Lista de Exercícios 3

Disciplina: Inferência Estatística
Professor: Philip Thompson
Monitores: Eduardo Adame & Ezequiel Braga

Setembro 2023

Caps. 7.2 (Distribuições a Priori e a Posteriori)

1. **(DeGroot e Schervish 2014, ex. 2, pág. 393)** Suponha que a proporção θ de itens defeituosos em um grande lote fabricado seja conhecida como sendo 0.1 ou 0.2, e a função de probabilidade a priori (prior p.f.) de θ é a seguinte:

$$\xi(0.1) = 0.7 \text{ e } \xi(0.2) = 0.3.$$

   Suponha também que, quando oito itens são selecionados aleatoriamente do lote, é encontrado que exatamente dois deles são defeituosos. Determine a função de probabilidade a posteriori (posterior p.f.) de θ.

2. **(DeGroot e Schervish 2014, ex. 3, pág. 394)** Suponha que o número de defeitos em um rolo de fita magnética de gravação segue uma distribuição de Poisson, para a qual a média λ é igual a 1 ou 1.5, e a função de probabilidade a priori (prior p.f.) de λ é a seguinte:

$$\xi(1) = 0.4 \text{ e } \xi(1.5) = 0.6.$$

   Se um rolo de fita selecionado aleatoriamente for encontrado com três defeitos, qual é a função de probabilidade a posteriori (posterior p.f.) de λ?

3. **(DeGroot e Schervish 2014, ex. 10, pág. 394)** Suponha que uma única observação X deve ser retirada da distribuição uniforme no intervalo $[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$, o valor de θ é desconhecido, e a distribuição a priori de θ é uniforme no intervalo $[10, 20]$. Se o valor observado de X é 12, qual é a distribuição a posteriori de θ?

Cap. 7.3 (Distribuições a Priori Conjugadas)

4. **(DeGroot e Schervish 2014, ex. 17, pág. 407)** Suponha que o número de minutos que uma pessoa deve esperar por um ônibus todas as manhãs segue uma distribuição uniforme no intervalo $[0, \theta]$, onde o valor do ponto final θ é desconhecido. Suponha também que a função de densidade de probabilidade (p.d.f.) a priori de θ é a seguinte:

$$\xi(\theta) = \begin{cases} \frac{192}{\theta^4} & \text{para } \theta \geq 4, \\ 0 & \text{caso contrário.} \end{cases}$$

Se os tempos de espera observados em três manhãs consecutivas são 5, 3 e 8 minutos, qual é a função de densidade de probabilidade a posteriori (posterior p.d.f.) de θ?

5. **(DeGroot e Schervish 2014, ex. 19, pág. 407)** Suponha que $X_1, X_2, \ldots, X_n$ formam uma amostra aleatória de uma distribuição para a qual a função de densidade de probabilidade $f(x|\theta)$ é a seguinte:

$$f(x|\theta) = \begin{cases} \theta x^{\theta-1} & \text{para } 0 < x < 1, \\ 0 & \text{caso contrário.} \end{cases}$$

   Suponha também que o valor do parâmetro θ é desconhecido ($\theta > 0$), e a distribuição a priori de θ é a distribuição gama com parâmetros α e β ($\alpha > 0$ e $\beta > 0$). Determine a média e a variância da distribuição a posteriori de θ.

6. **(DeGroot e Schervish 2014, ex. 21, pág. 407)** Suponha que $X_1, X_2, \ldots, X_n$ formem uma amostra aleatória da distribuição exponencial com parâmetro θ. Seja a distribuição a priori de θ imprópria com "p.d.f." $\frac{1}{\theta}$ para $\theta > 0$. Encontre a distribuição a posteriori de θ e mostre que a média a posteriori de θ é $\frac{1}{\bar{X}_n}$.

Cap. 7.4 (Estimadores de Bayes)

7. **(DeGroot e Schervish 2014, ex. 2, pág. 416)** Suponha que a proporção θ de itens defeituosos em um grande envio seja desconhecida, e a distribuição a priori de θ seja a distribuição beta com parâmetros $\alpha = 5$ e $\beta = 10$. Suponha também que 20 itens sejam selecionados aleatoriamente do envio e que exatamente um desses itens seja encontrado defeituoso. Se a função de perda de erro quadrático for utilizada, qual é o estimador de Bayes de θ?

8. **(DeGroot e Schervish 2014, ex. 4, pág. 416)** Suponha que uma amostra aleatória de tamanho n seja retirada da distribuição de Bernoulli com parâmetro θ, que é desconhecido, e que a distribuição a priori de θ seja uma distribuição beta cuja média é μ_0. Mostre que a média da distribuição a posteriori de θ será uma média ponderada com a forma $\gamma_n \bar{X}_n + (1 - \gamma_n)\mu_0$, e mostre que $\gamma_n \to 1$ conforme $n \to \infty$.

9. **(DeGroot e Schervish 2014, ex. 11, pág. 416)** Suponha que uma amostra aleatória de tamanho n seja retirada de uma distribuição exponencial para a qual o valor do parâmetro θ é desconhecido, a distribuição a priori de θ é uma distribuição gama especificada, e o valor de θ deve ser estimado usando a função de perda de erro quadrático. Mostre que os estimadores de Bayes, para $n = 1, 2, \ldots$, formam uma sequência consistente de estimadores de θ.

Cap. 7.7 (Estatística Suficiente)

Nos dois primeiros exercícios a seguir, suponha que as variáveis aleatórias $X_1, X_2, \ldots, X_n$ formem uma amostra aleatória de tamanho n da distribuição especificada no exercício e mostre que a estatística T especificada é uma estatística suficiente para o parâmetro.

10. **(DeGroot e Schervish 2014, ex. 4, pág. 448)** A distribuição normal, na qual a média μ é conhecida e a variância $\sigma^2 > 0$ é desconhecida; $T = \sum_{i=1}^{n} (X_i - \mu)^2$.

11. **(DeGroot e Schervish 2014, ex. 7, pág. 448)** A distribuição beta com parâmetros α e β, onde o valor de β é conhecido e o valor de α é desconhecido ($\alpha > 0$); $T = \prod_{i=1}^{n} X_i$.

12. **(DeGroot e Schervish 2014, ex. 16, pág. 448)** Seja θ em um espaço de parâmetros Ω, igual a um intervalo de números reais (possivelmente não limitado). Deixe X ter uma f.d.p. ou f.m.p. $f_n(x|\theta)$ condicional a θ. Seja $T = r(X)$ uma estatística. Assuma que T é suficiente. Prove que, para cada possível f.d.p. a priori para θ, a f.d.p. a posteriori de θ dada $X = x$ depende de x apenas através de $r(x)$.

Cap. 7.8 (Estatística Suficiente Mínima)

13. **(DeGroot e Schervish 2014, ex. 8, pág. 455)** Suponha que $X_1, X_2, \ldots, X_n$ formem uma amostra aleatória de uma distribuição exponencial para a qual o valor do parâmetro β é desconhecido ($\beta > 0$). O MLE (Estimador de Máxima Verossimilhança) de β é uma estatística minimamente suficiente?

14. **(DeGroot e Schervish 2014, ex. 12, pág. 455)** Suponha que $X_1, X_2, \ldots, X_n$ formem uma amostra aleatória de uma distribuição para a qual a f.d.p. é a seguinte:

$$f(x|\theta) = \begin{cases} \frac{2x}{\theta^2} & \text{para } 0 \leq x \leq \theta, \\ 0 & \text{caso contrário.} \end{cases}$$

Aqui, o valor do parâmetro θ é desconhecido ($\theta > 0$). Determine o M.L.E. (Estimador de Máxima Verossimilhança) da mediana dessa distribuição e mostre que esse estimador é uma estatística minimamente suficiente para θ.

15. **(DeGroot e Schervish 2014, ex. 16, pág. 455)** Suponha que $X_1, X_2, \ldots, X_n$ formem uma amostra aleatória de uma distribuição de Poisson para a qual o valor da média λ é desconhecido, e que a distribuição a prior de λ seja uma determinada distribuição gama especificada. O estimador de Bayes de λ em relação à função de perda de erro quadrático é uma estatística minimamente suficiente?

Cap. 7.9 (Custos e Riscos de Estimadores)

16. **(DeGroot e Schervish 2014, ex. 2, pág. 460)** Suponha que as variáveis aleatórias $X_1, X_2, \ldots, X_n$ formem uma amostra aleatória de tamanho n ($n \geq 2$) da distribuição uniforme no intervalo $[0, \theta]$, onde o valor do parâmetro θ é desconhecido ($\theta > 0$) e deve ser estimado. Suponha também que, para todo estimador $\delta(X_1, X_2, \ldots, X_n)$, o M.S.E. $R(\theta, \delta)$ seja definido como $R(\theta, \delta) = \mathbb{E}_\theta[(\delta(\boldsymbol{X}) - h(\theta))^2]$. Explique por que o estimador $\delta_1(X_1, X_2, \ldots, X_n) = 2X_n$ é inadmissível.

17. **(DeGroot e Schervish 2014, ex. 3, pág. 460)** Considere novamente as condições do exercício anterior e seja o estimador δ_1 definido anteriormente. Determine o valor do M.S.E. $R(\theta, \delta_1)$ para $\theta > 0$.

18. **(DeGroot e Schervish 2014, ex. 6, pág. 460)** Suponha que $X_1, X_2, \ldots, X_n$ formem uma amostra aleatória de tamanho n ($n \geq 2$) da distribuição gama com parâmetros α e β, onde o valor de α é desconhecido ($\alpha > 0$) e o valor de β é conhecido. Explique por que $\bar{X}_n$ é um estimador inadmissível da média dessa distribuição quando a função de perda de erro quadrático é usada.

19. **(DeGroot e Schervish 2014, ex. 10, pág. 461)** Suponha que $X_1, X_2, \ldots, X_n$ formem uma amostra aleatória de uma distribuição para a qual a f.d.p. ou a f.m.p. é $f(x|\theta)$, onde $\theta \in \Omega$. Suponha que o valor de θ deve ser estimado, e que T é uma estatística suficiente para θ. Seja δ um estimador arbitrário de θ, e δ_0 outro estimador definido pela relação $\delta_0 = \mathbb{E}(\delta|T)$. Mostre que para cada valor de $\theta \in \Omega$,

$$\mathbb{E}_\theta(|\delta_0 - \theta|) \leq \mathbb{E}_\theta(|\delta - \theta|).$$

Cap. 8.7 (Estimadores Não-viesados)

20. **(DeGroot e Schervish 2014, ex. 4, pág. 512)** Suponha que uma variável aleatória X tenha a distribuição geométrica (com suporte em $\{0, 1, 2, \cdots\}$) com parâmetro desconhecido p. Encontre uma estatística $\delta(X)$ que seja um estimador não-viesado de $1/p$.

21. **(DeGroot e Schervish 2014, ex. 11, pág. 513)** Suponha que um determinado medicamento deve ser administrado a dois tipos diferentes de animais, A e B. Sabe-se que a resposta média dos animais do tipo A é a mesma que a resposta média dos animais do tipo B, mas o valor comum θ dessa média é desconhecido e deve ser estimado. Também é sabido que a variância da resposta dos animais do tipo A é quatro vezes maior do que a variância da resposta dos animais do tipo B. Sejam $X_1, X_2, \ldots, X_m$ as respostas de uma amostra aleatória de m animais do tipo A, e $Y_1, Y_2, \ldots, Y_n$ as respostas de uma amostra aleatória independente de n animais do tipo B. Finalmente, considere o estimador $\hat{\theta} = \alpha \bar{X}_m + (1 - \alpha)\bar{Y}_n$.

    (a) Para quais valores de α, m e n $\hat{\theta}$ é um estimador não-viesado de θ?

    (b) Para valores fixos de m e n, qual valor de α gera um estimador não-viesado com variância mínima?

22. **(DeGroot e Schervish 2014, ex. 13, pág. 513)** Suponha que $X_1, X_2, \ldots, X_n$ formem uma amostra aleatória de uma distribuição para a qual a f.d.p. ou a f.m.p. é $f(x|\theta)$, onde o valor do parâmetro θ é desconhecido. Sejam $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ e T uma estatística. Suponha que $\delta(\boldsymbol{X})$ seja um estimador não tendencioso de θ tal que $\mathbb{E}_\theta[\delta(\boldsymbol{X})|T]$ não depende de θ. (Se T for uma estatística suficiente, como definido na Seção 7.7, isso será verdade para qualquer estimador δ. Essa condição também se aplica a outros exemplos.) Seja $\delta_0(T)$ a média condicional de $\delta(\boldsymbol{X})$ dado T.

    (a) Mostre que $\delta_0(T)$ também é um estimador não-viesado de θ.

    (b) Mostre que $\text{Var}_\theta(\delta_0) \leq \text{Var}_\theta(\delta)$, para todos os valores possíveis de θ.

Caps. 8.8 (Informação de Fisher e Estimadores Eficientes)

23. **(DeGroot e Schervish 2014, ex. 7, pág. 527)** Suponha que $X_1, X_2, \ldots, X_n$ formem uma amostra aleatória da distribuição de Bernoulli com parâmetro desconhecido p. Mostre que $\bar{X}_n$ é um estimador eficiente de p.

24. **(DeGroot e Schervish 2014, ex. 10, pág. 527)** Suponha que $X_1, X_2, \ldots, X_n$ formem uma amostra aleatória da distribuição normal com média 0 e desvio padrão desconhecido $\sigma > 0$. Encontre o limite inferior especificado pela desigualdade da informação de Fisher para a variância de qualquer estimador não tendencioso de $\log \sigma$.

Referências

DeGroot, Morris H. e Mark J. Schervish (2014). *Probability and Statistics*. 4ª. Pearson Education Limited.