Correções: Resolução das Listas

Inferência Bayesiana | 1º/2024

Tailine J. S. Nonato

September 26, 2024

Sumário

Itens com alterações estão em **azul**.

Abaixo também é possível encontrar um resumo de alterações:

1. Lista 2
    - Questão 1 (iii)
    - Questão 2 (ii) e (iii)
    - Questão 4 (ii), (iii) e (iv)
    - Questão 8 (ii)
    - Questão 11 (iii) e (iv)
2. Lista 3.1
    - Questão 1 (iii)
    - Questão 3 (a), (b), (c) e (d)
    - Questão 6 (b)
3. Lista 3.2
    - Questão 1 (a), (b), (c)
    - Questão 2 (a), (b), (c)
4. Lista 4
    - Questão 2 (c)

- Questão 3 (a)

Lista 2

Questão 1

O seu professor chega na sala de aula e mostra uma moeda. Você suspeita que a moeda possa ser falsa e ter duas caras. Considere a priori probabilidades iguais para os eventos da moeda ser falsa ou ser honesta (i.e. uma moeda bem equilibrada).

(i) Calcule a sua probabilidade de obter cara num lançamento dessa moeda.

- Entende-se que:

$C = \{cara\}$

$H = \{honesta\}$

$F = \{falsa\}$

- Assim, a priori é dada por:

$P(F) = P(H) = 0.5$

- A probabilidade de obter cara num lançamento é dada por:

$P(C) = P(C|F)P(F) + P(C|H)P(H)$

$P(C) = 1 \cdot 0.5 + 0.5 \cdot 0.5$

$P(C) = 0.75$

(ii) Se o professor lançar a moeda e o resultado for cara, qual é agora a probabilidade dela ser falsa?

- A probabilidade de ser falsa dado que o resultado foi cara é dada por:

$P(F|C) = \frac{P(C|F)P(F)}{P(C)}$

$P(F|C) = \frac{1 \cdot 0.5}{0.75}$

$P(F|C) = \frac{0.5}{0.75}$

$P(F|C) = \frac{2}{3}$

(iii) Se o professor lançar a moeda n vezes e obter n caras, qual é a probabilidade dela ser falsa? Estude o comportamento desta probabilidade para n grande.

- Pelo Teorema de Bayes, a probabilidade de ser falsa dado que o resultado foi cara n vezes é dada por:

$P(F|C^n) = \frac{P(C^n|F)P(F)}{P(C^n)}$

Onde $P(C^n|F) = 1$, $P(F) = P(H) = 0.5$ e $P(C^n|H) = 0.5^n$

- Pela regra da probabilidade total, tem-se:

$P(C^n) = P(C^n|F)P(F) + P(C^n|H)P(H)$

$P(C^n) = 1 \cdot 0.5 + 0.5^n \cdot 0.5$

$P(C^n) = 0.5 + 0.5^{n+1}$

- Logo, substituindo:

$P(F|C^n) = \frac{1 \cdot 0.5}{0.5+0.5^{n+1}}$

$P(F|C^n) = \frac{0.5}{0.5+0.5^{n+1}}$

- Estudando o comportamento da probabilidade para n grande, tem-se:

$P(F|C^n) = \frac{0.5}{0.5+0.5^{n+1}}$

$P(F|C^n) = \frac{0.5}{0.5+0}$

$P(F|C^n) = 1$

(iv) Se o professor lançar a moeda uma vez e obter cara, qual é a probabilidade do próximo lançamento ser cara?

- A probabilidade do próximo lançamento ser cara dado que o resultado foi cara é dada por:

$P(C|C) = \frac{P(C|F)P(F)}{P(C)}$

$P(C|C) = \frac{1 \cdot 0.5}{0.75}$

$P(C|C) = \frac{0.5}{0.75}$

$P(C|C) = \frac{2}{3}$

(v) Explique porque é falso neste contexto a afirmação "os dois lançamentos da moeda são independentes", e explique qual seria a afirmação correta.

- A probabilidade do segundo lançamento ser cara depende do resultado do primeiro lançamento. A afirmação correta seria "os dois lançamentos da moeda são condicionalmente independentes", ou seja, a probabilidade do segundo lançamento ser cara dado que o primeiro foi cara é igual à probabilidade do segundo lançamento ser cara dado que o primeiro foi coroa.

Questão 2

Seja $y_1, y_2, ..., y_n$ uma amostra da distribuição de Bernoulli com probabilidade de sucesso θ e considere uma distribuição a priori uniforme para θ.

(i) Ache a distribuição a posteriori de θ e a sua média e variância.

- Entende-se que:

$s = \sum_{i=1}^{n} y_i \sim Binomial(n, \theta)$

$P(\theta) \sim Uniforme(0, 1)$

- Assim, a priori é dada por:

$P(\theta) = \frac{1}{b-a} = \frac{1}{1-0} = 1, \qquad 0 < \theta < 1$

- A posteriori pode ser obtida por:

$\theta|x \propto P(x|\theta)P(\theta)$

$\theta|x \propto \theta^s(1-\theta)^{n-s} * 1$

- Manipulando a expressão, tem-se:

$\theta|x \propto \theta^{s+1-1}(1-\theta)^{n-s+1-1}$

- Assim, a posteriori é dada por:

$\theta|x \sim Beta(s+1, n-s+1)$

- Logo, pela distribuição, a média e a variância da posteriori são dadas por:

$E(\theta|x) = \frac{s+1}{s+1+n-s+1} = \frac{s+1}{n+2}$

$Var(\theta|x) = \frac{(s+1)(n-s+1)}{(s+1+n-s+1)^2(s+1+n-s+1+1)} = \frac{(s+1)(n-s+1)}{(n+2)^2(n+3)}$

(ii) Mostre que é possível expressar a esperança a posteriori de θ da forma $(1-w)E(\theta) + w\hat{\theta}$, onde $E(\theta)$ e $\hat{\theta}$ são respectivamente a esperança a priori e a estimativa máximo verossímil de θ, e interprete este resultado.

- A esperança a priori de θ é dada por:

$E(\theta) = \frac{1}{2}$

- A estimativa máximo verossímil de θ é dada por:

$\hat{\theta} = \frac{s}{n}$

- Assim, a esperança a posteriori de θ pode ser expressa como:

$E(\theta|x) = \frac{s+1}{n+2} = \frac{n}{n+2} \cdot \frac{s}{n} + \frac{2}{n+2} \cdot \frac{1}{2}$

$E(\theta|x) = \frac{n}{n+2}\hat{\theta} + \frac{2}{n+2}E(\theta)$

- Sendo $w = \frac{n}{n+2}$, tem-se:

$E(\theta|x) = (1-w)E(\theta) + w\hat{\theta}$

(iii) Se y_{n+1} é uma observação futura deste processo de Bernoulli, ache a distribuição preditiva $p(y_{n+1}|y_1, ..., y_n)$.

- A distribuição preditiva de y_{n+1} é dada por:

$p(y_{n+1}|y_1, ..., y_n) = \int p(y_{n+1}|\theta)p(\theta|y_1, ..., y_n)d\theta$

- Para $y_{n+1} = 0$, tem-se:

$p(y_{n+1} = 0|y_1, ..., y_n) = \int (1-\theta) \cdot p(\theta|y_1, ..., y_n)d\theta$

$p(y_{n+1} = 0|y_1, ..., y_n) = \frac{1}{B(s+1,n-s+1)} \int (1-\theta) \cdot \theta^s(1-\theta)^{n-s}d\theta$

- A integral se assemelha à uma função Beta com parâmetros $s+1$ e $n-s+2$, logo:

$p(y_{n+1} = 0|y_1, ..., y_n) = \frac{B(s+1,n-s+2)}{B(s+1,n-s+1)}$

$p(y_{n+1} = 0|y_1, ..., y_n) = \frac{\Gamma(s+1)\Gamma(n-s+2)/\Gamma(n+3)}{\Gamma(s+1)\Gamma(n-s+1)/\Gamma(n+2)}$

$p(y_{n+1} = 0|y_1, ..., y_n) = \frac{n-s+1}{n+2}$

- Para $y_{n+1} = 1$, tem-se:

$p(y_{n+1} = 1|y_1, ..., y_n) = \int \theta \cdot p(\theta|y_1, ..., y_n)d\theta$

$p(y_{n+1} = 1|y_1, ..., y_n) = \frac{1}{B(s+1,n-s+1)} \int \theta \cdot \theta^s(1-\theta)^{n-s}d\theta$

- A integral se assemelha à uma função Beta com parâmetros $s+2$ e $n-s+1$, logo:

$p(y_{n+1} = 1|y_1, ..., y_n) = \frac{B(s+2,n-s+1)}{B(s+1,n-s+1)}$

$p(y_{n+1} = 1|y_1, ..., y_n) = \frac{\Gamma(s+2)\Gamma(n-s+1)/\Gamma(n+3)}{\Gamma(s+1)\Gamma(n-s+1)/\Gamma(n+2)}$

$p(y_{n+1} = 1|y_1, ..., y_n) = \frac{s+1}{n+2}$

Questão 4

No exercício 2, calcule

(i) a estimativa bayesiana para Perda Quadrática

- A perda quadrática é dada por:

$L(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2$

- Assim, a estimativa bayesiana para perda quadrática é dada por:

$E[L(\theta, \hat{\theta})|x] = E[(\theta - \hat{\theta})^2|x]$

$E[L(\theta, \hat{\theta})|x] = E[\theta^2 - 2\theta\hat{\theta} + \hat{\theta}^2|x]$

$E[L(\theta, \hat{\theta})|x] = E[\theta^2|x] - 2E[\theta\hat{\theta}|x] + E[\hat{\theta}^2|x]$

$E[L(\theta, \hat{\theta})|x] = E[\theta^2|x] - 2\hat{\theta}E[\theta|x] + \hat{\theta}^2$

$E[L(\theta, \hat{\theta})|x] = Var(\theta|x) + E[\theta|x]^2 - 2\hat{\theta}E[\theta|x] + \hat{\theta}^2$

$E[L(\theta, \hat{\theta})|x] = \frac{(s+1)(n-s+1)}{(n+2)^2(n+3)} + \left(\frac{s+1}{n+2}\right)^2 - 2\hat{\theta}\frac{s+1}{n+2} + \hat{\theta}^2$

- Isolando $\hat{\theta}$, tem-se:

$\hat{\theta} = \frac{n}{n+2}\hat{\theta} + \frac{2}{n+2}\frac{1}{2}$

$\hat{\theta} = \frac{n}{n+2}\hat{\theta} + \frac{1}{n+2}$

$\hat{\theta} = \frac{n}{n+2}\frac{s}{n} + \frac{1}{n+2}$

$\hat{\theta} = \frac{s}{n+2} + \frac{1}{n+2}$

$\hat{\theta} = \frac{s+1}{n+2}$

(ii) o limite da estimativa bayesiana para Perda Zero-Um quando $\epsilon \to \theta$.

- A perda zero-um é dada por:

$$L(\theta, \hat{\theta}) = \begin{cases} 0, & se \quad \theta = \hat{\theta} \\ 1, & se \quad \theta \neq \hat{\theta} \end{cases}$$

- Assim, a estimativa bayesiana para perda zero-um é dada por:

$E[L(\theta,\hat{\theta})|\underline{y}] = E[I(\theta \neq \hat{\theta})|\underline{y}]$

$E[L(\theta,\hat{\theta})|\underline{y}] = P(\theta \neq \hat{\theta}|\underline{y})$

$E[L(\theta,\hat{\theta})|\underline{y}] = 1 - P(\theta = \hat{\theta}|\underline{y})$

No caso especial que $n = 12, s = \sum_{i=1}^{12} y_i = 9$, calcule (iii) a estimativa bayesiana sob Perda Absoluta.

- O objetivo é minimizar o valor esperado da perda absoluta, que é dada por:

$L(\theta,\hat{\theta}) = |\theta - \hat{\theta}|$

- O estimador de Bayes sob perda absoluta é a mediana da distribuição a posteriori de θ. Nesse caso, a distribuição a posteriori de θ é dada por:

$\theta|\underline{y} \sim Beta(s+1, n-s+1)$

- No caso em que $s = 9$ e $n = 12$ e $\underline{y}$ o vetor de observações $(y_1, y_2, ..., y_{12})$, tem-se:

$\theta|\underline{y} \sim Beta(10, 4)$

- Utilizando o R, é possível aproximar a mediana da distribuição a posteriori de θ de modo que:

```
qbeta(0.5, 10, 4)
```

```
[1] 0.7247242
```

- Logo, a estimativa bayesiana sob perda absoluta é dada por:

$\hat{\theta} = Mediana = 0.725$

(iv) o intervalo HPD com credibilidade 0.99.

- O intervalo HPD é dado pelos limites a e b tais que:

$\int_a^b f(\theta|x)d\theta = 0.99$

- Onde $f(\theta|x)$ é a distribuição a posteriori de θ.
- Assim,

$HPD = [\theta_{(1-\alpha)/2}, \theta_{(1+\alpha)/2}]$

$HPD = [\theta_{0.005}, \theta_{0.995}]$

- Com $s = 9$ e $n = 12$, tem-se:

$\theta|x \sim Beta(10, 4)$

- Logo, o intervalor é:

```
qgamma(c(0.005, 0.995), 10, 4)
```

```
[1] 0.9292305 4.9996058
```

Questão 8

Suponha que (x_1, x_2, x_3) dado p_1, p_2, p_3 segue uma distribuição Multinomial com parâmetros n e (p_1, p_2, p_3), onde $p_i \geq 0$ e $p_1 + p_2 + p_3 = 1$, e que, a priori, (p_1, p_2, p_3) segue uma distribuição de Dirichlet com parâmetros $(\alpha_1, \alpha_2, \alpha_3)$.

(i) Ache a distribuição a posteriori de p_1, p_2, p_3 e as distribuições a posteriori marginais de p_i $(i = 1, 2, 3)$

- Entende-se que:

$(x_1, x_2, x_3) \sim Multinomial(n, (p_1, p_2, p_3))$

$P(p_1, p_2, p_3) \sim Dirichlet(\alpha_1, \alpha_2, \alpha_3)$

- Assim, a priori é dada por:

$P(p_1, p_2, p_3) = \frac{1}{B(\alpha_1, \alpha_2, \alpha_3)} p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1} p_3^{\alpha_3 - 1}$

- A posteriori pode ser obtida por:

$p_1, p_2, p_3|x \propto P(x|p_1, p_2, p_3)P(p_1, p_2, p_3)$

$p_1, p_2, p_3|x \propto p_1^{x_1} p_2^{x_2} p_3^{x_3} p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1} p_3^{\alpha_3 - 1}$

$p_1, p_2, p_3|x \propto p_1^{x_1 + \alpha_1 - 1} p_2^{x_2 + \alpha_2 - 1} p_3^{x_3 + \alpha_3 - 1}$

- Assim, a posteriori é dada por:

$p_1, p_2, p_3|x \sim Dirichlet(x_1 + \alpha_1, x_2 + \alpha_2, x_3 + \alpha_3)$

- As distribuições a posteriori marginais de p_i são dadas por:

$p_i|x = \int p_1, p_2, p_3|x dp_1 dp_2 dp_3$

$p_i|x = \int p_1, p_2, p_3|x \prod_{j \neq i} dp_j$

$p_i|x = \int p_1^{x_1+\alpha_1-1} p_2^{x_2+\alpha_2-1} p_3^{x_3+\alpha_3-1} dp_1 dp_2 dp_3$

$p_i|x = \int p_i^{x_i+\alpha_i-1} \left(1-p_i\right)^{n-x_i+\alpha_{-i}-1} dp_i$

- Logo, é possível observar que:

$p_i|x = \frac{B(x_i+\alpha_i, n-x_i+\alpha_{-i})}{B(\alpha_i, \alpha_{-i})}$

- Então, as distribuições a posteriori marginais de p_i são dadas por:

$p_i|x \sim Beta(x_i + \alpha_i, n - x_i + \alpha_{-i})$

(ii) Calcule as estimativas bayesianas de p_i e de $p_j - p_i$ sob Perda Quadrática $(i, j = 1, 2, 3, i \neq j)$.

- As estimativas bayesianas de p_i sob perda quadrática são dadas por:

$E[L(p_i, \hat{p_i})|x_i] = \frac{\alpha_i+x_i}{\alpha_1+x_1+\alpha_2+x_2+\alpha_3+x_3}$

- Logo, as estimativas bayesianas de $p_j - p_i$ sob perda quadrática são dadas por:

$E[L(p_j - p_i, \hat{p_j} - \hat{p_i})|x] = \frac{\alpha_j+x_j}{\alpha_1+x_1+\alpha_2+x_2+\alpha_3+x_3} - \frac{\alpha_i+x_i}{\alpha_1+x_1+\alpha_2+x_2+\alpha_3+x_3}$

Questão 11

É conhecido que 25% dos pacientes de um certo grupo que sofrem de enxaqueca melhoram após duas horas de serem tratados com um placebo. Para verificar se uma droga nova é melhor que o placebo, $n = 20$ pacientes foram tratados com o placebo e verificou-se que após duas horas $s = 8$ deles relataram ter melhorado. Seja θ a probabilidade de um paciente tratado com a droga nova melhorar após duas horas.

(i) Especifique a hipótese nula H_0 e a alternativa H_1;

$$\begin{cases} H_0 : \theta \leq 0.25 \\ H_1 : \theta > 0.25 \end{cases}$$

(ii) Usando a distribuição a priori "não informativa" $\theta \sim Uniforme(0, 1)$, calcule as chances relativas a priori e a posteriori de H_1 e o correspondente Fator de Bayes;

- Entende-se que:

$\sum X_i \sim^{iid} Binomial(20, \theta)$

$P(\theta) \sim Uniforme(0, 1)$

- Assim, a priori é dada por:

$P(\theta) = \frac{1}{b-a} = \frac{1}{1-0} = 1, \qquad 0 < \theta < 1$

- Logo, as probabilidades a priori são dadas por:

$P(H_0) = P(\theta \leq 0.25) = \int_0^{0.25} 1 d\theta = 0.25$

$P(H_1) = P(\theta > 0.25) = \int_{0.25}^{1} 1 d\theta = 0.75$

- Odds a priori

$odds(\theta \leq 0.25) = \frac{P(\theta \leq 0.25)}{P(\theta > 0.25)} = \frac{0.25}{0.75} = \frac{1}{3}$

$odds(\theta > 0.25) = \frac{P(\theta > 0.25)}{P(\theta \leq 0.25)} = \frac{0.75}{0.25} = 3$

- Pode-se concluir assim que a priori a chance de H_1 é 3 vezes maior que a chance de H_0.
- Sabe-se que:

$\theta | x \propto P(x|\theta)P(\theta)$

$\theta | x \propto \theta^s (1-\theta)^{n-s} * 1$

- Manipulando a expressão, tem-se:

$\theta | x \propto \theta^{s+1-1} (1-\theta)^{n-s+1-1}$

- Assim, a posteriori é dada por:

$\theta | x \sim Beta(s+1, n-s+1)$

- Com $s = 8$ e $n = 20$, tem-se:

$\theta | x \sim Beta(9, 13)$

- Logo, as probabilidades a posteriori são dadas por:

$P(H_0 | x) = P(\theta \leq 0.25 | s = 8)$

```
ph0<-round(pbeta(0.25, 9, 13),3)
ph0
```

```
[1] 0.056
```

$P(H_1|x) = P(\theta > 0.25|s = 8)$

```r
ph1<-round(1 - pbeta(0.25, 9, 13),3)
ph1
```

```
[1] 0.944
```

- Odds a posteriori

$odds(\theta \leq 0.25) = 0.056/\ 0.944 = 0.059$

$odds(\theta > 0.25) = 0.944/\ 0.056 = 16.857$

- Pode-se concluir assim que a posteriori a chance de H_1 é 16.86 vezes maior que a chance de H_0.
- Fator de Bayes

$\beta_{1,0} = \frac{odds(\theta_1|x)}{odds(\theta_1)} = 16.8/3 = 5.6$ em favor de H_1.

$\beta_{0,1} = \frac{odds(\theta_0|x)}{odds(\theta_0)} = 3/16.8 = 0.18$ em favor de H_0.

- Assim, a interpretação do Fator de Bayes é que a evidência a favor de H_1 é 5.6 vezes mais forte que a evidência a favor de H_0.

(iii) Seja $d = 1$ a decisão de rejeitar H_0 e $d = 0$ a de não rejeitar. Considere a função de perda de Neyman para a qual é 5 vezes mais custoso rejeitar H_0 quando ela é verdadeira do que não rejeitar quando ela é falsa [isto é, $L(d = 1, \theta \in H_0) = 5L(d = 0, \theta \notin H_0)$, $L(d = 1, \theta \notin H_0) = L(d = 0, \theta \in H_0) = 0$]. Calcule a decisão ótima a posteriori;

- Entende-se que:

$$L(d = 1, \theta \in H_0) = L(d = 0, \theta \notin H_0) = 5$$

$$L(d = 1, \theta \notin H_0) = L(d = 0, \theta \in H_0) = 0$$

- A função de perda de Neyman é dada por:

$a = L(d = 1, \theta < 0.25) = 5 \cdot L(d = 0, \theta > 0.25) = 5b$

- Assim, a decisão ótima a posteriori é dada por:

$P(\theta \leq 0.25|x) > \frac{a}{a+b}$

$P(\theta \leq 0.25|x) > \frac{5}{5+1}$

$P(\theta \leq 0.25|x) > \frac{5}{6}$

$P(\theta \leq 0.25|x) > 0.833$

(iv) É razoável chamar essa distribuição a priori de "não informativa" nesse problema? Se a sua resposta for negativa, sugira uma outra distribuição a priori e refaça os cálculos anteriores.

- É razoável, mas é necessário levar em conta que ao assumir uma distribuição a priori uniforme, supõe-se que todas as probabilidades são igualmente prováveis, favorecendo a hipótese alternativa. Conforme o item (ii), a chance de H_1 é 3 vezes mais forte que a chance de H_0. Assim, uma possibilidade para equilibrar as hipóteses seria assumir uma distribuição a priori Beta(1,3). Nesse caso, a priori é dada por:

$P(\theta) = \frac{1}{B(1,3)}\theta^{1-1}(1-\theta)^{3-1}$

$P(\theta) = 3(1-\theta)^2$

- Logo, as probabilidades a priori são dadas por:

$P(H_0) = P(\theta \leq 0.25) = \int_0^{0.25} 3(1-\theta)^2 d\theta = 0.578$

$P(H_1) = P(\theta > 0.25) = \int_{0.25}^{1} 3(1-\theta)^2 d\theta = 0.422$

- Odds a priori

$odds(\theta \leq 0.25) = \frac{P(\theta \leq 0.25)}{P(\theta > 0.25)} = \frac{0.578}{0.422} = 1.37$

$odds(\theta > 0.25) = \frac{P(\theta > 0.25)}{P(\theta \leq 0.25)} = \frac{0.422}{0.578} = 0.73$

- Pode-se concluir assim que a priori a chance de H_0 é 1.37 vezes maior que a chance de H_1.
- Sabe-se que:

$\theta|x \sim Beta(s+1, n-s+1)$

- Com $s = 8$ e $n = 20$, tem-se:

$\theta|x \sim Beta(9, 13)$

- Logo, as probabilidades a posteriori são dadas por:

$P(H_0|x) = P(\theta \leq 0.25|s = 8)$

```r
ph0<-round(pbeta(0.25, 9, 13),3)
ph0
```

```
[1] 0.056
```

$P(H_1|x) = P(\theta > 0.25|s = 8)$

```r
ph1<-round(1 - pbeta(0.25, 9, 13),3)
ph1
```

```
[1] 0.944
```

- Odds a posteriori

$odds(\theta \leq 0.25) = 0.056/\ 0.944 = 0.059$

$odds(\theta > 0.25) = 0.944/\ 0.056 = 16.857$

- Pode-se concluir assim que a posteriori a chance de H_1 é 16.857 vezes maior que a chance de H_0.

- Fator de Bayes

$\beta_{1,0} = \frac{odds(\theta_1|x)}{odds(\theta_1)} = 16.8/0.73 = 23.01$ em favor de H_1.

$\beta_{0,1} = \frac{odds(\theta_0|x)}{odds(\theta_0)} = 0.059/1.37 = 0.043$ em favor de H_0.

Lista 3.1

Questão 1

1. Seja $x_1, x_2, ..., x_n$ uma amostra da distribuição de Poisson com média θ, e considere a priori que θ tem uma distribuição Gama com parâmetros α e β (ou seja, com média $\frac{\alpha}{\beta}$ e variância $\frac{\alpha}{\beta^2}$).

(i) Ache a distribuição a posteriori de θ e sua média e variância.

- Entende-se que:

$x_i \sim Poisson(\theta)$

$\theta \sim Gama(\alpha, \beta)$

- Assim, a priori é dada por:

$P(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}$

- A posteriori pode ser obtida por:

$\theta|x \propto P(x|\theta)P(\theta)$

$\theta|x \propto \theta^{\sum x_i}e^{-n\theta}\theta^{\alpha-1}e^{-\beta\theta}$

$\theta|x \propto \theta^{\sum x_i+\alpha-1}e^{-(n+\beta)\theta}$

- Assim, a posteriori é dada por:

$\theta|x \sim Gama(\sum x_i + \alpha, n + \beta)$

- Logo, a média e a variância da posteriori são dadas por:

$E(\theta|x) = \frac{\sum x_i+\alpha}{n+\beta}$

$Var(\theta|x) = \frac{\sum x_i+\alpha}{(n+\beta)^2}$

(ii) Mostre que é possível expressar a esperança a posteriori de θ da forma $w\bar{x} + (1-w)\frac{\alpha}{\beta}$, e interprete este resultado.

- A esperança a priori de θ é dada por:

$E(\theta) = \frac{\alpha}{\beta}$

- Assim, a esperança a posteriori de θ pode ser expressa como:

$E(\theta|x) = \frac{\sum x_i+\alpha}{n+\beta} = \frac{n}{n+\beta}\frac{\sum x_i}{n} + \frac{\beta}{n+\beta}\frac{\alpha}{\beta}$

$E(\theta|x) = \frac{n}{n+\beta}\bar{x} + \frac{\beta}{n+\beta}\frac{\alpha}{\beta}$

- Sendo $w = \frac{n}{n+\beta}$, tem-se:

$E(\theta|x) = w\bar{x} + (1-w)\frac{\alpha}{\beta}$

- Assim, a interpretação deste resultado é que a esperança a posteriori de θ é uma combinação linear entre a média amostral $\bar{x}$ e a média a priori $\frac{\alpha}{\beta}$, onde o peso w é dado pela razão entre o tamanho amostral n e a soma do tamanho amostral com o parâmetro β.

(iii) O que acontece na parte (ii) quando β é grande com $\frac{\alpha}{\beta}$ fixo? Interprete!

$$\lim_{\beta\to\infty} w = \lim_{\beta\to\infty} \frac{n}{n+\beta} = \frac{n}{\beta} = 0$$

- Logo, se β é grande, o peso w tende a 0, de forma que a esperança a posteriori de θ se aproxima da média a priori $\frac{\alpha}{\beta}$:

$E(\theta|x) = w\bar{x} + (1-w)\frac{\alpha}{\beta}$

$E(\theta|x) = 0 + (1-0)\frac{\alpha}{\beta}$

$E(\theta|x) = \frac{\alpha}{\beta}$

- Assim, a interpretação é que a medida que o parâmetro β aumenta, a influência da informação amostral diminui e a estimativa se aproxima da média a priori $\frac{\alpha}{\beta}$.

(iv) Mostre que existe um número c tal que a variância a posteriori é maior do que a variância a priori sempre que $\bar{x} > c$, ache c e interprete este resultado.

- A variância a priori de θ é dada por:

$Var(\theta) = \frac{\alpha}{\beta^2}$

- Assim, a variância a posteriori de θ é dada por:

$Var(\theta|x) = \frac{\sum x_i + \alpha}{(n+\beta)^2}$

- Para que a variância a posteriori seja maior que a variância a priori, é necessário que:

$\frac{\sum x_i + \alpha}{(n+\beta)^2} > \frac{\alpha}{\beta^2}$

$\frac{\sum x_i + \alpha}{n+\beta} > \frac{\alpha}{\beta}$

$\frac{\sum x_i}{n} + \frac{\alpha}{n+\beta} > \frac{\alpha}{\beta}$

$\frac{\sum x_i}{n} > \frac{\alpha}{\beta} - \frac{\alpha}{n+\beta}$

$\frac{\sum x_i}{n} > \frac{\alpha}{\beta} - \frac{\alpha}{n+\beta}$

$\frac{\sum x_i}{n} > \frac{\alpha(n+\beta) - \alpha\beta}{\beta(n+\beta)}$

$\frac{\sum x_i}{n} > \frac{\alpha n}{\beta(n+\beta)}$

$\frac{\sum x_i}{n} > \frac{\alpha}{\beta} \cdot \frac{n}{n+\beta}$

- Assim, o número c é dado por:

$c = \frac{\alpha}{\beta}$

- A interpretação deste resultado é que a variância a posteriori de θ é maior que a variância a priori sempre que a média amostral $\bar{x}$ for maior que a razão entre os parâmetros α e β, ou seja, a variância a posteriori é maior que a variância a priori quando a média amostral é maior que a média a priori.

Questão 3

Seja $x_1, x_2, ..., x_n$ uma amostra da distribuição de Poisson com média θ, e considerea priori que θ tem uma distribuição Gama com parâmetros $\alpha = 1$ e $\beta = 1$. Ache:

(a) a estimativa bayesiana de θ no caso de perda quadrática

- Para um priori Gama com $\alpha = 1$ e $\beta = 1$, tem-se uma posteriori:

$\theta|x \sim Gama(\sum x_i + 1, n + 1)$

- Assim, a estimativa bayesiana de θ no caso de perda quadrática é dada por:

$\hat{\theta} = \frac{\sum x_i + 1}{n+1}$

(b) o limite do estimativa bayesiana sob perda zero-um quando $\epsilon \to 0$.

- A perda zero-um é dada por:

$\hat{\theta} = \frac{\sum x_i + 1 - 1}{n+1}$

$\hat{\theta} = \frac{\sum x_i}{n+1}$

Para o caso $n = 10$ e $\bar{x} = 1.55$, ache: (c) a estimativa bayesiana sob perda absoluta

- Para n=10 e $\bar{x} = 1.55$, a estimativa bayesiana sob perda absoluta é dada por:

$\alpha = \sum x_i + 1 = 10 \cdot 1.55 + 1 = 16.5$

$\beta = n + 1 = 10 + 1 = 11$

- Logo, a estimativa bayesiana sob perda absoluta é dada pela mediana da distribuição a posteriori de θ:

```
qgamma(0.5, 16.5, 11)
```

```
[1] 1.469808
```

(d) o intervalo HPD para θ com nível 95%

- O intervalo HPD é dado pelos limites a e b tais que:

$\int_a^b f(\theta|x)d\theta = 0.95$

- Onde $f(\theta|x)$ é a distribuição a posteriori de θ, Gamma(16.5, 11).

```r
qgamma(c(0.025, 0.975), 16.5, 11)
```

```
[1] 0.8657573 2.3056855
```

Questão 5

Seja $x_1, x_2, ..., x_n$ uma amostra da distribuição Normal com média μ e variância ϕ^{-1} conhecida, e considere a distribuição a priori $\mu \sim N(\mu_0, \tau^{-1})$.

(i) Ache a distribuição a posteriori de μ.

- Entende-se que:

$x_i \sim Normal(\mu, \phi^{-1})$

$\mu \sim Normal(\mu_0, \tau^{-1})$

- Assim, a priori é dada por:

$P(\mu) = \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{1}{2\tau}(\mu-\mu_0)^2}$

- A posteriori pode ser obtida por:

$\mu|x \propto P(x|\mu)P(\mu)$

$\mu|x \propto \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\phi^{-1}}} e^{-\frac{1}{2\phi^{-1}}(x_i-\mu)^2} \cdot \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{1}{2\tau}(\mu-\mu_0)^2}$

$\mu|x \propto e^{-\frac{1}{2\phi^{-1}}\sum_{i=1}^{n}(x_i-\mu)^2} \cdot e^{-\frac{1}{2\tau}(\mu-\mu_0)^2}$

$\mu|x \propto e^{-\frac{1}{2\phi^{-1}}\sum_{i=1}^{n}(x_i^2-2x_i\mu+\mu^2)} \cdot e^{-\frac{1}{2\tau}(\mu^2-2\mu\mu_0+\mu_0^2)}$

$\mu|x \propto e^{-\frac{1}{2\phi^{-1}}\sum_{i=1}^{n}(x_i^2-2x_i\mu+\mu^2)-\frac{1}{2\tau}(\mu^2-2\mu\mu_0+\mu_0^2)}$

- Logo,

$\mu|x \sim Normal(\frac{\sum x_i/\phi+\mu_0/\tau}{n/\phi+1/\tau}, \frac{1}{n/\phi+1/\tau})$

(ii) Mostre que é possível expressar a esperança a posteriori de μ da forma $w\bar{x} + (1-w)\mu_0$, e interprete este resultado.

- A esperança a priori de μ é dada por:

$E(\mu) = \mu_0$

- Assim, a esperança a posteriori de μ pode ser expressa como:

$E(\mu|x) = \frac{\sum x_i/\phi + \mu_0/\tau}{n/\phi + 1/\tau}$

$E(\mu|x) = \frac{n\mu_0 + \sum x_i}{n + \phi/\tau}$

$E(\mu|x) = \frac{n}{n+\phi/\tau}\bar{x} + \frac{\phi/\tau}{n+\phi/\tau}\mu_0$

- Sendo $w = \frac{n}{n+\phi/\tau}$, tem-se:

$E(\mu|x) = w\bar{x} + (1-w)\mu_0$

(iii) Se $\bar{x}_m$ é a média de m observações futuras $x_{n+1}, ..., x_{n+m}$, condicionalmente independentes de $x_1, ..., x_n$, ache a distribuição preditiva $p(\bar{x}_m|x_1, ..., x_n)$.

- Entende-se que:

$\bar{x}_m = \frac{\sum_{i=n+1}^{n+m} x_i}{m}$

- Assim, a distribuição preditiva é dada por:

$p(\bar{x}_m|x) = \int p(\bar{x}_m|\mu)p(\mu|x)d\mu$

$p(\bar{x}_m|x) = \int Normal(\bar{x}_m|\mu, \phi^{-1}/m)Normal(\mu|\frac{\sum x_i/\phi + \mu_0/\tau}{n/\phi + 1/\tau}, \frac{1}{n/\phi + 1/\tau})d\mu$

$p(\bar{x}_m|x) = \int \frac{1}{\sqrt{2\pi\phi^{-1}/m}} e^{-\frac{1}{2\phi^{-1}/m}(\bar{x}_m - \mu)^2} \cdot \frac{1}{\sqrt{2\pi\frac{1}{n/\phi+1/\tau}}} e^{-\frac{1}{2\frac{1}{n/\phi+1/\tau}}(\mu - \frac{\sum x_i/\phi + \mu_0/\tau}{n/\phi+1/\tau})^2} d\mu$

$p(\bar{x}_m|x) = \int \frac{1}{\sqrt{2\pi\phi^{-1}/m}} e^{-\frac{1}{2\phi^{-1}/m}(\bar{x}_m - \mu)^2} \cdot \frac{1}{\sqrt{2\pi\frac{1}{n/\phi+1/\tau}}} e^{-\frac{1}{2\frac{1}{n/\phi+1/\tau}}(\mu - \frac{\sum x_i/\phi + \mu_0/\tau}{n/\phi+1/\tau})^2} d\mu$

- Assim,

$p(\bar{x}_m|x) \sim Normal(\bar{x}_m|\frac{\sum x_i/\phi + \mu_0/\tau}{n/\phi + 1/\tau}, \frac{1}{n/\phi + 1/\tau} + \phi^{-1}/m)$

(iv) Discuta o que acontece com os resultados anteriores quando a distribuição a priori $p(\mu) \propto 1$ (ou seja, o caso limite quando $\tau \rightarrow 0$).

- Quando a distribuição a priori é uniforme, a informação a priori é desconsiderada e a estimativa é baseada apenas na informação amostral. Assim, a média a posteriori de μ é dada pela média amostral $\bar{x}$, a variância a posteriori de μ é dada pela variância amostral ϕ^{-1}/n e a distribuição preditiva é dada por:

$p(\bar{x}_m|x) \sim Normal(\bar{x}_m|\bar{x}, \phi^{-1}/n + \phi^{-1}/m)$

- Assim, a interpretação é que a medida que o parâmetro τ tende a zero, a influência da informação a priori diminui e a estimativa se aproxima da média amostral, a variância a posteriori de μ diminui e a distribuição preditiva é dada pela média amostral e pela variância amostral.

Questão 6

Seja $x_1, x_2, ..., x_n$ uma amostra da distribuição Normal com média μ e variância ϕ^{-1} conhecida, e considere a distribuição a priori $\mu \sim N(\mu_0, \tau^{-1})$.

(a): Ache o estimador bayesiano de μ no caso de

(i) perda quadrática $(L(\hat{\mu}, \mu) = (\hat{\mu} - \mu)^2)$,

- Para encontrar $\hat{\mu}$ sob perda quadrática, é necessário minimizar a perda esperada:

$E[L(\hat{\mu}, \mu)|x] = E[(\hat{\mu} - \mu)^2|x]$

$E[L(\hat{\mu}, \mu)|x] = E[\hat{\mu}^2 - 2\hat{\mu}\mu + \mu^2|x]$

$E[L(\hat{\mu}, \mu)|x] = E[\hat{\mu}^2|x] - 2E[\hat{\mu}\mu|x] + E[\mu^2|x]$

$E[L(\hat{\mu}, \mu)|x] = Var(\hat{\mu}|x) + E[\hat{\mu}|x]^2 - 2\mu E[\hat{\mu}|x] + \mu^2$

$E[L(\hat{\mu}, \mu)|x] = \frac{1}{n/\phi + 1/\tau} + E[\hat{\mu}|x]^2 - 2\mu E[\hat{\mu}|x] + \mu^2$

- Assim, a estimativa bayesiana de μ sob perda quadrática é dada por:

$\hat{\mu} = E[\mu|x] = \frac{\sum x_i/\phi + \mu_0/\tau}{n/\phi + 1/\tau}$

(ii) perda absoluta $(L(\hat{\mu}, \mu) = |\hat{\mu} - \mu|)$

- Para encontrar $\hat{\mu}$ sob perda absoluta, o processo é similar:

$E[L(\hat{\mu}, \mu)|x] = E[|\hat{\mu} - \mu|x]$

- Logo,

$\hat{\mu} = E[\mu|x] = \frac{\sum x_i/\phi + \mu_0/\tau}{n/\phi + 1/\tau}$

(iii) perda zero-um $(L(\hat{\mu}, \mu) = I(|\hat{\mu} - \mu| \geq \epsilon))$.

- Para encontrar $\hat{\mu}$ sob perda zero-um, o processo também é similar:

$E[L(\hat{\mu}, \mu)|x] = E[I(|\hat{\mu} - \mu| \geq \epsilon)|x]$

- Assim,

$\hat{\mu} = E[\mu|x] = \frac{\sum x_i/\phi + \mu_0/\tau}{n/\phi + 1/\tau}$

(b): Ache o intervalo HPD para μ com nível $100(1 - \alpha)\%$

- O intervalo HPD é dado pelos limites a e b tais que:

$\int_a^b f(\theta|x) d\theta = 100(1 - \alpha)$

- Onde $f(\theta|x)$ é a distribuição a posteriori de θ. Nesse caso:

$$\mu|x \sim Normal(\frac{\sum x_i/\phi + \mu_0/\tau}{n/\phi + 1/\tau}, \frac{1}{n/\phi + 1/\tau})$$

- Assim, com $n = 10$, $\bar{x} = 1.55$, $\mu_0 = 1$, $\tau = 1$ e $\phi = 1$, o intervalo HPD para μ com nível $100(1-\alpha)\%$ é dado por:

```
arg_mu = (1 * (10*1.55) + 1 * 1) / (1 * 10 + 1)
arg_sd = 1/(sqrt(1 * 10 + 1))

qnorm(c(0.025,0.975), arg_mu, arg_sd)
```

```
[1] 0.9090486 2.0909514
```

Lista 3.2

Questão 1

Considere o modelo $X \sim$ Uniforme$(\theta, \theta+1)$ $(-\infty < \theta < \infty)$.

(a) Mostre que θ é um parâmetro de locação.

- Para mostrar que θ é um parâmetro de locação, é necessário verificar se a distribuição é invariante por translação, ou seja, se $Y = X - \theta$ tem a mesma distribuição que X.
- A densidade de probabilidade de X é dada por:

$$f(x|\theta) = \begin{cases} 1, & \text{se } \theta \leq x \leq (\theta+1) \\ 0, & \text{caso contrário} \end{cases}$$

- Logo, ao fazermos a transformação $Y = X - \theta$, o novo modelo é $Y \sim$ Uniforme$(0, 1)$.

$$P(Y) = \int_{\theta}^{\theta+1} \frac{1}{1} dx = 1$$

- Observa-se que X depende de θ apenas como uma translação. Mesmo se adicionarmos ou subtrairmos uma constante a θ, a distribuição se desloca, mas não perde sua forma. Portanto, podemos concluir que θ é um parâmetro de locação.

(b) Dada a priori não informativa usual para o modelo de locação, $p(\theta) \propto 1$ e uma amostra aleatória $X_1, ..., X_n$ do modelo acima, discuta a propriedade da distribuição a posteriori.

- A priori não informativa usual para o modelo de locação é dada por:

$p(\theta) \propto 1$

- Assim, a distribuição a posteriori de θ é dada por:

$\theta|x \propto P(x|\theta)P(\theta)$

$\theta|x \propto \prod_{i=1}^{n} \frac{1}{1} \cdot 1$

$\theta|x \propto 1$

- A função de verossimilhança é dada por:

$L(\theta|x) = \prod_{i=1}^{n} 1$, se $\theta \leq x \leq (\theta + 1)$

- Isso implica que o valor de θ deve ser pequeno o suficiente para que X_1 (menor valor de X) seja maior ou igual a θ e grande o suficiente para que X_n (ou o maior valor de X) seja menor ou igual a $\theta + 1$. Duas condições devem ser satisfeitas:

$$\begin{cases} \theta \leq min(X_1, X_2, ..., X_n) \\ \theta + 1 \geq max(X_1, X_2, ..., X_n) \end{cases}$$

Portanto,

- A posteriori é dada por:

$p(\theta|x) \propto L(\theta|x)p(\theta) = 1$

se $min(X_1, X_2, ..., X_n) \leq \theta \leq max(X_1, X_2, ..., X_n)$

- Pode se concluir então que a posteriori é dada por:

$\theta|x \sim U[min(X_1, ..., X_n), max(X_1, ..., X_n)]$

(c) Na situação da parte (b), ache:

(i) os estimadores bayesianos sob Perda Quadrática e sob Perda Absoluta;

- O estimador bayesiano de θ sob perda quadrática é dado por:

$\hat{\theta}_q = E[\theta|x] = \int \theta p(\theta|x) d\theta$

- O estimador bayesiano de θ sob perda absoluta é dado por:

$\hat{\theta}_a = \text{mediana}(\theta|x)$

- Sabendo que a distribuição a posteriori de θ é dada por $\theta|x \sim U[min(X_1, ..., X_n), max(X_1, ..., X_n)]$, tem-se que:

$\hat{\theta}_q = \hat{\theta}_a = \frac{min(X_1,...,X_n)+max(X_1,...,X_n)}{2}$

(ii) O intervalo HPD para θ com credibilidade $100(1-\alpha)\%$

- Nesse caso, o intervalo HPD será o intervalo central da distribuição de θ. Assim,

Limite inferior: $min(X_1, ..., X_n) - \frac{\alpha}{2} \cdot [min(X_1, ..., X_n) - (max(X_1, ..., X_n) - 1)]$

Limite superior: $max(X_1, ..., X_n) - \frac{\alpha}{2} \cdot [max(X_1, ..., X_n) - (min(X_1, ..., X_n) + 1)]$

- Ou seja, o intervalo HPD para θ com credibilidade $100(1-\alpha)\%$ é dado por:

$$HPD = [max(X_1, ..., X_n) - 1 + \frac{\alpha}{2}, min(X_1, ..., X_n) - \frac{\alpha}{2}]$$

- Considerando que é uma distribuição uniforme, se soubessemos os valores de $X_1, ..., X_n$, poderíamos calcular no R:

```
qunif(alpha/2, min(x), max(x))
qunif(1-alpha/2, min(x), max(x))
```

(iii) A probabilidade a posteriori do evento $\theta > \theta_0$.

Se $\theta_0 <$ max($X_1, ..., X_n$) - 1, então $P(\theta > \theta_0) = 1$.

Se $\theta_0 >$ mix($X_1, ..., X_n$), então $P(\theta > \theta_0) = 0$.

Se θ_0 estiver entre min($X_1, ..., X_n$) e max($X_1, ..., X_n$) - 1, então $P(\theta > \theta_0) = 1 - \frac{\theta_0 - min(X_1,...,X_n)}{max(X_1,...,X_n) - min(X_1,...,X_n)+1}$

Questão 2

Considere o modelo $X \sim$ Uniforme$(0, \theta)$ $(0 < \theta < \infty)$.

(a) Mostre que $\sigma = \theta^{-1}$ é um parâmetro de escala.

- Temos que um parâmetro de escala afeta o tamanho ou a amplitude da distribuição, sem alterar sua forma. Se σ é um parâmetro de escala, então multiplicar X por uma constante c deve alterar σ de forma proporcional a c.
- Fazendo uma transformação de variável $Y = cX$, onde c é uma constante, tem-se:

$Y \sim$ Uniforme$(0, c\theta)$

- Isso mostra que o parâmetro θ foi escalado por uma constante c. Assim, tem-se:

$\sigma_Y = c\theta^{-1} = \frac{1}{c} \cdot \sigma$

- σ foi escalado por uma constante $1/c$, o que mostra que $\sigma = \theta^{-1}$ é um parâmetro de escala.

(b) Dada a priori não informativa usual para o modelo de escala, $p(\sigma) \propto \sigma^{-1}$ e uma amostra aleatória $X_1, ..., X_n$ do modelo acima, discuta a propriedade da distribuição a posteriori.

- Com $X_i \sim \text{Uniforme}(0, \theta)$, tem-se que:

$p(\theta|X_i) \propto p(X_i|\theta)p(\theta)$

$p(\theta|X_i) \propto \prod_{i=1}^{n} \frac{1}{\theta} \cdot \frac{1}{\theta}$

$p(\theta|X_i) \propto \frac{1}{\theta^n} \cdot \frac{1}{\theta}$

$p(\theta|X_i) \propto \theta^{-(n+1)}$

- Considerando a condição de que $\theta \geq max(X_1, ..., X_n)$, tem-se que:

$p(\theta|X_i) \propto \theta^{-(n+1)}$, se $\theta \geq max(X_1, ..., X_n)$

- Integrando sobre o intervalo de θ, tem-se:

$\int_{max(X_1,...,X_n)}^{\infty} \theta^{-(n+1)} d\theta = -\frac{1}{n\theta^n}\Big|_{max(X_1,...,X_n)}^{\infty}$

$\int_{max(X_1,...,X_n)}^{\infty} \theta^{-(n+1)} d\theta = \frac{1}{n(max(X_1,...,X_n))^n}$

- Encontrando a constante de normalização:

$c = \frac{n(max(X_1,...,X_n))^n}{\theta^{n+1}}$

- Assim, a distribuição a posteriori é dada por:

$p(\theta|X_i) = \frac{n(\max(X_1,...,X_n))^n}{\theta^{n+1}}$

- Nota-se que a distribuição a posteriori se assemelha a uma distribuição Pareto, logo:

$\theta|X_i \sim Pareto(n, max(X_1, ..., X_n))$

(c) Na situação da parte (b), ache:

(i) os estimadores bayesianos sob Perda Quadrática e sob Perda Absoluta;

- O estimador bayesiano de θ sob perda quadrática é dado por:

$\hat{\theta}_q = E[\theta|x] = \int \theta p(\theta|x) d\theta$

$\hat{\theta}_q = E[\theta|x] = \int \theta \frac{n(\max(X_1,...,X_n))^n}{\theta^{n+1}} d\theta$

$\hat{\theta}_q = E[\theta|x] = \frac{n(\max(X_1,...,X_n))^n}{n-1}$, com $n > 1$.

- O estimador bayesiano de θ sob perda absoluta é dado por:

$\hat{\theta}_a = \text{mediana}(\theta|x)$

$\hat{\theta}_a = \max(X_1, ..., X_n) \cdot 2^{1/n}$

(ii) O intervalo HPD para θ com credibilidade $100(1-\alpha)\%$

- O intervalo HPD é dado pelos limites a e b tais que:

$P(\theta_a < \theta < \theta_b | X_1, ..., X_n) = 1 - \alpha$

- Considerando que a distribuição a posteriori de θ é dada por $\theta|x \sim Pareto(n, max(X_1, ..., X_n))$, tem-se que:

$F(\theta_a) - F(\theta_b) = 1 - \alpha$

- Sabe-se que:

$F(\theta) = 1 - \left(\frac{\max(X_1,...,X_n)}{\theta}\right)^n$

- Assim, o intervalo HPD para θ com credibilidade $100(1-\alpha)\%$ é dado por:

$\theta_a = \max(X_1, ..., X_n)$

$\theta_b = \max(X_1, ..., X_n) \cdot \alpha^{-(\frac{1}{n+1})}$

- Ou seja, o intervalo HPD para θ com credibilidade $100(1-\alpha)\%$ é dado por:

$$HPD = [\max(X_1, ..., X_n), \max(X_1, ..., X_n) \cdot \alpha^{-(\frac{1}{n+1})}]$$

(iii) A probabilidade a posteriori do evento $\theta > \theta_0$.

- Tem-se que a probabilidade de $\theta > \theta_0$ é dada pela complementar à função de distribuição acumulada em θ_0:

$P(\theta > \theta_0 | X_1, ..., X_n) = 1 - F(\theta_0)$

$P(\theta > \theta_0 | X_1, ..., X_n) = 1 - \left[1 - \left(\frac{\max(X_1,...,X_n)}{\theta_0}\right)^n\right]$

$P(\theta > \theta_0 | X_1, ..., X_n) = \left(\frac{\max(X_1,...,X_n)}{\theta_0}\right)^n$, se $\theta_0 > \max(X_1, ..., X_n)$

Questão 3

Considere o modelo $x_1, ..., x_n|\theta \sim$ Ber(θ).

(a) Calcule a priori de Jeffreys e mostre que ela é própria.

- A priori de Jeffreys é dada por:

$p(\theta) \propto \sqrt{I(\theta)}$

$I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2} \log f(x|\theta)\right]$

$I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2} \log \theta^x(1-\theta)^{1-x}\right]$

$I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2} (x \log \theta + (1-x)\log(1-\theta))\right]$

$I(\theta) = -E\left[\frac{\partial}{\partial\theta}\left(\frac{x}{\theta} - \frac{1-x}{1-\theta}\right)\right]$

$I(\theta) = -E\left[\frac{-x}{\theta^2} + \frac{1-x}{(1-\theta)^2}\right]$

$I(\theta) = E\left[\frac{x}{\theta^2} + \frac{1-x}{(1-\theta)^2}\right]$

$I(\theta) = \frac{1}{\theta^2}E[x] + \frac{1}{(1-\theta)^2}E[1-x]$

$I(\theta) = \frac{1}{\theta^2}\theta + \frac{1}{(1-\theta)^2}(1-\theta)$

$I(\theta) = \frac{1}{\theta} + \frac{1}{1-\theta}$

- Assim, a priori de Jeffreys é dada por:

$p(\theta) \propto \sqrt{\frac{1}{\theta} + \frac{1}{1-\theta}}$

- Para mostrar que a priori de Jeffreys é própria, é necessário verificar se a integral da priori é finita:

$\int \sqrt{\frac{1}{\theta} + \frac{1}{1-\theta}} d\theta$

```r
integrate(function(theta) sqrt(1/theta + 1/(1-theta)), lower = 0, upper = 1)$value
```

```
[1] 3.141593
```

- Assim, como a integral é finita, a priori de Jeffreys é própria.

(b) Considere uma amostra aleatória $X_1, ..., X_n$ do modelo acima e a priori de Jeffreys. Ache o estimador bayesiano de θ sob Perda Quadrática e explique como achar um intervalo HPD para θ com credibilidade $100(1-\alpha)\%$.

- O estimador bayesiano de θ sob perda quadrática é dado por:

$\hat{\theta} = E[\theta|x] = \int \theta p(\theta|x) d\theta$

- Um exemplo de cálculo do intervalo HPD para θ com credibilidade $100(1-\alpha)\%$ pode ser dado por:

```
qbeta(alpha/2, sum(x), n - sum(x))
```

```
qbeta(1-alpha/2, sum(x), n - sum(x))
```

Lista 4

Questão 1

Considere uma amostra $y_1, ..., y_n$ da distribuição Normal com média μ e variância $\sigma^2 = 1/\tau$ desconhecidas, e suponha que a priori a distribuição de (μ, τ) é especificada da seguinte forma: $\mu|\tau \sim N(\mu_0, 1/\lambda_0\tau)$ e $\tau \sim$ Gama(α_0, β_0), onde λ_0, α_0 e β_0 são positivas.

(a) Ache a distribuição a posteriori de $p(\mu, \tau|D)$ e as distribuições marginais $p(\mu|D)$ e $p(\tau|D)$.

- A distribuição a posteriori de $p(\mu, \tau|D)$ é dada por:

$p(\mu, \tau|D) \propto p(D|\mu, \tau)p(\mu|\tau)p(\tau)$

$p(\mu, \tau|D) \propto \prod_{i=1}^n \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{1}{2\tau}(y_i-\mu)^2} \cdot \frac{1}{\sqrt{2\pi\lambda_0\tau}} e^{-\frac{1}{2\lambda_0\tau}(\mu-\mu_0)^2} \cdot \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} \tau^{\alpha_0-1} e^{-\beta_0\tau}$

$p(\mu, \tau|D) \propto \frac{1}{\sqrt{2\pi\tau}^n} e^{-\frac{1}{2\tau}\sum_{i=1}^n (y_i-\mu)^2} \cdot \frac{1}{\sqrt{2\pi\lambda_0\tau}} e^{-\frac{1}{2\lambda_0\tau}(\mu-\mu_0)^2} \cdot \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} \tau^{\alpha_0-1} e^{-\beta_0\tau}$

$p(\mu, \tau|D) \propto \tau^{n/2-1} e^{-\frac{1}{2\tau}\sum_{i=1}^n (y_i-\mu)^2} \cdot \tau^{1/2} e^{-\frac{1}{2\lambda_0\tau}(\mu-\mu_0)^2} \cdot \tau^{\alpha_0-1} e^{-\beta_0\tau}$

$p(\mu, \tau|D) \propto \tau^{n/2+1/2+\alpha_0-1} e^{-\frac{1}{2\tau}\sum_{i=1}^n (y_i-\mu)^2 - \frac{1}{2\lambda_0\tau}(\mu-\mu_0)^2 - \beta_0\tau}$

$p(\mu, \tau|D) \propto \tau^{n/2+1/2+\alpha_0-1} e^{-\frac{1}{2\tau}(\sum_{i=1}^n (y_i-\mu)^2 + \frac{1}{\lambda_0}(\mu-\mu_0)^2 + 2\beta_0)}$

- Assim, a distribuição a posteriori de $p(\mu, \tau|D)$ é dada por:

$p(\mu, \tau|D) \sim \text{Gama}(\alpha_0 + n/2, \frac{1}{2}(\sum_{i=1}^n y_i^2 + \frac{1}{\lambda_0}(\mu-\mu_0)^2 + 2\beta_0))$

(b) Discuta neste contexto o uso da distribuição a priori não informativa $p(\mu, \tau) \propto 1/\tau$. A distribuição a posteriori é própria? Qual é a relação da distribuição $p(\mu|D)$ e os resultados clássicos?

- A priori não informativa $p(\mu, \tau) \propto 1/\tau$ é uma distribuição imprópria, de forma que a distribuição a posteriori é própria. A distribuição a posteriori de μ é dada por:

$p(\mu|D) = \int p(\mu, \tau|D) d\tau$

- A distribuição a posteriori de μ é dada por:

$p(\mu|D) \sim t_{2\alpha_0+n}(\mu_0, \frac{1}{\lambda_0(\alpha_0+n)})$

Questão 2

Seja $n = (n_1, ..., n_k)$ um vetor aleatório com distribuição multinomial e densidade $p(n|\theta) \propto \prod_{i=1}^k \theta_i^{n_i}$, onde $\theta = (\theta_1, ..., \theta_k)$, $\theta_i > 0$ e $\sum_i \theta_i = 1$. Considere a priori para θ uma distribuição de Dirichlet com parâmetro $\alpha = (\alpha_1, ..., \alpha_k)$, isto é $p(\theta) \propto \prod_{i=1}^k \theta_i^{\alpha_i - 1}$.

(c) No caso particular $\alpha = (0, 0, ..., 0)$ a distribuição a priori de θ é imprópria. Mostre que a distribuição a posteriori é própria se e somente se $n_i > 0$ para $i = 1, 2, ..., k$.

- A densidade a posteriori é dada por:

$p(\theta|n) \propto p(n|\theta)p(\theta)$

$p(\theta|n) \propto \prod_{i=1}^k \theta_i^{n_i} \cdot \prod_{i=1}^k \theta_i^{-1}$

$p(\theta|n) \propto \prod_{i=1}^k \theta_i^{n_i - 1}$

- Sabe-se que para que uma distribuição Dirichlet com parâmetros $n_i - 1$ (os expoentes na posteriori) seja própria todos os parâmetros devem ser maiores que zero. Para a posteriori integre para 1 sobre o simplex $\{\theta \in \mathbb{R}^k : \sum_{i=1}^k \theta_i = 1, \theta_i > 0\}$, é necessário que todos os expoentes sejam positivos. Portanto, a posteriori será própria se:

$$\begin{cases} n_i - 1 > 0, & \text{para todo } i = 1, 2, ..., k \\ n_i > 1, & \text{para todo } i = 1, 2, ..., k \end{cases}$$

Questão 3

Na véspera do primeiro turno para a eleição de governador do DF em 2010, a Datafolha divulgou uma pesquisa indicando que, de 891 eleitores entrevistados que já tinham decidido em quem votar, Agnelo Queiroz tinha a preferência de 467, Weslian Roriz de 315 e outros candidatos de 109 eleitores. Formule um modelo para analisar esses dados. O interesse centra fundamentalmente em três perguntas:

(a) a eleição poderia ser definida no primeiro turno?

- Podemos iniciar a análise entendendo que:

$n = (n_1, n_2, n_3) \sim \text{Multinomial}(891, \theta)$, onde $n_1 = 467$, $n_2 = 315$ e $n_3 = 109$; $\theta = (\theta_1, \theta_2, \theta_3)$ e $\sum_{i=1}^{3} \theta_i = 1$.

- A distribuição a priori pode ser dada por:

$\theta \sim \text{Dirichlet}(\alpha)$, onde $\alpha = (\alpha_1, \alpha_2, \alpha_3)$.

- Para garantir que seja uma priori não informativa, podemos considerar $\alpha = (1, 1, 1)$. Assim, a distribuição a posteriori de θ é dada por:

$\theta|n \sim \text{Dirichlet}(n + \alpha)$

$\theta|n \sim \text{Dirichlet}(467 + 1, 315 + 1, 109 + 1)$

$\theta|n \sim \text{Dirichlet}(468, 316, 110)$

- Agora, sabe-se que eleição poderia ser definida no primeiro turno se o candidato Agnelo tivesse mais de 50% dos votos válidos. Portanto, devemos calcular $P(\theta > 0.5|n)$.

```r
pacman::p_load(gtools)

amostras <- rdirichlet(10000, c(468, 316, 110))
mean(amostras[, 1] > 0.5)
```

```
[1] 0.921
```

(b) O candidato Agnelo poderia ser eleito no primeiro turno?

- Assim, a probabilidade de Agnelo ser eleito no primeiro turno é dada por:

$P(\theta > 0.5) = 1 - \Phi\left(\frac{\log[\frac{0.5}{1-0.5}]-0.75}{\sqrt{1+782}}\right)$

$P(\theta > 0.5) = 1 - \Phi\left(\frac{\log[\frac{1}{2}]-0.75}{\sqrt{783}}\right)$

$P(\theta > 0.5) = 1 - \Phi\left(\frac{-0.6931-0.75}{\sqrt{783}}\right)$

$P(\theta > 0.5) = 1 - \Phi\left(\frac{-1.4431}{\sqrt{783}}\right)$

$P(\theta > 0.5) = 1 - \Phi\left(-0.0516\right)$

$P(\theta > 0.5) = 1 - 0.4794$

$P(\theta > 0.5) = 0.5206$

(c) qual será a diferença na porcentagem dos votos válidos entre os dois primeiros colocados?

- A diferença na porcentagem dos votos válidos entre os dois primeiros colocados é dada por:

$P(\theta_1 - \theta_2 > 0) = 1 - \Phi\left(\frac{\log[\frac{467}{782}1-\frac{467}{782}]-0.75}{\sqrt{1+782}}\right)$

- Calculando, o resultado é:

$P(\theta_1 - \theta_2 > 0) = 1 - \Phi\left(\frac{\log[\frac{467}{782}1-\frac{467}{782}]-0.75}{\sqrt{1+782}}\right)$

Lista 5

Questão 2

Considere $n = 12$ ensaios de Bernoulli com probabilidade de sucesso p e $y = 9$ sucessos. Suponha que a distribuição a priori de p é especificada de forma que $\eta = \log\left[\frac{p}{1-p}\right]$ segue uma distribuição Normal com média $\mu = 0$ e variância $\sigma^2 = 1$. Obtenha aproximações para $\mathbb{E}(p|y = 9)$, DP$(p|y = 9)$ e $\Pr(p > \frac{1}{2}|y = 9)$. Repita o exercício para os casos que $\sigma^2 = 4$ e 9 e compare com o caso $\sigma^2 = 1$.

- A distribuição a priori de η é dada por:

$p(\eta) = \frac{1}{\sqrt{2\pi}}e^{-\frac{\eta^2}{2}}$

- Assim, a distribuição a posteriori de η é dada por:

$p(\eta|y) \propto p(y|\eta)p(\eta)$

$p(\eta|y) \propto \binom{12}{9}e^{9\eta}(1-e^{\eta})^3e^{-\frac{\eta^2}{2}}$

- A distribuição a posteriori de η é dada por:

$p(\eta|y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\eta - 0.75)^2}{2(1+12)}}$

- Assim, a distribuição a posteriori de p é dada por:

$p(p|y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\log\left[\frac{p}{1-p}\right] - 0.75)^2}{2(1+12)}}$

- A esperança a posteriori de p é dada por:

$\mathbb{E}(p|y) = \frac{1}{1+e^{-0.75}}$

- A variância a posteriori de p é dada por:

$\text{Var}(p|y) = \frac{1}{(1+e^{-0.75})^2} \cdot \frac{e^{-0.75}}{(1+e^{-0.75})^2}$

- A probabilidade a posteriori de $p > \frac{1}{2}$ é dada por:

$\Pr(p > \frac{1}{2}|y) = 1 - \Phi\left(\frac{\log\left[\frac{1}{2}1 - \frac{1}{2}\right] - 0.75}{\sqrt{1+12}}\right)$

- Assim, para $\sigma^2 = 4$:

$p(p|y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\log\left[\frac{p}{1-p}\right] - 0.75)^2}{2(4+12)}}$

$\mathbb{E}(p|y) = \frac{1}{1+e^{-0.75}}$

$\text{Var}(p|y) = \frac{1}{(1+e^{-0.75})^2} \cdot \frac{e^{-0.75}}{(4+12)^2}$

$\Pr(p > \frac{1}{2}|y) = 1 - \Phi\left(\frac{\log\left[\frac{1}{2}1 - \frac{1}{2}\right] - 0.75}{\sqrt{4+12}}\right)$

- E, para $\sigma^2 = 9$:

$p(p|y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\log\left[\frac{p}{1-p}\right] - 0.75)^2}{2(9+12)}}$

$\mathbb{E}(p|y) = \frac{1}{1+e^{-0.75}}$

$\text{Var}(p|y) = \frac{1}{(1+e^{-0.75})^2} \cdot \frac{e^{-0.75}}{(9+12)^2}$

$\Pr(p > \frac{1}{2}|y) = 1 - \Phi\left(\frac{\log\left[\frac{1}{2}1 - \frac{1}{2}\right] - 0.75}{\sqrt{9+12}}\right)$