ME 705: Inferência Bayesiana

1 Fundamentos

1. Comente o "Principio da Incerteza" (de Heisenberg). Por que este principio é importante ?.

2. Fundamente o conceito de "Caos" desde uma visão matemática.

3. Comente o "Efeito Hawthorne". Por que este "efeito" resulta interessante?.

4. Comente o "Efeito Placebo". Por que este "efeito" resulta interessante?.

5. Explique a diferencia entre aditividade finita e aditividade infinita.

6. Explique a importância dos axiomas de Rényi na abordagem Bayesiana.

7. Que significa "Coerência" ?.

8. Que significa "Verossimilhança Relativa"?

9. Que significa "long-run frequency approach of probability"? (veja Von-Mises).

10. Explique, usando algum exemplo, a visão frequentista de probabilidade.

11. Interprete em termos probabilisticos a seguinte afirmação: "Eu acredito que existe um 40 % da chance de que no ano 2020, a medicina tenha desenvolvido uma cura para todo tipo de câncer ". Explique por que a interpretação frequentista de probabilidade não pode ser aplicada neste contexto.

12. Por que o livro "Ars Conjectandi" de James Bernoulli, apresenta particular interesse para a comunidade Bayesiana?

13. Qual é o significado do sistema axiomático de Savage?. Interprete o sentido da função de Utilidade.

14. Que significa probabilidade subjetiva?.

15. Discuta a diferencia (se existir) entre distribuição a priori e distribuição a posteriori.

16. Justifique a seguinte afirmação: "Probabilidade matemática é um caso particular da Probabilidade Subjetiva".

17. Que significa "Inferência estatística axiomática" e qual é a diferencia entre este tipo de inferência e a inferência frequentista.

Sugestão de referência: S. James Press (1989). Bayesian Statistics: Principles, Models, And Applications. John Wiley & Sons. New York.

2 Estimação

1. Determine as posterioris nos seguintes cenários,

- $x|\sigma \sim$ N(0,σ^2), $1/\sigma^2 \sim$ Gama(1,2).
- $x|\lambda \sim$ Poisson(λ), $\lambda \sim$ Gama(2,1).
- $x|p \sim$ BNeg(10,p), $p \sim$ Beta (1/2,1/2).

2. Considere duas observações desde $P_\theta(x = \theta - 1) = P_\theta(x = \theta + 1) = 0.5$, $\theta \in \mathbb{R}$. assuma que a quantidade de interesse é $\theta \in \Theta = \mathbb{R}$; assuma ainda que o espaço de decisões é $\mathcal{D} = \mathbb{R}$ e que $\forall \delta \in \mathcal{D}$, a perda associada é dada por $L(\theta, \delta) = 1 - \mathbb{I}_\theta(\delta)$, onde $\mathbb{I}_{\mathrm{u}}(\mathrm{v}) = 1$ se $u = v$ e $\mathbb{I}_{\mathrm{u}}(\mathrm{v}) = 0$ se $u \neq v$. Considere as seguintes regras de decisão: $\delta_0(x_1, x_2) = \frac{x_1+x_2}{2}$ e $\delta_1(x_1, x_2) = x_1 + 1$ e mostre que os riscos de ambas decisões coincidem, isto é $R(\theta, \delta_0) = R(\theta, \delta_1) = 0.5$.
Assuma que $\forall \delta \in \mathcal{D}$, $R(\theta, \delta) = E_\theta(L(\theta, \delta(x)))$.

3. Considere a perda quadrática ponderada $L_\omega(\theta, \delta) = \omega(\theta)(\theta - \delta)^2$, onde $\omega(\theta) = \mathbb{I}_{(-\infty, \frac{1}{2})}(\theta)$ e a distribuição a priori $\pi(\theta) = \mathbb{I}_{(0,1)}(\theta)$. Considere ainda, a verossimilhança $f(x|\theta) = \theta x^{\theta-1}\mathbb{I}_{(0,1)}(\mathrm{x})$, $\theta > 0$. Determine o estimador de Bayes δ^π.
Considere a perda quadrática $L(\theta, \delta) = (\theta - \delta)^2$, e a distribuição a priori $\pi_1(\theta)$ dada pela distribuição uniforme continua em $(0, \frac{1}{2})$. Considere ainda, a verossimilhança $f(x|\theta) = \theta x^{\theta-1}\mathbb{I}_{(0,1)}(\mathrm{x})$, $\theta > 0$. Determine o estimador de Bayes δ^{π_1}.
Compare as decisões δ^π e δ^{π_1}.

4. Considere a distribuição a priori $\pi(\theta) = (1/3)(U_{[0,1]}(\theta) + U_{[2,3]}(\theta) + U_{[4,5]}(\theta))$ e $f(x|\theta) = \theta e^{-\theta x}$.

Assumindo a perda multilinear $L_{k_1,k_2}(\theta,\delta) = \begin{cases} k_2(\theta-\delta) & \text{se } \theta > \delta \\ k_1(\delta-\theta) & \text{se } \theta \leq \delta \end{cases}$, mostre que o estimador de Bayes não é único.

5. No contexto de teste de hipótese $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \notin \Theta_0$. Determine a relação existente entre a perda "0-1," definida por

$$L(\theta,\delta) = \begin{cases} 1-\delta & \text{se } \theta \in \Theta_0 \\ \delta & \text{se } \theta \notin \Theta_0 \end{cases}$$

e os erros tipo I e II da abordagem de Neyman-Pearson.

6. Considere que a priori de θ é uma Exp(α_0), onde α_0 é uma constante conhecida e positiva. Assuma ainda que a verossimilhança de x dado θ pode ser representada por uma Exp(θ). Determine nesse caso a decisão de Bayes associada a perda quadrâtica usual.

7. Determine a distribuição a posteriori de θ dado x assumindo a priori N(0,1) para θ e a verossimilhança de x dado θ como sendo N(θ,1). Nesse casso determine o estimador de Bayes, considerando uma perda "0-1", onde $\Theta_0 = \{\theta : \theta > \theta_0\}$, e θ_0 é um valor fixo e conhecido.

8. Determine nos modelos do exercício 1. a decisão de Bayes, assumindo perda quadrática.

Sugestão de referência: Christian P. Robert (1994). The Bayesian Choice: A Decision - Theoretic Motivation. Springer-Verlag. New York.

3 Priori-Posteriori

1. Demonstre que a verossimilhança de $x|\theta$, $\sigma^2 \sim$ N($\theta, \sigma^2 I_p$) apresenta a forma exponencial. Determine nesse caso os parametros naturais da verossimilhança (reparametrização de (θ, σ^2), que permite afirmar que $f(x|\theta, \sigma^2)$ é familia exponencial natural).

2. Determine a familia conjugada natural associada com as seguintes verossimilhanças:

- $f(x|\theta)$ dada por uma distribuição N(θ, σ^2), σ^2 conhecido;

- $f(x|\theta)$ dada por uma distribuição Poisson(θ);
- $f(x|\theta)$ dada por uma distribuição Gama(ν, θ);
- $f(x|\theta)$ dada por uma distribuição Binom(n, θ);
- $f(x|\theta)$ dada por uma distribuição NegBin(m, θ);
- $f(x|\theta)$ dada por uma distribuição Multinom$(\theta_1, \cdots, \theta_k)$;
- $f(x|\theta)$ dada por uma distribuição N$(\mu, 1/\theta)$.

3. Assuma que $x|\theta \sim \mathrm{N}(\theta, 1)$ e que a distribuição a priori para θ é dada por $\pi(\theta) = \exp(-|\theta|)/2$. Determine a distribuição a posteriori para θ.

4. Assuma $f(x|\theta, \phi) = \theta^{\phi} x^{\phi-1} e^{-\theta x}/\Gamma(\phi)$ onde $\theta > 0, \phi > 0$. Demonstre que na familia conjugada natural associada a (θ, ϕ) a densidade marginal de ϕ apresenta a forma $f(\phi) \propto c^{\phi}\Gamma(1+d\phi)/(\Gamma(\phi))^d$, onde c e d representam valores conhecidos.

5. A distribuição a posteriori derivada desde uma priori impropria, pode resultar numa posteriori impropria:
Se n e p são as quantidades aleatorias de interesse e x representa o numero de sucessos em n ensaios independentes (com probabilidade de sucesso p). Assumindo ainda, que a priori $\pi(n, p) = \pi(n)\pi(p)$ e adicionalmente, que $\pi(p)$ é dada como sendo uma $U(0, 1)$. Por outro lado, considerando $\pi(n) \propto 1$, $\forall n \in \mathbb{R}$ (ou seja uma priori impropria). Demonstre que a posteriori para n é impropria.

4 Estimação e Modelo Normal

1. Verifique os seguintes resultados, considerando uma perda quadratica sob a quantidade de interesse θ.

Tabela 1: Conjugada Natural e Estimador de Bayes

Verossimilhança	Conjugada	Estimador
$\mathrm{N}(\theta, \sigma^2)$	$\mathrm{N}(\mu, \tau^2)$	$\frac{\mu\sigma^2+\tau^2 x}{\sigma^2+\tau^2}$
$\mathrm{Poisson}(\theta)$	$\mathrm{Gama}(\alpha, \beta)$	$\frac{\alpha+x}{\beta+1}$
$\mathrm{Gama}(\nu, \theta)$	$\mathrm{Gama}(\alpha, \beta)$	$\frac{\alpha+\nu}{\beta+x}$
$\mathrm{Bin}(n, \theta)$	$\mathrm{Beta}(\alpha, \beta)$	$\frac{\alpha+x}{\alpha+\beta+n}$
$\mathrm{NegBin}(n, \theta)$	$\mathrm{Beta}(\alpha, \beta)$	$\frac{\alpha+n}{\alpha+\beta+x+n}$
$\mathrm{Multi}(n; \theta_1, \cdots, \theta_k)$	$\mathrm{Dirichlet}(\alpha_1, \cdots, \alpha_k)$	$\frac{\alpha_i+x_i}{(\sum_j \alpha_j)+n}$
$\mathrm{N}(\mu, 1/\theta)$	$\mathrm{Gama}(\alpha/2, \beta/2)$	$\frac{\alpha+1}{\beta+(\mu-x)^2}$

2. Considere uma amostra $x_1, \cdots, x_n$ i.i.d. $\mathrm{N}(\theta, \sigma^2)$ onde (θ, σ^2) é a quantidade de interesse. Assuma a priori não informativa de Jeffreys π e determine,

- A distribuição de $\theta|\sigma, \overline{x}, s^2$.
- A distribuição de $\sigma^2|\overline{x}, s^2$.
- A distribuição de $\theta|\overline{x}, s^2$.

Onde, $\overline{x} = \frac{\sum_{i=1}^n x_i}{n}$, $s^2 = \sum_{i=1}^n (x_i - \overline{x})^2$.

5 Testes bayesianos

1. Considere x como sendo uma observação desde uma distribuição $\mathrm{Bin}(n, p)$, $\Theta_o = [0, 1/2]$. Assuma que o valor n é conhecido e use para p uma priori não informativa $\pi(p) \propto 1$, determine a distribuição a posteriori de H_0. Estude a variação da resposta Bayesiana em função de n para $x = 0$ e $x = n/2$.

2. Seja x $\mathrm{N}(\theta, \sigma^2)$, σ^2 conhecido, θ aleatório. Postule a hipótese $H_0 : \theta < 0$.

2.1. Considere $\theta \sim \mathrm{N}(\mu, \tau^2)$ e estude a posteriori de H_0 quando $x = 0$ e $\frac{\tau}{\sigma} \to \infty$.

2.2. Considere $\theta \sim \pi(\theta) \propto 1$, determine a posteriori de H_0 e compare com o resultado obtido no item anterior.

3. Assuma $x \sim \text{Poisson}(\lambda)$ e considere o teste $H_0 : \lambda \leq 1$ vs $H_1 : \lambda > 1$. Apresente a posteriori de H_0 quando $x = 1$ e $\lambda \sim \text{Gama}(\alpha, \beta)$. α e β hiperparametros conhecidos e fixos.

3.1. Estude a posteriori, nos casos particulares: a)$\alpha \to 0$, b) $\beta \to 0$, c) $\alpha \to 0$ e $\beta \to$, d) $\frac{\alpha}{\beta} \to 0$, e) $\frac{\beta}{\alpha} \to 0$.

3.2. Compare as conclusões anteriores com a inferência resultante da substituição da priori Gamma por uma priori não informativa $\pi(\lambda) \propto 1/\lambda$. É sempre possível assumir uma priori não informativa como esta?

4. Considere x como sendo uma observação desde uma distribuição $\text{Bin}(n, p)$, $H_0 : p = 1/2$, e $H_1 : p \neq 1/2$. Assuma a priori $\pi(p)$ como uma distribuição $\text{Beta}(\alpha, \alpha)$. α hiperparametro conhecido e fixo.
Determine o limite da probabilidade a posteriori de H_0 quando $n = 10, x = 5$ e $n = 15, x = 7$ no caso em que $\alpha \to \infty$. Compare os resultados anteriores com os obtidos assumindo a priori não informativa de Jeffreys em p.

5. Resolva os dois exercícios anteriores usando Fator de Bayes.

6. Quando $x \sim \text{N}(\theta, 1)$ e $\theta \sim \text{N}(0, \sigma^2)$, compare as respostas Bayesianas dos seguintes testes,

- $H_0^1 : \theta = 0$ vs $H_1^1 : \theta \neq 0$,
- $H_0^2 : |\theta| \leq \epsilon$ vs $H_1^2 : |\theta| > \epsilon$.

Observe e compare as respostas quando ϵ e σ variam.

7. Considere $x \sim \text{N}(\theta, 1)$ e o teste $H_0 : |\theta| \leq c$ vs $H_1 : |\theta| > c$, quando $\pi(\theta) \propto 1$.

7.1. Elabore um gráfico da probabilidade a posteriori de H_0 em função de c.

7.2. Determine o valor de c para o qual a probabilidade a posteriori de H_0 é 0.95. Calcule nesse caso o Fator de Bayes.

8. Considere $x_1, \cdots, x_n$ como independentes e identicamente distribuídos $\mathrm{N}(\theta, v)$. Sob o modelo 1: $v = 1$ e sob o modelo 2: $v = 2$. Assumamos ainda que a priori $\theta \sim \mathrm{N}(m, w)$. Mostre que quando $w \to \infty$ o Fator de Bayes do modelo 1 contra o modelo 2 tende a

$$2^{n/2} \exp(-\sum_{i=1}^{n}(x_i - \overline{x})^2/4).$$

9. Considere $x_1, \cdots, x_n$ como independentes e identicamente distribuídos $\mathrm{N}(\theta, v)$. Sob o modelo 1: $\theta = 0$. Sob o modelo 2: $\theta \sim \mathrm{N}(0, w)$. Mostre que o Fator de Bayes para o modelo 1 contra o modelo 2 tende a infinito quando $n \to \infty$, ou seja que para um tamanho amostral suficientemente grande é dada uma alta probabilidade a posteriori para o modelo 1.

10. Assuma $x \sim$ Poisson (θ). Sob o modelo 1: $\theta = 1$, sob o modelo 2: $\theta \sim$ Gamma, i.e. $f_2(\theta) = a^b \theta^{b-1} e^{-a\theta} / \Gamma(b)$.

10.1. Demonstre que o Fator de Bayes para o modelo 1 contra o modelo 2 é maximizado quando x assume o valor da parte inteira de $a - b + 2$.

10.2. No caso em que a moda a priori é igual a 1, mostre que o Fator de Bayes resulta dado pela seguinte forma:

$$e^{-1}(a+1)^{a+x+1} a^{-(a+1)} \Gamma(a+1)/\Gamma(a+x+1),$$

note ainda, que esta quantidade tende a infinito quando $a \to 0$.

11. Complemento: teste para a media de uma população Normal.
Consulte: Peter Lee (1989). Bayesian Statistics: an introduction. Co published in the Americas by Halsted Press an imprint of John Wiley & Sons Inc.

11.1. Conduzir o teste para media de uma população Normal, assumindo variância desconhecida.

11.2. Compare a proposta com o caso visto na aula (assumindo variância conhecida).

11.3. Apresente um exemplo numérico do item 11.1..

12. Complemento: teste para a diferença das medias de duas populações Normais.
Consulte: Peter Lee (1989). Bayesian Statistics: an introduction. Co published in the Americas by Halsted Press an imprint of John Wiley & Sons Inc.

12.1. Conduzir o teste para a diferença das medias, assumindo variâncias desconhecidas, mas iguais.

12.2. Apresente um exemplo numérico do item 12.1..

12.3. Conduzir o teste para a diferença das medias, assumindo variâncias desconhecidas, sem a suposição da igualdade entre elas.

12.4. Apresente um exemplo numerico do item 12.3..

12.5. Compare as respostas dos itens 12.1., 12.3. e a apresentada na aula (variâncias conhecidas).

6 Métodos Computacionais

1. Seja $g(\theta)$ uma função proporcional à distribuição normal de parâmetros μ e σ^2. Considere que é de interesse estimar

$$I = \int g(\theta)d\theta.$$

Utilize o Método de Monte Carlo, para obter esta integral. Para implementar o procedimento, utilize uma amostra gerada desde uma normal de parâmetros μ_1 e σ_1^2. Analise os resultados em função da relação entre σ^2 e σ_1^2.
Demonstre que a variância do estimador de Monte Carlo $\hat{I}$ é infinita se $\sigma_1^2 \leq \frac{\sigma^2}{2}$.

2. Seja $g(\underline{\theta})$ uma função proporcional à distribuição normal $p-$dimensional, $N_p(m, V)$. Considere que é de interesse estimar

$$I = \int g(\underline{\theta})d\underline{\theta}.$$

Utilize o Método de Monte Carlo, para obter esta integral. Para implementar o procedimento, utilize uma amostra gerada desde uma normal

$N_p(m, kV)$. Analise os resultados em função de k.
Demonstre que a variância do estimador de Monte Carlo $\hat{I}$ é dada pela forma

$$n^{-1}\left\{\left(\frac{k^2}{2k-1}\right)^{p/2} - 1\right\},$$

onde n representa o tamanho amostral utilizado. Demonstre que esta variância é infinita se $k \leq 1/2$.

3. Seja $g(\theta)$ uma função proporcional à distribuição normal de parâmetros μ e σ^2. Considere que é de interesse estimar

$$I = \int g(\theta)d\theta.$$

Utilize o Método de Monte Carlo, para obter esta integral. Para implementar o procedimento, utilize uma amostra gerada desde uma t-student. Analise o resultado $\hat{I}$ em relação ao verdadeiro valor de I e em relação a variância de $\hat{I}$.