MÉTODOS DE ELICITAÇÃO DA DISTRIBUIÇÃO A PRIORI BETA

Fernado Antônio MOALA[1]
Débora Luzia PENHA[1]

▪ RESUMO: Uma característica distintiva da Inferência Bayesiana é que nos permite introduzir a informação de um especialista na análise estatística. Portanto, o desenvolvimento de métodos para construção de distribuições a priori que represente esta informação de forma adequada é muito importante para a aplicação de métodos Bayesianos. Contudo, a dificuldade para construir um procedimento para quantificar a informação a priori tem contribuido para o aparecimento de vários procedimentos para a obtenção de distribuições a priori. Então, um estudo para verificar quão informativo estes procedimentos são é de grande interesse prático. Neste artigo, os métodos propostos por O'Hagan (1998), Fox (1966), Gross (1971) e Bissecção para elicitar a distribuição a priori Beta são discutidos. Um exemplo real é apresentado para mostrar a aplicação desses métodos em uma situação prática.

▪ PALAVRAS-CHAVE: Elicitação; distribuição Beta; distribuição a priori; análise Bayesiana; especialista.

1 Introdução

A elicitação é o processo de transformação do conhecimento de um especialista sobre alguma quantidade desconhecida na forma de uma distribuição de probabilidade. No contexto da análise estatística Bayesiana, essa distribuição de probabilidade é geralmente utilizada como distribuição a priori para um ou mais parâmetros desconhecidos de um modelo estatístico.

A elicitação é uma parte importante do paradigma Bayesiano, porém, ainda pouco utilizada. Apesar de trazer grandes contribuições à análise estatística, muitos acreditam que a elicitação é inviável. Um dos fatores é a diculdade em representar adequadamente a opinião do especialista na forma probabilística, outro é o vício

[1]Universidade Estadual Paulista - UNESP, Departamento de Estatstica, CEP: 19060-900, Presidente Prudente, SP, Brasil. E-mail: *femoala@fct.unesp.br*

algumas vezes encontrado nas respostas dos especialistas quando questionados sobre situações que envolvem incertezas.

O desafio no processo de elicitação consiste na elaboração de perguntas adequadas que permitam capturar o conhecimento do especialista e amenizem o possível vício de suas respostas. Além disso, é necessário uma fase de treinamento para que o especialista se familiarize com o formato do processo de elicitação.

Existem diversas aplicações das técnicas de elicitação no estudo de situações raras, ou seja, situações em que não existe uma quantidade relevante de dados disponível. Nesses casos, a opinião do especialista é fundamental no processo de decisão. Gustafson et al. (2003) elicitou a opinião de um especialista sobre mudança organizacional, nessa situação não existiam dados disponíveis. Dominitz (1998) elicitou a opinião de um especialista sobre futuras remunerações, novamente não existiam dados. Ang e Buttery (1997) elicitaram avaliações sobre a segurança de usinas nucleares. Laws e O'Hagan (2002) elicitaram a opinião de um especialista sobre erros no processo de auditoria financeira. Alguns outros exemplos podem ser encontrados em Grisley e Kellogg (1982) sobre aplicações na agricultura e Coolen et al. (1992) em aplicações na engenharia. Martz e Waller (1982), Fox (1966), Gross (1971) e Van Noortwijk et al. (1992) apresentam procedimentos de elicitação no contexto de confiabilidade.

É assumido que o especialista apenas seja capaz de definir algumas medidas de sua densidade, tais como mediana, moda, média e percentis. Para traduzir essas informações na forma de uma distribuição de probabilidade podemos utilizar tanto a metodologia paramétrica quanto a não-paramétrica. Na elicitação paramétrica, o estatístico ajusta as informações provenientes do especialista a uma distribuição pertencente a alguma família de distribuições paramétricas. A elicitação não-paramétrica é um procedimento mais flexível, contudo, podem ser computacionalmente impraticável para grandes conjuntos de dados (O'Hagan, 1991 e Rasmussen & Ghahramani, 2002).

Nosso interesse nesse artigo é discutir alguns métodos de elicitação propostos na literatura com ênfase na elicitação da distribuição Beta. A família de distribuições Beta é frequentemente utilizada como distribuição a priori para a confiabilidade de um componente ou proporções. Nos concentraremos nos mtodos de elicitao propostos por O'Hagan (1998), Fox (1966), Gross (1971) e Bissecção. Para ilustrar esses métodos de elicitação e compará-los nós consideramos dados simulados como fossem elicitados de um especialista.

Esse artigo está organizado da seguinte forma. Na Seção 2, as etapas do processo de elicitação são descritas. Seção 3 descreve as dificuldades com o uso de distribuições a priori mais simples para representação da informação do especialista. Algumas técnicas paramétricas para elicitação da distribuio Beta são abordadas na Seção 4. A Seção 5 fornece uma simulação das abordagens consideradas no artigo e comparações. Na Seção 6 descrevemos uma aplicação real dos metodos estudados e finalmente a Seção 7 traz as conclusões.

2 O Processo de elicitação

Uma metodologia confiável a fim de elicitar a informação proporcionada pelo especialista é um pré-requisito para especificarmos a distribuição a priori dos parâmetros de um dado modelo estatístico. A literatura mostra que é importante, ao aplicar um procedimento de elicitação da opinião do especialista, fazê-lo de uma forma estruturada, clara e transparente. Nesta seção discutimos como estruturar a elicitação do especialista em etapas, de forma que possamos obter uma distribuição a priori defensável. Segundo Garthwaite, Kadane e O'Hagan (2005) o processo de elicitação pode ser realizado em quatro etapas.

Etapa 1. Seleção e treinamento do especialista.

O treinamento do especialista é realizado para investigar a precisão de suas especificações probabilísticas e familiarizá-lo com o procedimento estatístico, para aqueles que não os conhecem.

Etapa 2. Elicitação das especificações probabilísticas do especialista.

As especificações escolhidas podem ser momentos da distribuição ou qualquer outra quantidade que o estatístico acredita que o especialista possa ser capaz de fornecer. Nesse trabalho, alguns percentis serão elicitados nesse estágio do processo.

Etapa 3. Ajuste de uma distribuição de probabilidade às especificações probabilísticas fornecidas pelo especialista.

A distribuição ajustada é conhecida como densidade estimada do especialista, pois tal distribuição representa a opinião do especialista.

Etapa 4. Feedback.

Os resultados do ajuste são apresentados ao especialista que avalia se sua opinião está sendo bem representada. Se o especialista ficar insatisfeito com os resultados, então retorna-se á etapa 2 e elicita-se mais informações.

A etapa 2 tem sido foco de muitos trabalhos na área da psicometria. Apesar de existir uma vasta literatura sobre os aspectos psicométricos no processo de elicitação, não faz parte do objetivo desse trabalho tal estudo. Informações adicionais sobre a estruturação e aplicação de um processo de elicitação podem ser obtidas em Meyer e Booker (1991).

3 Técnicas paramétricas de elicitação

Uma das suposições presentes na maior parte das técnicas de elicitação é a de que a densidade do especialista pertence a alguma família paramétrica. Assim, o foco do estudo torna-se a elicitação dos hiperparâmetros da densidade de probabilidade selecionada.

Uma vez obtidas as especificações probabilísticas do especialista, o processo de elicitação se completa com o ajuste de uma distribuição a esses dados elicitados.

A distribuição Uniforme é a distribuição mais simples que pode ser ajustada às crenças do especialista.

Nesse caso, o especialista só precisa especificar um intervalo [a; b] no qual ele acredita que o parâmetro esteja. Na maioria das vezes, essa distribuição é bastante criticada por ser muito simples. Ela atribui probabilidade zero para os valores que se encontram fora do intervalo especificado e mesma probabilidade para todos aqueles que pertencem ao intervalo.

Assim, valores próximos a extremidade do intervalo possuem a mesma probabilidade daqueles que se situam no centro do mesmo.

A distribuição triangular evita o problema de atribuir mesma probabilidade para valores centrais e extremos. Para essa distribuição, o especialista também precisa especificar a moda, denotada por 'm'. Então a distribuição assumida possui densidade

$$f(x) = \begin{cases} 2\frac{x-a}{(b-a)(m-a)} & \text{se} \quad a \leq x \leq m \\ 2\frac{b-x}{(b-a)(b-m)} & \text{se} \quad m \leq x \leq b \end{cases} \tag{1}$$

O ajuste dessa distribuição à crença do especialista pode ser melhor do que o encontrado com a distribuição uniforme, porém a probabilidade de um valor que não pertence ao intervalo ainda é zero.

Geralmente, métodos de elicitação mais complicados impõem uma estrutura sobre a opinião de um especialista, assumindo que seu conhecimento pode ser representado adequadamente por uma distribuição de probabilidade mais complexa do que a uniforme e a triangular.

Essas distribuições de probabilidade são controladas por um ou mais parâmetros (conhecidos como hiperparâmetros). Assim, o processo de elicitação consiste em escolher adequadamente valores para esses hiperparâmetros que consigam capturar as principais característica da opinião do especialista.

4 Ajustando a distribuição Beta

A distribuição Beta é bastante utilizada para representar o conhecimento do especialista para determinadas quantidades de interesse, por exemplo a função de confiabilidade num instante "t", denotada por $R(t)$.

Se a variável aleatória X possui distribuição Beta com parâmetros de forma l e k, então sua função densidade de probabilidade é dada por

$$f(x) = \frac{1}{B(l,k)} x^{l-1}(1-x)^{k-1} \qquad 0 < x < 1, \tag{2}$$

onde $B(l,\, k) = \frac{\Gamma(l)\Gamma(k)}{\Gamma(l+k)}$ e $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ é a função gama.

A esperança e a variância são dadas respectivamente por:

$$E(X) = \frac{l}{l+k} \quad \text{e} \quad Var(X) = \frac{lk}{(l+k)^2(l+k+1)}. \tag{3}$$

A distribuição Beta possui dois parâmetros que não devem ser elicitados diretamente, visto que é impraticável esperar que o especialista entenda a função dos parâmetros de uma distribuição de probabilidade. Porém, a média e a variância da distribuição Beta possuem relação direta com seus parâmetros. Assim, elicitando os valores da média e variância podemos determinar os valores de l e k. Contudo, a elicitação dos momentos de uma distribuição é muito criticada (veja Garthwaite, Kadane e O'Hagan, 2005).

Alguns pesquisadores como Beach e Swenson (1966), Peterson e Miller (1964) e Spencer (1961) investigaram a habilidade das pessoas em fornecer medidas de tendência central.

Uma amostra de números foi exibida para alguns indivíduos para que estimassem a média, a moda e a mediana. Os pesquisadores observaram que quando a distribuição da amostra era aproximadamente simétrica, as estimativas dessas três medidas eram precisas. Porém, o mesmo não acontecia quando a distribuição da amostra era assimétrica. Por exemplo, os pesquisadores Peterson e Miller (1964) observaram que quando a distribuição era assimétrica à direita, as avaliações da mediana e da moda foram razoavelmente precisas, porém as avaliações da média foram tendenciosas.

A elicitação da variância é ainda mais criticada. Em estudos onde a estimativa da variância foi requisitada ao especialista, os pesquisadores perceberam que as pessoas têm muita dificuldade em entender o signicado da variância e atribuir valores numéricos a ela. Assim, a variância não deve ser elicitada diretamente.

Um método proposto por O'Hagan (1998) e utilizado para ajustar qualquer distribuição à opinião do especialista, evita a elicitação da variância. Em um primeiro momento, esse método requer a elicitação da moda ($\hat{m}$), do limite inferior ($\hat{i}$) e do superior ($\hat{s}$) da distribuição do especialista. No próximo estágio as seguintes probabilidades são requeridas do especialista

$$\begin{aligned}
p_1 &= P(\hat{i} < X < \hat{m}), \\
p_2 &= P\left(\hat{i} < X < \frac{\hat{i}+\hat{m}}{2}\right), \\
p_3 &= P\left(\frac{\hat{m}+\hat{s}}{2} < X < \hat{s}\right), \\
p_4 &= P\left(\hat{i} < X < \frac{\hat{i}+3\hat{m}}{4}\right), \\
p_5 &= P\left(\frac{\hat{s}+3\hat{m}}{4} < X < \hat{s}\right).
\end{aligned} \tag{4}$$

As probabilidades foram requisitadas nessa ordem para minimizar o vício. Além disso, a escolha desses intervalos evita que o especialista avalie pequenas probabilidades. Para fornecer uma melhor elicitação do centro da distribuição, os intervalos são menores em torno da moda.

Posteriormente as respostas são convertidas em seis probabilidades, descritas como:

$$\begin{aligned}
q_1 = P\left(\hat{i} < X < \frac{\hat{i}+\hat{m}}{2}\right) &= p_2, \\
q_2 = P\left(\frac{\hat{i}+\hat{m}}{2} < X < \frac{\hat{i}+3\hat{m}}{4}\right) &= p_4 - p_2, \\
q_3 = P\left(\frac{\hat{i}+3\hat{m}}{4} < X < \hat{m}\right) &= p_1 - p_4, \\
q_4 = P\left(\hat{m} < X < \frac{3\hat{m}+\hat{s}}{4}\right) &= 1 - p_1 - p_5, \\
q_5 = P\left(\frac{3\hat{m}+\hat{s}}{4} < X < \frac{\hat{m}+\hat{s}}{2}\right) &= p_5 - p_3, \\
q_6 = P\left(\frac{\hat{m}+\hat{s}}{2} < X < \hat{s}\right) &= p_3.
\end{aligned} \tag{5}$$

Para o caso da distribuição Beta, temos $\hat{i} = 0$ e $\hat{s} = 1$. Para obter a distribuição Beta$(\hat{l}, \hat{k})$ que melhor representa a opinião do especialista, determinamos o par de valores de $(\hat{l}, \hat{k})$ que minimize a soma de quadrados das diferenças entre as probabilidades elicitadas e as probabilidades da distribuição Beta(l, k).

Entre os métodos descritos, podemos destacar o de Fox (1966) e Gross (1971).

No método de Fox (1966) o especialista é solicitado a fornecer uma estimativa para a moda de sua distribuição, denotada por $\hat{m}$ e a probabilidade p de X pertencer ao intervalo $[\hat{m} - Z\hat{m}, \hat{m} + Z\hat{m}]$, em que $0 < Z < 1$ é dado ao especialista. As estimativas dos parâmetros, l e k, são encontradas através das seguintes equações

$$\frac{\hat{l}-1}{\hat{l}+\hat{k}-2} = \hat{m}, \tag{6}$$

$$\int_{\hat{m}-Z\hat{m}}^{\hat{m}+Z\hat{m}} \frac{1}{B(\hat{l},\hat{k})} x^{\hat{l}-1}(1-x)^{\hat{k}-1} dx = p. \tag{7}$$

O método de Gross (1971) sugere que o especialista forneça a média de sua distribuição $(\overline{x})$ e a probabilidade p de X pertencer ao intervalo $[0, Z\overline{x}]$, em que $0 < Z < 1$ é dado ao especialista pelo estatístico. As equações obtidas são

$$\frac{\hat{l}}{\hat{l}+\hat{k}} = \overline{x}, \tag{8}$$

$$\int_{0}^{Z\overline{x}} \frac{1}{B(\hat{l},\hat{k})} x^{\hat{l}-1}(1-x)^{\hat{k}-1} dx = p. \tag{9}$$

Em comparao ao método proposto por O'Hagan o qual requer seis probabilidades elicitadas tanto o mtodo de Fox quanto Gross exigem apenas uma.

Outro método proposto na literatura é o método da Bissecção. Este método requer que o especialista forneça os quartis, particioando o intervalo [0, 1] em quatro intervalos de probabilidades iguais. Segundo O'Hagan et al.(2005) esse método consiste na seguinte sequência de perguntas:

Seja X a quantidade desconhecida de interesse. Determine um valor (a mediana do especialista) tal que X seja igualmente provável ser maior ou menor do que esse ponto.

Supondo que o valor de X esteja abaixo da mediana especificada anteriormente, determine um novo valor (quartil inferior) tal que seja igualmente provável que X seja menor ou maior do que esse valor.

Supondo que o valor de X esteja acima da mediana especificada na pergunta 1, determine um novo valor (quartil superior) tal que seja igualmente provável X ser menor ou maior do que esse valor.

5 Simulação

Nesta seção analisaremos o comportamento da distribuição a priori beta elicitada através dos métodos propostos neste artigo.

Para efeito de comparação dessas distribuições a priori com a verdadeira distribuição a priori Beta utilizamos diferentes valores dos parâmetros que fornecem diferentes formas da densidade da Beta.

As Tabelas 1 a 3 fornecem os estimadores dos parâmetros obtidos da informação dada pelo especialista em forma de probabilidades, média e moda. Assumindo que a informação do especialiasta está sujeita a imprecisão, consideramos então um erro ε_1 ocorrido nas probabilidades elicitadas e um erro ε_2 na média e moda. Sem perda de generalidade consideramos esses erros iguais em cada tabela. Assim, as probabilidades, média e moda consideradas elicitadas nesta simulação são dadas por:

$$P^*\{\theta \in A\} = P\{\theta \in A\} + \varepsilon_1, \hat{m} = m + \varepsilon_2, \overline{x} = E + \varepsilon_2, \tag{10}$$

onde $P\{\theta \in A\}$, m e E são probabilidades, moda e média da distribuição Beta, respectivamente. A é o intervalo contendo a varivel cuja probabilidade ser elicitada. Os erros tomam valores $\varepsilon_i = 0.01, 0.05$ e 0.1 para avaliar sua influência na estimação. Como os valores a serem elicitados estão entre 0 e 1 assume-se que o especialista seja capaz de fornecer valores com uma margem de erro pequena. Desta forma, o especialista estará fornecendo os valores de $P^*\{\theta \in A\}$, $\hat{m}$ e $\overline{x}$ onde A constitui diferentes intervalos sugeridos pelo estatístico para que ele forneça as respectivas probabilidades. Nesta Seção, estes valores são gerados da distribuição Beta para cada par de parâmetros considerados nas Tabelas a seguir.

Os gráficos das distribuições a priori são mostrados após cada correspondente tabela.

Tabela 1 - Estimação dos parâmetros da distribuição a priori Beta elicitada com $\varepsilon_1 = \varepsilon_2 = 0.01$

(l, k)	O'Hagan	Fox	Gross	Bissecção
(1 , 3)	(0.96, 2.75)	(1.03, 4.45)	(0.86, 2.45)	(1.07, 3.07)
(2, 2)	(1.84, 1.85)	(2.09, 2.05)	(1.75, 1.68)	(2.03, 1.96)
(2, 5)	(1.83, 4.55)	(2.16, 5.36)	(1.72, 4.11)	(2.11, 5.05)
(5, 1)	(4.45, 0.89)	(5.41, 0.95)	(3.65, 0.68)	(4.69, 0.89)

Tabela 2 - Estimação dos parâmetros da distribuição a priori Beta elicitada com $\varepsilon_1 = \varepsilon_2 = 0.05$

(l, k)	O'Hagan	Fox	Gross	Bissecção
(1 , 3)	(0.72, 1.89)	(1.25, 5.79)	(0.45, 1.05)	(1.37, 3.31)
(2, 2)	(1.38, 1.42)	(2.47, 2.20)	(0.97, 0.79)	(2.16, 1.80)
(2, 5)	(1.27, 3.07)	(2.84, 6.51)	(0.93, 1.84)	(2.57, 5.19)
(5, 1)	(3.27, 0.69)	(20.00, 0.1)	(1.19, 0.15)	(3.35, 0.51)

Tabela 3 - Estimação dos parâmetros da distribuição a priori Beta elicitada com $\varepsilon_1 = \varepsilon_2 = 0.1$

(l, k)	O'Hagan	Fox	Gross	Bissecção
(1 , 3)	(0.50, 1.19)	(1.68, 7.1)	(0.15, 0.27)	(1.76, 3.46)
(2, 2)	(1.00, 1.07)	(2.90, 2.26)	(0.32, 0.21)	(2.26, 1.58)
(2, 5)	(0.79, 1.80)	(3.64, 7.16)	(0.36, 0.57)	(3.15, 5.18)
(5, 1)	(2.68, 0.64)	(12.00, 3.6e-05)	(0.01, 0.001)	(1.86, 0.23)

Utilizamos o parâmetro Z requerido nos procedimentos propostos por Fox e Gross com o valor $Z = 0.5$.

Obviamente quando a margem de erro é pequena a precisão é muito maior na estimação da distribuição a priori independentemente do método utilizado. Contudo, quando esse erro aumenta, alguns métodos tornam-se completamente inapropriados produzindo inclusive distribuições a priori consideradas não-informativas, como o método de Gross mostrado na Figura 3. As distribuições a priori Beta(1, 1) e B(0.5, 0.5) são consideradas distribuiçõs a priori não-informativas.

Pelas Figuras 1 a 3 observamos também que a distribuição a priori proposta por O'Hagan é a menos influencida pela precisão da informação do especialista.

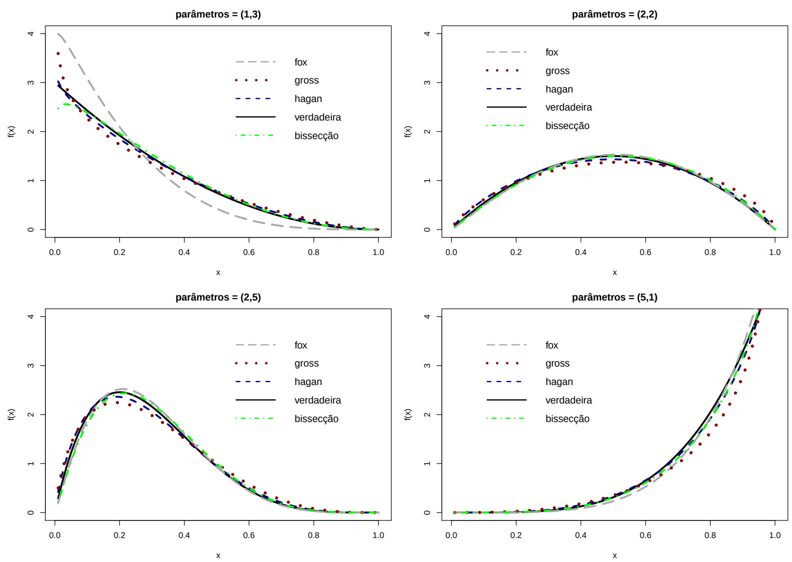

Figura 1 - Comparação gráfica das densidades elicitadas e verdadeira para $\varepsilon_1 = \varepsilon_2 = 0.01$.

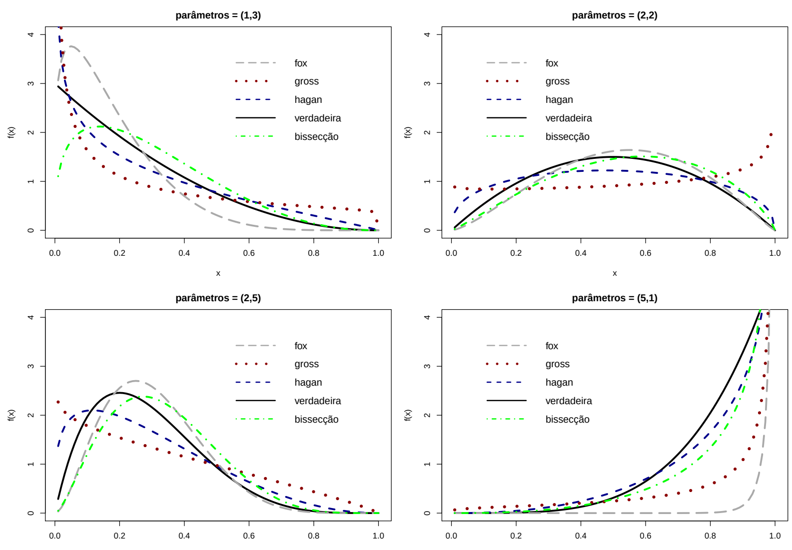

Figura 2 - Comparação gráfica das densidades elicitadas e verdadeira para $\varepsilon_1 = \varepsilon_2 = 0.05$.

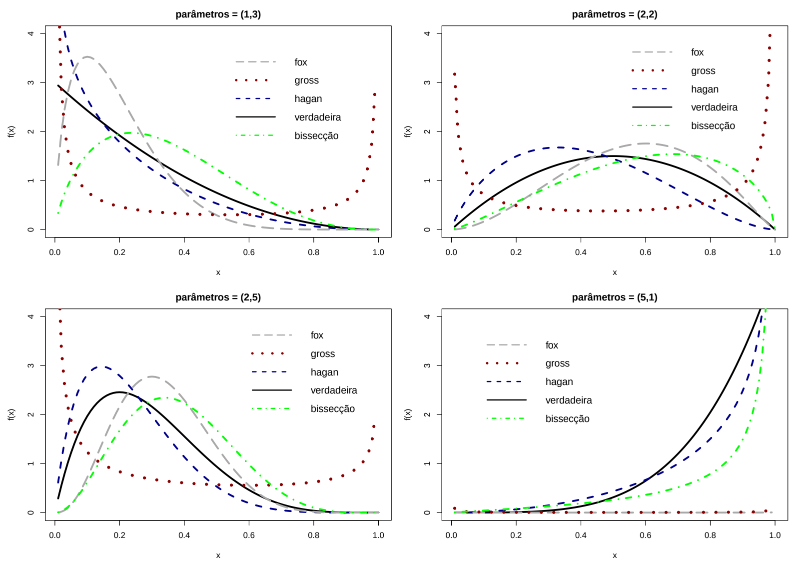

Figura 3 - Comparação gráfica das densidades elicitadas e verdadeira para $\varepsilon_1 = \varepsilon_2 = 0.1$.

Os métodos propostos por Fox e Gross dependem do parâmetro k, cujo valor deve estar no intervalo [0, 1], fornecido pelo estatístico. A Tabela 4 mostra a influência do valor de Z na estimação da distribuição a priori do especialista.

Podemos observar que as estimativas dos parâmetros da distribuição a priori Beta são melhores quando $Z = 0.5$.

Tabela 4 - Análise do parâmetro Z requerido nos métodos de Fox e Gross

		FOX			GROSS	
(l, k)	Z=0.1	Z=0.5	Z=0.8	Z=0.1	Z=0.5	Z=0.8
(1 , 3)	(1.95, 19.10)	(1.03, 4.45)	(1.21, 5.07)	(0.65, 1.52)	(0.45, 1.05)	(0.16, 0.38)
(2, 2)	(3.42, 2.98)	(2.09, 2.05)	(18.02, 14.93)	(0.87, 0.71)	(0.97, 0.79)	(0.01, 0.01)
(2, 5)	(4.07, 10.21)	(2.16, 5.36)	(3.14, 7.44)	(1.00, 1.98)	(0.93, 1.84)	(0.26, 0.52)
(5, 1)	(4.45, 0.83)	(5.41, 0.95)	(10.18, 0.56)	(8.57, 1.13)	(1.19, 0.15)	(1.20, 0.22)

6 Exemplo real

Discutimos agora uma aplicação prática apresentada em Moala (2009) para a estimação da distribuição a priori do risco de fraturas em pacientes que sofrem de osteoporose.

Para um paciente recebendo um particular tratamento há risco incerto de fratura do osso e então nosso interesse será a construção de uma distribuição a priori para este risco sob um determinado tratamento. A informação sobre o risco para qualquer tratamento virá das informações do clínico (especialista em osteoporose).

É pedido ao clínico fornecer o valor mais provável do risco (moda $\widehat{m}$), o risco esperado/médio $\overline{x}$ e os quartis $q_{0.25}$, $q_{0.50}$ e $q_{0.75}$. No caso dos quartis, por exemplo do 1ºquartil, o estatístico pede ao clínico o valor $q_{0.25}$ para qual ele sente haver uma chance de 25% que o risco seja menor que $q_{0.25}$.

Então, baseado nessas informações podemos estimar os parâmetros da distribuição Beta através dos quatro métodos de elicitação propostos.

A partir dos quartis, média e moda fornecidos pelo clnico e das probabilidades ajustadas em Moala (2009) podemos complementar a informação necessária para a utilização dos métodos propostos aqui.

Os quartis elicitados são dados por $q_{0.25} = 0.55$, $q_{0.5} = 0.62$ e $q_{0.75} = 0.75$, moda $m = 0.6$ e média $\overline{x} = 0.7$.

Daí, as probabilidades requiridas na utilização do método de O'Hagan são dadas por

$$p_1 = P\{0 \leq X \leq m\} = P\{0 \leq X \leq 0.6\} = 0.42$$

$$p_2 = P\{0 \leq X \leq \tfrac{L+m}{2}\} = P\{0 \leq X \leq 0.3\} = 0.03$$

$$p_3 = P\{\tfrac{m+U}{2} \leq X \leq U\} = P\{0.8 \leq X \leq 1\} = 1 - P\{X \leq 0.8\} = 0.20$$

$$p_4 = P\{0 \leq X \leq \tfrac{L+3m}{4}\} = P\{0 \leq X \leq 0.45\} = 0.093$$

$$p_5 = P\{\tfrac{U+3m}{4} \leq X \leq 1\} = P\{0.7 \leq X \leq 1\} = 1 - P\{X \leq 0.7\} = 0.30$$

Considerando $Z = 0.5$, então a probabilidade requerida pelo Método de Fox é dada por

$$P = P\{m - mZ \leq X \leq m + mZ\} = P\{0.3 \leq X \leq 0.9\} = P\{X \leq 0.9\} - P\{X \leq 0.3\} = 0.887 - 0.03 = 0.857.$$

e para o Método de Gross a probabilidade é dada por

$$P = P\{0 \leq X \leq \overline{x}Z\} = P\{0 \leq X \leq 0.35\} = 0.044 \quad \text{, para } Z = 0.5.$$

A Tabela 5 apresenta os parâmetros estimados da distribuição a priori Beta pelos quatro métodos de elicitação considerados neste artigo e a Figura 4 apresenta as correspondentes densidades estimadas.

Tabela 5 - Elicitação da distribuição Beta

Métodos	l	k
O'Hagan	5.406	3.244
Fox	2.912	2.274
Gross	3.800	1.628
Bissecção	6.917	3.992

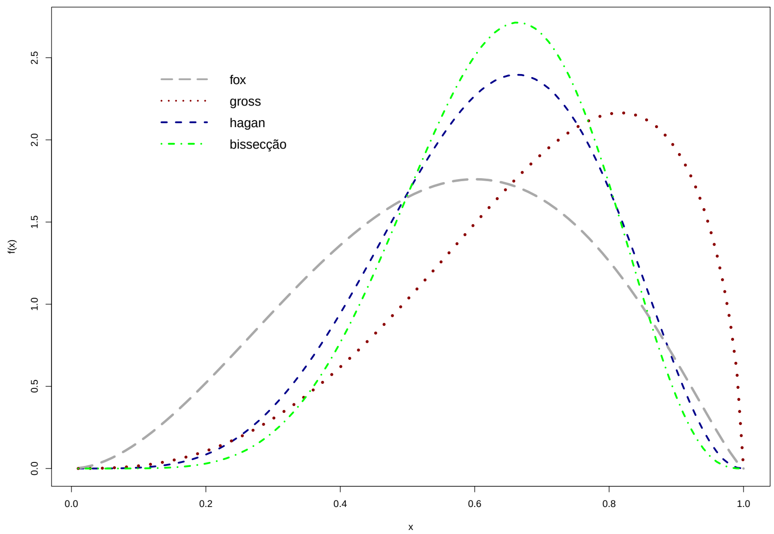

Figura 4 - Distribuições a priori elicitadas para o risco de osteoporose.

Concluses

A Elicitação de distribuições a priori é um importante componente da Inferência Bayesiana e ainda pouco pesquisada. A opinião do especialista tem um importante papel na modelagem de problemas para os quais não há dados suficientes. Além disso, em qualquer análise estatstica haver certamente alguma forma de conhecimento disponvel sobre o assunto, além dos dados. Nós apresentamos os resultados de uma simulação comparando quatro populares métodos para a elicitação da distribuição Beta. Mostramos uma superioridade dos métodos propostos por O'Hagan e Bissecção sobre os métodos propostos por Fox e Gross. O método de O'Hagan mostra uma performance melhor que da Bissecção quando a imprecisão do especialista e aumenta. Um exemplo foi apresentados para ilustrar as metodologias analisadas. Uma proposta de estudo futuro seria estender a comparação desses métodos analisados neste artigo sob a distribuição Beta comparando-os sob a distribuição de Kumarasawmy (1980).

Agradecimentos

Os autores agradecem FAPESP (Fundao de Amparo Pesquisa do Estado de So Paulo) pelo apoio financeiro.

MOALA, F. A.; PENHA, D. L. Elicitation methods for Beta prior distribution. *Rev. Bras. Biom.,* Lavras, v.34, n.1, p.49-62, 2016.

- *ABSTRACT: A distinctive feature of Bayesian inference is that allows us to use the the expert's information in a statistical analysis. Therefore, the development of methods for construction of prior distributions that represent this information in an appropriate manner is very important for the application of Bayesian inferences. However, the difficulty to build a procedure to quantify this prior information has contributed to the development of several procedures for obtaining prior distributions. Then a study to verify how informative these procedures are is of great practical interest. In this paper, the methods proposed by O'Hagan (1998), Fox (1966), Gross (1971) and Bisection to elicit the prior distribution Beta are discussed. A real example is presented to show the application of these methods in a practical situation.*

- *KEYWORDS: Elicitation; Beta distribution; prior distribution; Bayesian analysis; expert.*

Referências

ANG, M.L.; BUTTERY, N.E. An approach to the application of subjective probabilities in level 2 PSAs. *Reliability Engineering and System Safety*, v.58, p.145-156, 1997.

BEACH, L.R.; SWENSON, R.G. Intuitive estimation of means. *Psychonomic Science*, v.5, p.161-162, 1966.

COOLEN, F.P.A.; MERTENS, P.R.; NEWBY, M.J. A Bayes-competing risk model for the use of expert judgement in reliability estimation. *Reliability Engineering and System Safety*, v.35, p.23-30, 1992

COOKE, R. *Experts in Uncertainty*. Oxford University Press, 1991.

DOMINITZ, J. Earning expectations, revisions, and realisations. *The Review of Economics and Statistics*, v.80, n.3, p.374-388, 1998.

FOX, B.L. *A Bayesian approach to reliability assessment.* Memorandum RM-5084-NASA, The Rand Corporation, Santa Monica, USA, 1966.

GARTHWAITE, P.H.; KADANE, J.B.; O'HAGAN, A. Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, v.100, p.680-701, 2005.

GRISLEY, W.; KELLOG, E.D. Farmers subjective probabilities in Northern Thailand: an elicitation analysis. *American Journal of Agricultural Economics*, v.65, p.74-82, 1982.

GROSS, A.J. The application of exponential smoothing to reliability assessment. *Technometrics*, v.65, p.877-883, 1971.

GUSTAFSON, D.H.; SAINFORT, F.; EICHLER, M.; ADAMS, L.; BISOGNANO, M.; STEUDEL, H. Developing and testing a model to predict outcomes of organizational change. *Health Services Research*, v.38, p.751-776, 2003.

HUGHES, G.; MADDEN, L. V. Some methods for eliciting expert knowledge of plant disease epidemics and their application in cluster sampling for disease incidence. *Crop Protection*, v.21, p.203-215, 2002.

KUMARASWAMY, P. A generalized probability density function for double-bounded random processes. *Journal of Hydrology* v.46, n.1-2, p.79-88, 1980.

LAWS, D.J.; O'HAGAN, A. A hierarchical Bayes model for multiloca-tional auditing. *Journal of the Royal Statistical Society*, v.51, p.431-450, 2002.

MARTZ, H.F. ; WALLER, R.A. *Bayesian Reliability Analysis*. Wiley & Sons. New York, 1982.

MEYER, M. A. and BOOKER, J. M. *Eliciting and analyzing expert judgment: a practical guide*. London: Academic Press, 1991

MOALA, F.A. Elicitao de uma distribuio a priori em uma pesquisa mdica. *Revista Brasileira de Estatstica*. v.70, p.57-74, 2009.

O'HAGAN, A. Eliciting expert beliefs in substantial practical applications. *Journal of Royal Statistical Society*, v.47, p.21-35, 1998

O'HAGAN, A. Bayes-Hermite quadrature. *Journal of Statistical Planning and Inference*, v.29, p.245-260, 1991.

PETERSON, C.R.; MILLER, A. Mode, median and mean as optimal strategies. *Journal of Experimental Psychology*, v.68, p.363-367, 1964

RASMUSSEN, E. GHAHRAMANI, Z. Infinite mixtures of Gaussian process experts. In T. G. Dietterich, S. Becker, and Z. Ghahramani, editors, *Advances in Neural Information Processing Systems*, 14, p.881-888, 2002.

SPENCER, J. Estimating averages. *Ergonomics*, v.4, p.317-328, 1961.

Van NOORTWIJK, J. M., DEKKER, R., COOKE, R. M., MAZZUCHI, T. et al. Expert judgment in maintenance optimization, *Reliability, IEEE Transactions on*, v.41, n.3, p.427-432, 1992.

WINKLER, R.L. The assessment of prior distributions in Bayesian analysis. *Journal of the American Statistical association*, v.62, p.776-800, 1967

Recebido em 22.07.2015.

Aprovado após revisão em 25.11.2015.