = \left(\frac{\theta_0}{\theta_1}\right)^n \exp\left\{-\left(\theta_0 - \theta_1\right)Y\right\}$  é a exponencial de uma variável com distribuição Gama de parâmetros  $(n, \theta)$ . Expressando  $Y$  como função de  $B$ , tem-se que:

$$g^{-1}(b) = \frac{1}{(\theta_1 - \theta_0)} \ln \left[ \left(\frac{\theta_1}{\theta_0}\right)^n b \right] \Rightarrow \frac{d}{db} g^{-1}(b) = \frac{1}{(\theta_1 - \theta_0)b}, \quad (14)$$

$$f_Y(g^{-1}(b)) = \frac{\theta^n}{\Gamma(n)} \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[ \left( \frac{\theta_1}{\theta_0} \right)^n b \right] \right\}^{n-1} e^{-\theta \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[ \left( \frac{\theta_1}{\theta_0} \right)^n b \right] \right\}}. \quad (15)$$

Utilizando (14) e (15), em (10) tem-se:

$$f_B(b) = \frac{1}{(\theta_1 - \theta_0)b} \frac{\theta^n}{\Gamma(n)} \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[ \left( \frac{\theta_1}{\theta_0} \right)^n b \right] \right\}^{n-1} e^{-\theta \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[ \left( \frac{\theta_1}{\theta_0} \right)^n b \right] \right\}}. \quad (16)$$

Para estabelecer o valor crítico  $b_c$  para que se tenha um determinado valor de erro tipo I  $\alpha$ , é necessário resolver a equação integral

$$\int_0^{b_c} \frac{1}{(\theta_1 - \theta_0)b} \frac{\theta^n}{\Gamma(n)} \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[ \left( \frac{\theta_1}{\theta_0} \right)^n b \right] \right\}^{n-1} e^{-\theta \left\{ \frac{1}{(\theta_1 - \theta_0)} \ln \left[ \left( \frac{\theta_1}{\theta_0} \right)^n b \right] \right\}} db = \alpha. \quad (17)$$

Usando substituição de variáveis:

$$u = \frac{1}{(\theta_1 - \theta_0)} \ln \left[ \left( \frac{\theta_1}{\theta_0} \right)^n b \right] \quad b = 0 \Rightarrow u = -\infty$$

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

A equação integral (17) se reduz a:

$$\frac{1}{(\theta_1 - \theta_0)} \ln \left[ \left( \frac{\theta_1}{\theta_0} \right)^n b \right] \int_0^{\frac{1}{(\theta_1 - \theta_0)} \ln \left[ \left( \frac{\theta_1}{\theta_0} \right)^n b \right]} \frac{\theta^n}{\Gamma(n)} u^{n-1} e^{-\theta u} du = \alpha \quad (18)$$

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  $w = \theta_0 u$ :

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

$$F_B^0(b) = \int_0^{\frac{\theta_0}{(\theta_1 - \theta_0)} \ln \left[ \left( \frac{\theta_1}{\theta_0} \right)^n b \right]} \frac{1}{\Gamma(n)} w^{n-1} e^{-w} dw \quad (19)$$

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; então,

$$F_B^0(b) = P_{\theta=\theta_0}(B(X) \leq b) = \Gamma \left( n, \frac{\theta_0}{(\theta_1 - \theta_0)} \ln \left[ \left( \frac{\theta_1}{\theta_0} \right)^n b \right] \right). \quad (20)$$

De maneira análoga, obtém-se para  $\theta = \theta_1$ ,

$$F_B^1(b) = P_{\theta=\theta_1}(B(X) \leq b) = \Gamma \left( n, \frac{\theta_1}{(\theta_1 - \theta_0)} \ln \left[ \left( \frac{\theta_1}{\theta_0} \right)^n b \right] \right). \quad (21)$$

Para o caso particular de  $H_0: \theta = 1$  versus  $H_1: \theta = 2$ , (12) fica da forma:

$$B(x_1, \dots, x_n) = \left( \frac{1}{2} \right)^n \exp \left\{ \sum_{i=1}^n x_i \right\}. \quad (22)$$

Então, de (20) e (21) tem-se:

$$F_B^0(b) = P_{\theta=1}(B(X) \leq b) = \Gamma \left( n, \ln [2^n b] \right), \quad (23)$$

$$F_B^1(b) = P_{\theta=2}(B(X) \leq b) = \Gamma \left( n, 2 \ln [2^n b] \right). \quad (24)$$

Para obter erros iguais, satisfazendo (11) no caso de, por exemplo, uma amostra de tamanho 4, o valor de  $b_c$  é 0,797, o que fornece probabilidade de erro de 25,2% (Figura 9).

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

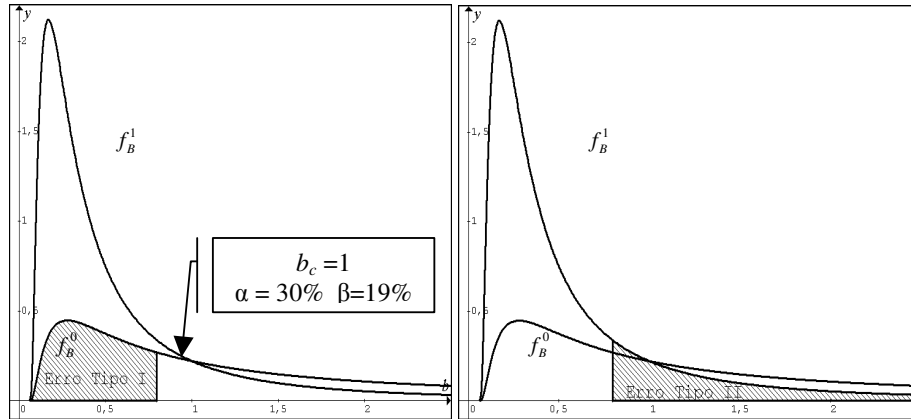


Figura 9 - Gráficos exibindo as áreas correspondentes às probabilidades de erros tipos I e II iguais a 25,2% para o caso de  $n = 4$ .

A seguir apresenta-se o teste  $\Upsilon_{\pi}^*$  com a inclusão da RND para diferentes tamanhos de amostra do teste de hipóteses do parâmetro da distribuição exponencial.

### 6.1 Análise da RND via dados simulados: caso não-simétrico

Para construir a "região de não-decisão" foram considerados os tamanhos amostrais:  $n = 1$ ;  $n = 4$ ;  $n = 10$ ;  $n = 30$  e  $n = 100$ .

Simulou-se o procedimento de obtenção de amostras da distribuição Exponencial (1) versus Exponencial (2) de acordo com uma rotina desenvolvida no programa SAS em Brighenti (2007) verificando o número de vezes em que não houve tomada de decisão. Inicialmente um número aleatório uniforme  $u \sim U(0,1)$  foi gerado. Caso  $u$  fosse menor que  $\frac{1}{2}$ , gerava-se uma amostra de tamanho  $n$  da distribuição Exponencial (1); se  $u$  fosse maior que  $\frac{1}{2}$ , a amostra era gerada da distribuição Exponencial (2). Utilizou-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.

Considerou-se, para o condicionamento, a função particionante  $H^{-1}(z) = [F_B^1]^{-1}(1 - F_B^0(z))$ . Conforme apresentado na Tabela 1, para encontrar o valor crítico  $b_c$  e os limites da RND em cada caso utilizou-se o software Maple 10. Lembrando que, 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:

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

Então, calculou-se, para cada tamanho de amostra  $n$ , a respectiva taxa de erros Tipo I e II, satisfazendo a condição  $\alpha = \beta$ . Como esperado, as taxas de erros são menores à medida que o tamanho de amostra aumenta. Na Tabela 1 são apresentados também os

valores críticos em função da desigualdade equivalente  $\sum_{i=1}^n x_i = \frac{1}{(\theta_1 - \theta_0)} \ln \left[ b_c \left( \frac{\theta_1}{\theta_0} \right)^n \right]$ .

Tabela 1 - 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 ]r, 1[
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

Na Tabela 1 são apresentados os valores críticos  $b_c$  no teste não-condicional. Como,  $F_B^0(1) > 1 - F_B^1(1)$  os limites da região de não-decisão são dados por  $r = [F_B^1]^{-1}(1 - F_0(1))$  e  $a = 1$  e a RND = ]r, 1[, bastando calcular apenas o limite inferior da RND para cada caso.

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 2).

Tabela 2 - 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 qualidades desejáveis como o controle simultâneo das taxas de erros tipo I e tipo II.

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.

## Conclusões

O teste unificado apresentado não acrescenta maiores dificuldades matemáticas ao teste tradicional.

Com a utilização da função particionante dada por  $H(z) = [F_B^0]^{-1}(1 - F_B^0(z))$  o cálculo da região de não decisão do teste unificado é relativamente simples.

O exemplo com a distribuição Beta, em que ocorre razão de verossimilhanças simétrica, ilustra o caso em que a região de não decisão é vazia.

No exemplo com a distribuição Exponencial, envolvendo razão de verossimilhanças não-simétrica, a probabilidade de ocorrência da região de não-decisão decresce rapidamente com o tamanho da amostra. Além disso, o teste controla bem os erros tipos I e II, implicando assim que o teste tem uma boa performance e pode ser recomendado para uso prático.

BRIGHENTI, C. R. G.; CHAVES, L. M. No-decision region in tests with frequentists conditional and bayesian interpretation. *Rev. Bras. Biom.*, São Paulo, v.25, n.4, p.23-47, 2007.

- **ABSTRACT:** *The Bayesian Test is a Likelihood Ratio Test with critical values given by product of the ratio between prior densities and the ratio between losses. In this way, for some determined values of  $p$  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  $p$  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. In this work, it was studied size of the NDR to hypothesis test of the parameter Exponential distribution with samples of different sizes. It was concluded that the size of the NDR depends on the size of the sample and that some situations can make the test unified inadequate.*
- **KEYWORDS:** *Conditional error probability; no-decision region; partitioning function; posterior probability.*



## Referências

- BERGER, J. O. *Statistical decision theory and Bayesian analysis*. 2nd. ed. New York: Springer-Verlag, 1985.618p.
- BERGER, J. O.; BROWN, L. D.; WOLPERT, R. L. A unified conditional frequentist and Bayesian test for fixed and sequential simple hypothesis testing. *Ann. Stat.*, Hayward, v.22, n.4, p.1787-1807, 1994.
- BERGER, J.; BOUKAI, B.; WONG, Y. Unified frequentist and Bayesian testing of a precise hypothesis. *Stat. Sci.*, Hayward, v.12, n.3, p.133-148, 1997. (Com discussão: p.149-160)
- BERGER, J. O. Could Fisher, Jeffreys and Neyman have agreed on testing? *Stat. Sci.*, Beachwood, v.18, n.1, p.1-12, 2003. (Com discussão: p.13-32)
- BRIGHENTI, C. R. G. *Testes freqüentistas condicionais e testes com interpretação bayesiana e freqüentista condicional*. 2007. 212f. Tese (Doutorado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, 2007.
- GARTHWAITE, P. H.; JOLLIFFE, I. T.; JONES, B. *Statistical inference*. London: Prentice Hall, 1995. 291p.
- HUBBARD, R.; BAYARRI, M. J. Confusion over measures of evidence (p's) versus errors ( $\alpha$ 's) in classical statistical testing. *Am. Stat.*, Alexandria, v.57, n.3, p.171- 182, 2003.
- HUBBARD, R.; ARMSTRONG, J. S. *Why we don't really know what "statistical significance" means: a major educational failure*. 2005. Disponível em: < <http://www.wharton.upenn.edu> >. Acesso em: 21 mar. 2007.
- KIEFER, J. Conditional confidence statements and confidence estimators. *J. Am. Stat. Assoc.*, Alexandria, v.72, n.360, p.789-807, 1977. (Com discussão: p.808-827).
- KILLEN, P. R. An alternative to null-hypothesis significance tests. *Psychol. Sci.*, Oxford, v.16, n.5, p.345-353, 2005.
- LARSON, R. E.; HOSTETLER, R. P.; EDWARDS, B. H. *Cálculo com aplicações* 4. ed. Rio de Janeiro: LTC, 2001. 711p.
- MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. *Introduction to the theory of statistics*. 3<sup>rd</sup>. ed. Singapore: McGraw-Hill International, 1974. 480p.
- NEYMAN, J.; PEARSON, E. S. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Trans. R. Soc.*, Ser. A, London, v.231, p.289-337, 1933.
- STERNE, J. A. C.; SMITH, G. D. Sifting the evidence - what's wrong with significance tests? *Br. Med. J.*, London, v.322, n.7280, p.226-231, 2001.
- TONHASCA JÚNIOR, A. The "magic limiar" and the test power. *Ciê. Cult.*, São Paulo, v.43, n.6, p.417- 422, 1991.

Recebido em 10.09.2007.

Aprovado após revisão em 12.02.2008.