



Universidade de Brasília
IE – Departamento de Estatística

Automatização da criação de regiões de credibilidade HPD e do Teste de Significância Genuinamente Bayesiano – FBST

Lucas Barbosa Cusinato Rodrigues

Orientador: Professor Doutor Eduardo Yoshio Nakano

Julho de 2014

1. Sumário

2.	Resumo.....	3
3.	Introdução.....	4
4.	Inferência Bayesiana	5
	Intervalos e regiões críveis.....	5
	FBST	5
5.	O algoritmo automático	7
	Regiões de credibilidade HPD	7
	Máximo <i>a posteriori</i>	8
	Teste de Significância Genuinamente Bayesiano (FBST).....	9
	Simulando sob H_0	10
6.	Exemplos de aplicações.....	11
	Teste de uma proporção	11
	Teste de homogeneidade em tabelas 2x2	12
	Lei do equilíbrio de Hardy-Weinberg	13
	Teste de independência em uma tabela de contingência 2x2.....	15
	Comparação de duas distribuições Gama	16
7.	Conclusão	18
8.	Referências Bibliográficas	19

2. Resumo

Propõe-se, neste trabalho, uma automatização do cálculo da medida de evidência de hipóteses precisas pelo procedimento FBST. Esse procedimento foi apresentado inicialmente no artigo *Evidence and Credibility: Full Bayesian Significance Test for Precise Hypotheses* (PEREIRA e STERN, 1999). A aplicabilidade do algoritmo proposto foi comprovada após a comparação de seus resultados sobre os exemplos do artigo.

3. Introdução

O problema de testar uma hipótese composta num cenário bayesiano é conceitualmente mais simples que no clássico. Atendendo à interpretação probabilística direta das hipóteses em confronto, não se tem mais do que calcular as respectivas probabilidades *a posteriori* e optar por uma delas em função de algum critério baseado em sua grandeza relativa. No entanto, esse esquema bayesiano de testar hipóteses fica inviabilizado no caso de uma hipótese simples, caso frequentemente abordado do tipo $H_0: \theta = \theta_0$ vs $H_a: \theta \neq \theta_0$, em que o espaço paramétrico é um conjunto infinito não enumerável.

Com isso, surgiram outras formas para o teste de hipóteses precisas (hipótese com dimensão menor que a dimensão do espaço paramétrico) como, por exemplo, o Teste de Significância Genuinamente Bayesiano (FBST - *Full Bayesian significance Test*) proposto por Pereira e Stern (1999). O FBST testa as hipóteses precisas baseadas no cálculo da probabilidade *a posteriori* da região HPD (*Highest Posterior Density*) tangente¹ ao conjunto que define a hipótese nula. O FBST é um teste intuitivo e de fácil caracterização geométrica e o mesmo pode ser implementado por meio de técnicas de otimização e integração numérica.

O FBST é um procedimento relativamente novo e ainda não é disponibilizado pelos principais softwares estatísticos. Nesse contexto, o objetivo deste trabalho é apresentar um novo algoritmo, devidamente implementado no software R, (R Development Core Team, 2012) que automatize a construção da região HPD e calcule o e-valor (valor de evidência) do FBST. O R possui algumas bibliotecas que fornecem as regiões de credibilidade HPD em algumas situações específicas. Atualmente esses programas fornecem os resultados considerando que a distribuição *a posteriori* é bem comportada (unimodal, por exemplo). O procedimento proposto visa também acabar com essa restrição, apresentando regiões de credibilidade HPD para distribuições *a posteriori* multimodais.

¹ Uma região HPD T é dita tangente ao conjunto A se contiver o ponto de máxima densidade de A . Ela é composta pelos pontos do espaço paramétrico que são mais consistentes (de maior densidade *a posteriori*) com os dados observados do que H_0 . (PEREIRA e STERN, 1999)

4. Inferência Bayesiana

A Inferência Estatística utiliza o conceito de probabilidade para modelar quantidades desconhecidas de interesse (θ). No paradigma bayesiano, todas as quantidades desconhecidas possuem natureza estocástica, sendo a probabilidade interpretada como uma medida de incerteza que associamos a cada conjunto possível de valores para essas quantidades.

Visando aumentar a informação disponível sobre θ , definida em sua distribuição de probabilidades *a priori* $\pi(\theta)$, observamos empiricamente uma quantidade X a ela associada. A forma dessa associação é definida pela verossimilhança do modelo adotado, $f(X|\theta)$, e a regra de Bayes é então utilizada para atualizar essa informação, o que leva à distribuição de probabilidades *a posteriori*:

$$\pi(\theta|X) = \frac{g(x, \theta)}{m(x)} = \frac{f(x|\theta)\pi(\theta)}{m(x)} = \frac{f(x|\theta)\pi(\theta)}{\int g(x, \theta)d\theta} \propto f(x|\theta)\pi(\theta),$$

em que $m(x)$ denota a densidade marginal de X ; $f(X|\theta)$ a função de verossimilhança e $g(x, \theta)$ a densidade conjunta de X e θ . As inferências bayesianas serão realizadas exclusivamente pela distribuição *a posteriori* de θ .

Intervalos e regiões críveis

As estimativas intervalares no paradigma bayesiano também são baseadas na distribuição de probabilidade *a posteriori*. Considera-se (a, b) um intervalo de credibilidade $C = 100 * (1 - \alpha)\%$ para uma quantidade desconhecida θ se

$$P(a < \theta < b|X) = C$$

Uma solução simples e a mais regularmente adotada é o cálculo de a e b como os quantis de ordem $\alpha/2$ e $1 - \alpha/2$ respectivamente. Para distribuições assimétricas, esse método pode resultar em intervalos muito compridos. Na distribuição *Gama*(2, 5), por exemplo, os percentis p_5 e p_{95} levam ao intervalo (0,07; 0,95), com 90% de credibilidade e 0,88 de comprimento. Já o intervalo (0,02; 0,79) possui a mesma credibilidade, mas somente 0,77 de comprimento.

Podemos buscar, então, a menor região possível, fixada a credibilidade desejada. Pode-se mostrar que esse intervalo contém os valores de θ que possuem maior densidade *a posteriori*. Denominado HPD, ele é definido como

$$R = \{\theta \in \Theta: \pi(\theta|X) \geq \pi\}, \text{ com } P(\theta \in R|X) = C$$

Apesar de estar definida pela constante π , costumamos determinar R variando os limites de intervalos de credibilidade C até encontrar o de menor comprimento possível, segundo a precisão numérica adotada. Essa metodologia exige que a região crível seja determinada por um intervalo e inviabiliza sua obtenção para casos de distribuições a

posteriori multimodais ou multidimensionais. O método de obtenção de R via Monte Carlo proposta na seção seguinte segue diretamente da definição do HPD e supera essas duas grandes limitações do método mais tradicional.

FBST

Testar hipóteses compostas no paradigma bayesiano é conceitualmente mais simples do que no clássico. Em geral, basta optar por uma das hipóteses tomando como base suas respectivas probabilidades *a posteriori*. Já para hipóteses precisas (*e.g.*, $H_0: \theta = \theta_0$ vs $H_a: \theta \neq \theta_0$) esse esquema fica inviabilizado, visto que elas possuem probabilidade igual a zero. Nesse contexto, surgiu o Teste de Significância Genuinamente Bayesiano (FBST), proposto em (PEREIRA e STERN, 1999). No teste FBST, associamos uma medida de evidência à hipótese testada e, caso essa medida seja grande, optamos por essa hipótese.

Para o cálculo da evidência da hipótese, considere:

- $T_\varphi = \{\theta \in \Theta \mid \pi(\theta|X) > \varphi\}$ a região do espaço paramétrico cujos elementos possuem densidade a posteriori maior que φ ;
- $k = \int_{T_\varphi} \pi(\theta|X) d\theta$ a credibilidade dessa região;
- f^* o valor máximo da posteriori sobre a hipótese nula H_0 , alcançado com o argumento θ^* :

$$\theta^* \in \arg \max_{\theta \in \Theta_0} f(\theta), \quad f^* = f(\theta)$$

Então $T^* = T_{f^*}$ é o conjunto tangente à hipótese nula, de credibilidade k^* . A medida de evidência de H_0 é definida como a probabilidade do conjunto complementar a T^* :

$$Ev(H_0) = 1 - k^*$$

5. O algoritmo automático

Para construir o algoritmo proposto, considere as seguintes notações: $\pi_i = \pi(\theta_i|X)$, $i = 1, 2, \dots, M$ é o conjunto de pontos simulados da posteriori avaliados na densidade e $\pi_{(1)} \leq \pi_{(2)} \leq \dots \leq \pi_{(M)}$ são as estatísticas de ordem desse conjunto.

Regiões de credibilidade HPD

Da definição, sabemos que HPD é uma região cujo bordo é um conjunto de pontos em que a densidade *a posteriori* assume o mesmo valor. No caso bivariado, por exemplo, é limitada por uma curva de nível de $\pi(\theta|X)$. Isso implica que, para determinar o HPD, basta escolher qual será a imagem $\pi(\theta_a|X)$ dessa fronteira. A credibilidade real desse HPD poderá ser aproximada via Monte Carlo pelo número de pontos simulados cujo valor da densidade é maior que $\pi(\theta_a|X)$ dividido pelo total de simulações (M):

$$1 - \alpha \approx \frac{1}{M} \sum I(\pi_i > \pi(\theta_a|X))$$

Então, se tomarmos a altura na fronteira como $\pi(\theta_a|X) = \pi_{(k)}$, a credibilidade da região $\{\theta \in R^p \mid \pi(\theta|X) > \pi_{(k)}\}$ será aproximada por

$$\frac{1}{M} \sum_{i=1}^M I(\pi_i > \pi_{(k)}) = 1 - \frac{k}{M}$$

Portanto, para uma credibilidade fixada em $(1 - \alpha) * 100\%$, devemos escolher $\pi_{(k)}$ de modo que

$$1 - \frac{k}{M} = 1 - \alpha \Leftrightarrow k = \alpha * M$$

Repare que, no cálculo do HPD, só precisamos das estatísticas de ordem das observações da *posteriori*, que também podem ser obtidas por meio de uma função proporcional a $\pi(\theta|X)$. Isso porque se $\pi_{(1)} < \pi_{(2)} < \dots < \pi_{(M)}$ então $c\pi_{(1)} < c\pi_{(2)} < \dots < c\pi_{(M)}$, $\forall c \in R^+$.

Considere os seguintes dados sobre o tempo de falha (horas) de 15 filmes isolantes: 50,25 87,75 87,76 87,77 92,90 92,91 95,96 108,30 108,30 117,90 123,90 124,30 129,70 135,60 135,60. Ao ajustar um modelo *Weibull* a esses dados, temos interesse estimar seus parâmetros de forma e escala.

Ao invés de calcularmos intervalos HPD marginalmente para cada parâmetro, podemos obter uma região de credibilidade conjunta. A Figura 1 apresenta a região de 90% de credibilidade conjunta (área vermelha) obtida pelo método descrito acima e a tradicional (azul) obtida pela aproximação de Bonferroni, com 95% de credibilidade marginal HPD em cada parâmetro. O percentual de pontos simulados (10^6) dentro dos

limites de Bonferroni foi de 90,53%; sendo essa, portanto, a credibilidade correta dessa região retangular.

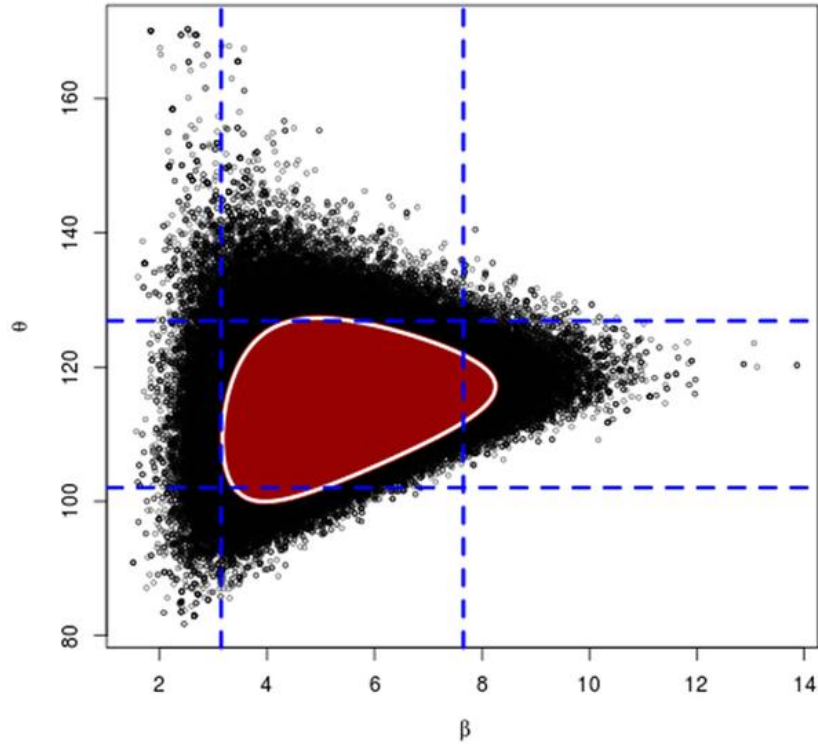


Figura 1: Regiões de credibilidade HPD obtidas por diferentes metodologias.

Máximo *a posteriori*

Ao minimizar a função perda 0-1, definida como

$$L_{01} = \begin{cases} 1, & \text{se } |d - g(\theta)| > \epsilon \\ 0, & \text{se } |d - g(\theta)| \leq \epsilon \end{cases}$$

obtemos o estimador pontual bayesiano, que corresponde ao ponto médio do intervalo de credibilidade HPD de comprimento 2ϵ . Pode-se mostrar que esse estimador converge para o ponto de máxima densidade posteriori (MAP) quando $\epsilon \rightarrow 0$. Então, para maximizar a distribuição *a posteriori*, basta fazer um intervalo HPD com o comprimento tendendo a zero.

Logo, podemos aproximar o MAP via Monte Carlo com o algoritmo apresentado na seção anterior se tomarmos o limite

$$\epsilon \rightarrow 0 \Leftrightarrow (1 - \alpha) \rightarrow 0 \xleftrightarrow{k=\alpha*M} k = M$$

O que leva à conclusão de que a aproximação do MAP corresponde ao valor simulado de maior densidade posterior avaliada, $\pi_{(M)}$. Note que essa aproximação também pode ser encontrada apenas com o núcleo da densidade posterior.

Teste de Significância Genuinamente Bayesiano

Se C é a credibilidade da região HPD tangente ao conjunto definido por H_0 , a evidência a favor dessa hipótese é

$$ev = 1 - C$$

Podemos aproximar o valor de C via simulações Monte Carlo e o problema se reduz a obter esse conjunto tangente, ou, equivalentemente, o ponto de máxima densidade posterior sob a hipótese H_0 .

Em geral, esse processo de maximização pode ser realizado por meio de algoritmos de otimização de uso geral, que normalmente exigem adaptações na função a ser maximizada. É muito comum, por exemplo, termos de aplicar a transformação logarítmica (entre outras reparametrizações) em um parâmetro com o intuito de permitir sua variação ao longo de toda a reta para satisfazer as condições de algum desses procedimentos de maximização. Além disso, há casos de hipóteses precisas não lineares como a do equilíbrio de Hardy-Weinberg, $\theta_3 = (1 - \sqrt{\theta_1})^2$, em que maximizar a densidade posterior não pode ser feito diretamente por esses algoritmos.

Todo esse cuidado necessário para a maximização tornava o cálculo da evidência da hipótese um processo muito artesanal. Uma alternativa mais automatizada a esse processo seria simular observações da distribuição *a posteriori* sob H_0 e aproximar seu máximo (MAP_{H_0}) via Monte Carlo como visto no tópico anterior.

Porém, ao invés de simular diretamente da distribuição posterior condicionada em H_0 , a alternativa encontrada foi “ajustar” pontos simulados da *posteriori* por um procedimento numérico. Em cada um desses valores simulados vamos basicamente utilizar as componentes dependentes (ou fixadas) dos valores simulados como chute inicial para chegar a valores que satisfaçam H_0 .

Simulando sob H_0

Considere particionar o vetor de parâmetros da distribuição posterior em dois subvetores, $\boldsymbol{\theta}_{\text{IND}}$ e $\boldsymbol{\theta}_{\text{DEP}}$, de modo que $\boldsymbol{\theta}_{\text{DEP}}$ contenha as dimensões que se reduzem ao condicionar em H_0 . Para a execução do algoritmo aqui proposto, é preciso definir um sistema de equações cujas variáveis sejam as componentes de $\boldsymbol{\theta}_{\text{DEP}}$. Além disso, H_0 deve ser o conjunto solução desse sistema e as componentes de $\boldsymbol{\theta}_{\text{IND}}$ consideradas constantes.

A título de ilustração, considere uma análise de regressão em que se tenha interesse de testar a hipótese $\{\beta_1 = \beta_2 = \log(\beta_3)\}$. Essa hipótese reduz duas dimensões no espaço paramétrico, sendo necessárias duas variáveis para $\boldsymbol{\theta}_{\text{DEP}}$. Nesse caso, definimos o seguinte sistema nas variáveis β_1 e β_2

$$\begin{cases} \beta_1 - \beta_2 = 0 \\ \beta_1 - \log(\beta_3) = 0 \end{cases}$$

Apesar de a escolha de fixar $\boldsymbol{\theta}_{\text{IND}} = \beta_3$ ser arbitrária, é interessante por tornar esse um sistema de equações lineares.

Após a definição de um sistema de equações adequado, vamos fixar $\boldsymbol{\theta}_{\text{IND}i}$ e utilizar $\boldsymbol{\theta}_{\text{DEP}i}$ ($i = 1, 2, \dots, M$) como chute inicial para buscar uma raiz do sistema de equações (em $\boldsymbol{\theta}_{\text{DEP}}$) utilizando o método de Newton-Raphson para solução numérica de sistemas não lineares. Ao atualizar $\boldsymbol{\theta}_{\text{DEP}i}$, o ponto passa a pertencer à região H_0 de acordo com a precisão numérica adotada. Esse procedimento é equivalente a determinar as dimensões dependentes ($\boldsymbol{\theta}_{\text{DEP}}$) por meio das independentes ($\boldsymbol{\theta}_{\text{IND}}$) juntamente com a hipótese (H_0). O elevado número de diferentes chutes iniciais $\boldsymbol{\theta}_{\text{DEP}}$ próximos a pontos de alta probabilidade *a posteriori* é uma indicação de que esse algoritmo de fato resulte em uma boa aproximação para o MAP_{H_0} , mas essa convergência para o máximo global ainda não foi provada formalmente.

Seja g o núcleo da densidade *a posteriori* e ${}_M\boldsymbol{\theta}_p$ a matriz com as M simulações dos p parâmetros de g . Então o algoritmo para o FBST pode ser formulado como:

- 1) Definimos o sistema de equações (H_0) e o vetor indicador das dimensões a se reduzir (Dep);
- 2) para cada linha $\boldsymbol{\theta}_i$ de $\boldsymbol{\theta}$, armazenamos na matriz $\boldsymbol{\theta}_{H_0}$ a solução² do sistema H_0 nas variáveis $\boldsymbol{\theta}_{\text{DEP}}$, fixado $\boldsymbol{\theta}_{\text{IND}}$;
- 3) obtemos então o vetor $\boldsymbol{\pi}_{H_0}$ avaliando g em cada linha de $[\boldsymbol{\theta}_{H_0}, \boldsymbol{\theta}_{\text{IND}}]$;
- 4) sendo j a posição do maior valor de $\boldsymbol{\pi}_{H_0}$, definimos $\text{MAP}_{H_0} = [\boldsymbol{\theta}_{H_0}, \boldsymbol{\theta}_{\text{IND}}]_j$;
- 5) por fim, o e-valor será aproximado pela proporção de pontos simulados situados fora do conjunto tangente:

$$ev = \frac{1}{M} \sum_{k=1}^M I_{[0; g(\text{MAP}_{H_0})]}(\boldsymbol{\pi}_k)$$

² Obtida pelo algoritmo de Newton-Raphson.

6. Exemplos de aplicações

O algoritmo automático foi aplicado aos cinco problemas apresentados em (PEREIRA e STERN, 1999) para comparar seus resultados. Os quatro primeiros dizem respeito à distribuição multinomial e o quinto à comparação de modelos Gama. No artigo original, foram utilizadas distribuições *a priori* de referência de modo que as distribuições *a posteriori* fossem proporcionais às funções de verossimilhanças, integráveis no espaço paramétrico.

Em ordem de complexidade da hipótese, as situações são:

1. Teste bilateral para proporção;
2. homogeneidade em tabelas de contingência 2x2;
3. equilíbrio de Hardy-Weinberg;
4. independência em tabelas de contingência 2x2;
5. comparação de distribuições Gamas.

A seguir a tradução livre desses problemas e seus resultados comparados. As dimensões fixadas como dependentes na execução do algoritmo estão em vermelho.

Teste de uma proporção

Sejam $D \sim \text{Binomial}(20, \theta)$, $\Theta = [0,1]$ o espaço paramétrico, d o número de sucessos observados e $H_0 = \{\theta \in \Theta \mid \theta = 0,5\}$ a hipótese a ser testada. O sistema de equações de solução H_0 utilizado foi

$$\theta - 0,5 = 0$$

As medidas de evidência (e-valores) obtidas para cada número de sucessos possível estão apresentadas na Tabela 1.

d	Evidência da hipótese	
	Pereira e Stern	Algoritmo automático
0	0,00	0,000
1	0,00	0,000
2	0,00	0,000
3	0,00	0,001
4	0,01	0,004
5	0,02	0,018
6	0,06	0,061
7	0,16	0,162
8	0,35	0,352
9	0,64	0,642
10	1,00	1,000

Tabela 1: Evidência de H_0 para $d=0,1,\dots,10$.

Teste de homogeneidade em tabelas 2x2

Duas amostras são retiradas de duas populações binomiais visando testar quando ambas as taxas de sucesso são iguais. Sejam x e y o número de sucessos em cada amostra de vinte (20) elementos. A densidade a *posteriori* para esse modelo multinomial é $f(\theta | x) \propto \theta_1^x \theta_2^{n-x} \theta_3^y \theta_4^{m-y}$, em que $\theta_1 = 1 - \theta_2$ e $\theta_3 = 1 - \theta_4$ representam a probabilidade de sucesso de cada população.

O espaço paramétrico e o conjunto da hipótese nula são

$$\Theta = \{0 \leq \theta \leq 1 \mid \theta_1 + \theta_2 = 1 \wedge \theta_3 + \theta_4 = 1\}$$

$$\Theta_0 = \{\theta \in \Theta \mid \theta_1 = \theta_3\}$$

O sistema H_0 utilizado foi

$$\theta_1 - \theta_3 = 0$$

As medidas de evidência obtidas estão apresentadas na Tabela 2. Alguns valores simulados (quando $x=4$ e $y=10$) estão apresentados na Figura 2, a variação promovida pelo algoritmo sobre cada ponto simulado está indicada pelas setas em vermelho. Note que essa variação somente ocorre na dimensão definida como dependente, θ_3 .

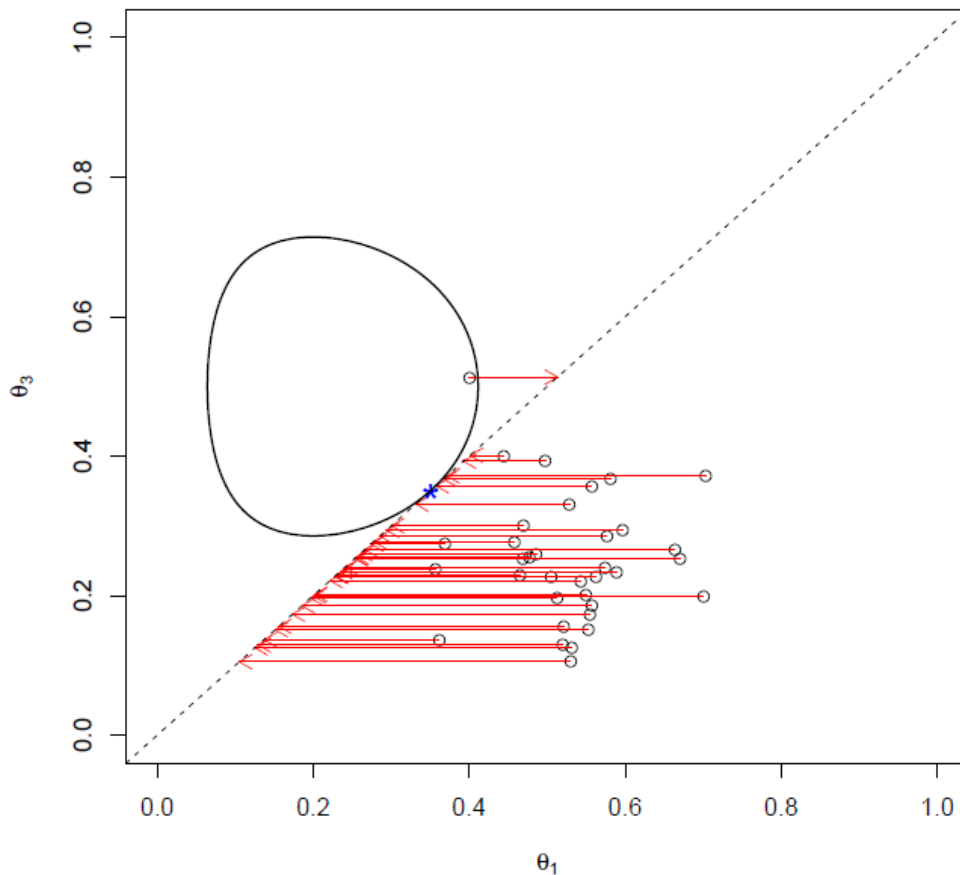


Figura 2: Teste de homogeneidade com $x=10$, $y=4$ e $n=m=20$. Região HPD tangente (linha cheia), hipótese nula (linha pontilhada), máximo a posteriori sob H_0 (em azul) e exemplos de valores simulados antes e depois do algoritmo.

Lei do equilíbrio de Hardy-Weinberg

Nessa aplicação em Biologia há uma amostra de $n = 20$, em que x_1 e x_3 são as contagens de elementos homozigotos e $x_2 = n - x_1 - x_3$ a contagem de elementos heterozigotos na amostra. $\theta = [\theta_1, \theta_2, \theta_3]$ é o vetor de parâmetros com a probabilidade de ocorrência de cada um dos três possíveis fenótipos. A densidade *a posteriori* para esse modelo trinomial é

$$f(\theta | x) \propto \theta_1^{x_1} \theta_2^{x_2} \theta_3^{x_3}$$

O espaço paramétrico e o conjunto da hipótese nula são

$$\Theta = \{ \theta \geq 0 \mid \theta_1 + \theta_2 + \theta_3 = 1 \}$$

$$\Theta_0 = \{ \theta \in \Theta \mid \theta_3 = (1 - \sqrt{\theta_1})^2 \}$$

O sistema de equações utilizado:

$$H_0: \sqrt{\theta_1} + \sqrt{\theta_3} - 1 = 0$$

A Tabela 2 contém os e-valores para diferentes possíveis valores de x_1 e x_2 obtidos pelo algoritmo aqui proposto e os apresentados originalmente.

		Homogeneidade		Equilíbrio de Hardy-Weinberg			
X	Y	Evidência da hipótese		X ₁	X ₃	Evidência da hipótese	
		Pereira e Stern	Algoritmo automático			Pereira e Stern	Algoritmo automático
5	0	0,05	0,045	1	2	0,01	0,004
5	1	0,18	0,185	1	3	0,01	0,013
5	2	0,43	0,435	1	4	0,04	0,039
5	3	0,71	0,718	1	5	0,09	0,091
5	4	0,93	0,927	1	6	0,18	0,181
5	5	1,00	1,000	1	7	0,31	0,313
5	6	0,94	0,936	1	8	0,48	0,482
5	7	0,77	0,775	1	9	0,66	0,664
5	8	0,58	0,576	1	10	0,83	0,832
5	9	0,39	0,387	1	11	0,95	0,953
5	10	0,24	0,235	1	12	1,00	1,000
10	0	0,00	0,000	1	13	0,96	0,961
10	1	0,00	0,003	1	14	0,84	0,828
10	2	0,01	0,014	1	15	0,66	0,663
10	3	0,05	0,046	1	16	0,47	0,469
10	4	0,12	0,115	1	17	0,27	0,266
10	5	0,24	0,236	1	18	0,12	0,121
10	6	0,41	0,406	5	0	0,02	0,024
10	7	0,61	0,609	5	1	0,09	0,093
10	8	0,81	0,805	5	2	0,29	0,282
10	9	0,95	0,947	5	3	0,61	0,612
10	10	1,00	1,000	5	4	0,89	0,886
12	0	0,00	0,000	5	5	1,00	1,000
12	1	0,00	0,000	5	6	0,90	0,905
12	2	0,00	0,002	5	7	0,66	0,662
12	3	0,01	0,008	5	8	0,40	0,405
12	4	0,03	0,026	5	9	0,21	0,205
12	5	0,07	0,064	5	10	0,09	0,086
12	6	0,14	0,138	9	0	0,21	0,198
12	7	0,26	0,256	9	1	0,66	0,664
12	8	0,42	0,421	9	2	0,99	0,993
12	9	0,62	0,614	9	3	0,86	0,852
12	10	0,81	0,804	9	4	0,49	0,493
				9	5	0,21	0,206
				9	6	0,06	0,064
				9	7	0,01	0,015

Tabela 2: Resultados dos testes de homogeneidade e do equilíbrio de Hardy-Weinberg.

Teste de independência em uma tabela de contingência 2x2

Suponha que testes laboratoriais sejam utilizados para auxiliar no diagnóstico de uma doença. Deve ser interessante verificar se os resultados dos testes estão realmente relacionados à condição de saúde do paciente. Um paciente amostrado de uma clínica é classificado como um dos quatro estados do conjunto

$$\Theta = \{(h, t) \mid h, t = 0 \text{ ou } 1\}$$

de modo que h seja variável indicadora de ocorrência da doença e t indicadora de resultado positivo no teste. Para uma amostra de tamanho n registramos no vetor $(x_{00}, x_{01}, x_{10}, x_{11})$ cujas componentes são a frequência de cada possibilidade de (t, h) na amostra. O espaço paramétrico é o *simplex*

$$\Theta = \left\{ (\theta_{00}, \theta_{01}, \theta_{10}, \theta_{11}) \mid \theta_{ij} > 0 \wedge \sum_{i,j} \theta_{ij} = 1 \right\}$$

E a hipótese nula de que h e t são independentes é definida por

$$\Theta_0 = \{\theta \in \Theta \mid \theta_{00} = \theta_{0\cdot} \cdot \theta_{\cdot 0}; \theta_{0\cdot} = \theta_{00} + \theta_{01}; \theta_{\cdot 0} = \theta_{00} + \theta_{10}\}$$

A equação de solução H_0 utilizada foi

$$H_0: \theta_{00} - (\theta_{00} + \theta_{01}) * (\theta_{00} + \theta_{10}) = 0$$

Os e-valores para os cinco diferentes cenários estão apresentados na Tabela 3.

X ₀₀	X ₀₁	X ₁₀	X ₁₁	Evidência da hipótese	
				Pereira e Stern	Algoritmo automático
12	6	95	35	0,96	0,957
48	25	9	10	0,54	0,538
96	50	18	20	0,24	0,231
18	5	39	30	0,29	0,288
36	10	78	60	0,06	0,059

Tabela 3: Evidência de H_0 para cada uma das cinco tabelas de contingência.

Comparação de duas distribuições Gama

O modelo Gama pode ser útil ao comparar duas distribuições de sobrevivência. Por exemplo, procedimentos médicos e eficiência farmacológica, confiabilidade de peças, ativos no mercado financeiro. A estatística suficiente para a distribuição Gama é o tamanho da amostra (n), a soma (s) e o produto (p) das observações.

A Tabela 5 contém o tempo até a falha de comparadores de moedas (um componente de caça níqueis) de duas marcas diferentes. Um empresário se ofereceu para substituir a marca 1 pela marca mais barata 2. O empresário testou 10 comparadores de moedas de cada marca e computou as médias amostrais e desvios-padrão. A distribuição Gama se encaixa muito bem nesse tipo de tempo de falha e foi usada para modelar o processo. Sua função de verossimilhança é

$$f(n_1, \alpha_1, \beta_1, n_2, \alpha_2, \beta_2 | \text{dados}) \propto \frac{\beta_1^{\alpha_1 n_1}}{\Gamma(\alpha_1)^{n_1}} \frac{\beta_2^{\alpha_2 n_2}}{\Gamma(\alpha_2)^{n_2}} p_1^{\alpha_1 - 1} e^{-\beta_1 s_1} p_2^{\alpha_2 - 1} e^{-\beta_2 s_2}$$

Os tempos de falha da segunda marca possuem média e desvio de 31,42 e 8,65; e não os 29,25 e 3,65 exibidos na tabela 4. Além disso, os cinco primeiros elementos da segunda amostra são os últimos da primeira. Apesar de as informações sobre média e desvio-padrão parecerem mais confiáveis, a estatística suficiente para esse modelo envolve o produto dos tempos até a falha em cada amostra. Portanto, optou-se por replicar o exemplo com os dados disponibilizados.

Sendo a média e o desvio-padrão denotados por $\mu = \alpha/\beta$ e $\sigma = \mu/\beta$, as hipóteses testadas foram $H'_0: \mu_1 = \mu_2$ e $H_0: \mu_1 = \mu_2 \wedge \sigma_1 = \sigma_2$. Os sistemas de equações referentes a essas duas hipóteses:

$$H'_0: \begin{cases} \alpha_1 - \alpha_2 = 0 \\ \beta_1 - \beta_2 = 0 \end{cases} \quad \text{e} \quad H_0: \begin{cases} \alpha_1 - \alpha_2 = 0 \\ \beta_1 - \beta_2 = 0 \end{cases}$$

As medidas de evidência resultantes estão apresentadas na Tabela 5. O aumento na média e no desvio-padrão (segunda amostra) do artigo original para o aqui exposto está de acordo com os resultados apresentados. A evidência de que as médias são iguais foi maior do que no artigo original, provavelmente por conta da manutenção da proximidade das médias.

O crescimento da evidência de que as médias são iguais, H' , com relação ao artigo original, está associado ao aumento da média observada da marca 2, que passou a ficar mais próxima da primeira marca. O mesmo argumento vale para a hipótese H , visto que a diferença entre os desvios-padrão também se reduziu.

Hipótese	Pereira e Stern	Algoritmo automático
H'	0,89	0,977
H	0,01	0,545

Tabela 4: Evidências de H e H'

Brand 1 sample				
39.27	31.72	12.33	27.67	56.66
28.32	53.72	29.71	23.76	33.55
mean ₁ =33.67		std ₁ =13.33		
Brand 2 sample				
28.32	53.72	29.71	23.76	33.55
24.07	33.79	33.10	26.93	27.23
mean ₂ =29.25		std ₂ =3.62		
Evidence				
$Ev(H') = 0.89$		$Ev(H) = 0.01$		

Tabela 5: Tempos de falha dos 20 comparadores de moeda. (PEREIRA e STERN, 1999)

7. Conclusão

Este trabalho comprovou a aplicabilidade do algoritmo proposto de automatizar a construção da região HPD e o cálculo do e-valor (valor de evidência) do FBST. Isso porque os resultados, comparados os deste trabalho com os do artigo original, foram iguais.

No futuro próximo, pretende-se criar uma biblioteca para o software R com a implementação desse algoritmo.

8. Referências Bibliográficas

Gamerman, D. (1997). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman & Hall . New York , ISBN: 0412818205.

Paulino, CD; Turkman, MAA; Murteira, B. (2003). *Estatística Bayesiana*. Fundação Calouste Gulbenkian, Lisboa.

Pereira, CAB.; Stern, JM. (1999). Evidence and credibility: full bayesian significance test of precise hypothesis. *Entropy*, 1:99-110.

R Development Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.