

Universidade de São Paulo
Escola de Artes, Ciências e Humanidades

ACH2053 – Introdução à Estatística – 1º sem. 2023

Professor: José Ricardo G. Mendonça

Gabarito — Prova de recuperação — Data: 11 ago. 2023 — 19h00 às 20h45

Na resolução dos problemas, explique seu raciocínio e o que você está fazendo de forma que eu possa acompanhá-lo(a). Soluções “mágicas” ou “geniais” não serão aceitas sem explicações.

Problemas

1. Um lote de 100 rebimbocas contém 9 peças defeituosas. Duas rebimbocas são selecionadas do lote ao acaso, sem reposição.

- (a) Qual é a probabilidade da segunda rebimboca ser defeituosa dado que a primeira rebimboca era defeituosa?

Após a retirada da primeira rebimboca defeituosa, restam 99 rebimbocas, das quais 8 são defeituosas. Assim, a probabilidade da segunda rebimboca retirada (B) ser defeituosa dado que a primeira rebimboca era defeituosa (A) vale

$$P(B|A) = \frac{8}{99} \simeq 8,1\%.$$

- (b) Qual é a probabilidade das duas rebimbocas serem defeituosas?

A probabilidade das duas rebimbocas serem defeituosas ($A \cap B$) vale

$$P(A \cap B) = P(B|A)P(A) = \frac{8}{99} \cdot \frac{9}{100} \simeq 0,73\%.$$

2. Uma das maneiras de localizar um *bug* em um programa de computador envolve executar o programa com conjuntos de entradas aleatórias até que o *bug* se manifeste. Determinada empresa estimou que a probabilidade de revelar um *bug* de um subsistema em particular dessa forma é de 12%. Encontre o valor esperado μ aproximado do número de testes necessários até encontrar um *bug* e seu desvio padrão. Dica: Use a distribuição geométrica.

A probabilidade do *bug* se manifestar no primeiro teste vale 12%, de se manifestar no segundo teste vale $(100\% - 12\%) \cdot 12\%$ e assim sucessivamente, da maneira que a probabilidade do *bug* se manifestar no k -ésimo teste vale $P(k) = 0,88^{k-1} \cdot 0,12$, uma

distribuição geométrica de parâmetro $p = 0,12$. O valor esperado de uma variável aleatória $X \sim \text{Geom}(p)$ vale $E(X) = 1/p$, no nosso caso, $E(X) \simeq 8,3$. Já o desvio padrão de X vale $\sigma = \sqrt{\text{Var}(X)} = \sqrt{(1-p)/p^2}$, no nosso caso $\sigma \simeq 7,8$. Assim, devemos esperar ter de realizar $8,3 \pm 7,8$ testes até que o *bug* se manifeste, isto é, algo entre 1 e 16 testes.

3. Seja $X \sim \text{Poisson}(\lambda)$. Calcule $E(X)$.

O valor esperado de uma variável aleatória $X \sim \text{Poisson}(\lambda)$ é dado por

$$E(X) = \sum_{k=0}^{\infty} k \cdot P(X = k) = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!},$$

onde na última igualdade fatoramos a constante $e^{-\lambda}$, cancelamos o termo com $k = 0$ e simplificamos a razão $k/k!$. O somatório restante pode ser escrito como

$$\sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \sum_{k=0}^{\infty} \frac{\lambda^{k+1}}{k!} = \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda e^{\lambda},$$

onde na última passagem reconhecemos a série de Taylor para a função e^{λ} . Assim, o valor esperado de uma variável aleatória $X \sim \text{Poisson}(\lambda)$ vale $E(X) = e^{-\lambda} \lambda e^{\lambda} = \lambda$.

4. A capacidade máxima de um elevador é de 480 kg. Se as massas (“pesos”) dos passageiros for dada por uma distribuição normal $N(75, 120)$, qual é a probabilidade de 7 passageiros ultrapassarem o limite de carga do elevador?

Sete passageiros, cada um com distribuição de massa $N(75, 120)$, possuem distribuição de massa total $X = X_1 + \dots + X_7 \sim N(7 \cdot 75, 7 \cdot 120) = N(525, 840)$. Essa é uma propriedade da distribuição normal: uma soma de variáveis aleatórias normais é novamente normal. Queremos estimar $P(X > 480)$, que podemos calcular centralizando e normalizando a variável aleatória X ,

$$P(X > 480) = P\left(\frac{X - 525}{\sqrt{840}} > \frac{480 - 525}{\sqrt{840}}\right) \simeq P(Z > -1,55) = P(Z < 1,55).$$

Consultando uma tabela encontramos $P(Z < 1,55) \simeq 94\%$, uma probabilidade razoável.

5. Seja X_1, \dots, X_n uma amostra aleatória simples de uma v.a. $X \sim f(x; \theta) = \frac{1}{2\theta^3} x^2 e^{-x/\theta}$ de parâmetro $0 < \theta < \infty$ no intervalo $0 < x < \infty$.

- (a) Encontre o estimador de máxima verossimilhança $\hat{\theta}$ de θ .

A função de máxima verossimilhança para um conjunto de dados x_1, \dots, x_n vale

$$L(\theta; x_1, \dots, x_n) = \prod_{i=1}^n f(x_i; \theta) = \prod_{i=1}^n \frac{1}{2\theta^3} x_i^2 e^{-x_i/\theta} = \frac{1}{2^n \theta^{3n}} \left(\prod_{i=1}^n x_i^2 \right) e^{-\sum_{i=1}^n x_i/\theta}.$$

Encontramos o estimador de máxima verossimilhança $\hat{\theta}$ para θ quando

$$\left. \frac{dL(\theta; x_1, \dots, x_n)}{d\theta} \right|_{\theta=\hat{\theta}} = 0.$$

Antes de calcular a derivada acima, vamos reescrever $L(\theta; x_1, \dots, x_n)$ como

$$L(\theta; x_1, \dots, x_n) = \underbrace{\frac{1}{2^n} \left(\prod_{i=1}^n x_i^2 \right)}_{\text{independe de } \theta} \theta^{-3n} \exp \left(-\frac{1}{\theta} \sum_{i=1}^n x_i \right),$$

onde destacamos a parte independente de θ com a qual não precisamos nos preocupar ao efetuar a derivação em relação a θ e que vamos simplesmente denotar por C . Dessa forma obtemos

$$\left. \frac{dL(\theta; x_1, \dots, x_n)}{d\theta} \right|_{\theta=\hat{\theta}} = C \left[-3n\hat{\theta}^{-3n-1} + \hat{\theta}^{-3n} \frac{1}{\hat{\theta}^2} \sum_{i=1}^n x_i \right] \exp \left(-\frac{1}{\hat{\theta}} \sum_{i=1}^n x_i \right) = 0,$$

isto é,

$$-3n + \frac{1}{\hat{\theta}} \sum_{i=1}^n x_i = 0 \Rightarrow \hat{\theta} = \frac{1}{3n} \sum_{i=1}^n x_i = \frac{1}{3} \bar{x}.$$

O estimador $\hat{\theta}$ de máxima verossimilhança para o parâmetro θ da distribuição $f(x; \theta)$ vale, portanto, dado um conjunto de dados x_1, \dots, x_n , $1/3$ do valor médio dos dados.

(b) Encontre o estimador de momentos $\hat{\theta}$ de θ .

Como temos somente um parâmetro para estimar, basta resolver $\mu_1 = m_1$, onde

$$\mu_1 = E(X) = \int_{-\infty}^{\infty} x f(x; \theta) dx$$

é o primeiro momento da distribuição de probabilidades e

$$m_1 = \frac{1}{n} \sum_{i=1}^n x_i$$

é o primeiro momento empírico obtido a partir dos dados. O cálculo de μ_1 fornece

$$\mu_1 = E(X) = \int_0^{\infty} x \frac{1}{2\theta^3} x^2 e^{-x/\theta} dx = \frac{1}{2\theta^3} \int_0^{\infty} x^3 e^{-x/\theta} dx.$$

Substituindo $u = x/\theta$ e integrando o resultado sucessivamente por partes obtemos

$$\mu_1 = E(X) = \frac{\theta}{2} \int_0^{\infty} u^3 e^{-u} du = \frac{\theta}{2} \cdot 3 \int_0^{\infty} u^2 e^{-u} du = \frac{\theta}{2} \cdot 6 \int_0^{\infty} u e^{-u} du = 3\theta.$$

Assim, o estimador de momentos para o parâmetro θ da pdf $f(x; \theta)$ é dado por

$$\mu_1 = m_1 \Rightarrow 3\hat{\theta} = \frac{1}{n} \sum_{i=1}^n x_i \Rightarrow \hat{\theta} = \frac{1}{3} \bar{x},$$

exatamente como antes.

Formulário

Axiomas da teoria das probabilidades

Para quaisquer eventos (subconjuntos) A e B de um espaço amostral Ω valem $P(A) \geq 0$, $P(\Omega) = 1$ e $P(A \cup B) = P(A) + P(B)$ se $A \cap B = \emptyset$.

Como consequências dos axiomas da teoria das probabilidade valem $P(\bar{A}) = 1 - P(A)$, $P(\emptyset) = 0$ e $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ para quaisquer eventos $A, B \subseteq \Omega$.

Identidades de Bayes

$$P(A|B)P(B) = P(B|A)P(A), \quad P(B) = \sum_{i=1}^n P(B|A_i)P(A_i),$$

onde $A_i \cap A_j = \emptyset$ se $i \neq j$ e $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$.

Algumas distribuições de probabilidade

- Binomial $\text{Bin}(n, p)$: $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, \dots, n$;
 $E(X) = np$, $\text{Var}(X) = np(1-p)$.
- Geométrica $\text{Geom}(p)$: $p_X(k) = (1-p)^{k-1} p$, $k = 1, 2, \dots$; $E(X) = \frac{1}{p}$, $\text{Var}(X) = \frac{1-p}{p^2}$.
- Poisson (λ) : $p_X(k) = \frac{\lambda^k e^{-\lambda}}{k!}$, $k = 0, 1, \dots$; $E(X) = \lambda$, $\text{Var}(X) = \lambda$.
- Uniforme $U(a, b)$: $f_X(x) = \frac{1}{b-a}$, $a \leq x \leq b$; $E(X) = \frac{1}{2}(a+b)$, $\text{Var}(X) = \frac{1}{12}(b-a)^2$.
Às vezes, por conveniência o domínio da distribuição uniforme é dado como $a < x < b$.
- Exponencial $\text{Exp}(\beta)$: $f_X(x) = \frac{1}{\beta} e^{-x/\beta}$, $x \geq 0$; $E(X) = \beta$, $\text{Var}(X) = \beta^2$.
- Normal $N(\mu, \sigma^2)$: $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$, $-\infty < x < +\infty$;
 $E(X) = \mu$, $\text{Var}(X) = \sigma^2$.

Determinação de intervalo de confiança e do tamanho de uma amostra

Para que $P(|\bar{X} - \mu| \leq \varepsilon) \geq \gamma$, onde $\varepsilon > 0$ é o erro amostral tolerado e $0 < \gamma < 1$ é o coeficiente de confiança, usando o CLT para \bar{X} encontramos que devemos ter $P(-z_\gamma \leq Z \leq z_\gamma) \simeq \gamma$, onde $Z \sim N(0, 1)$ e $z_\gamma = \sqrt{n}\varepsilon/\sigma$ com σ^2 a variância da população a partir da qual \bar{X} foi obtido, de onde podemos extrair o intervalo de confiança $\text{IC}(\mu; \gamma) = (\bar{X} - \varepsilon, \bar{X} + \varepsilon)$ de nível γ para a média da população ou o valor do tamanho necessário n da amostra. No caso de proporções, a população $X \sim \text{Bernoulli}(p)$ e sua variância vale $\sigma^2 = p(1-p)$. Caso não conheçamos o valor de p podemos maximizar o valor de σ^2 na fórmula para z_γ ou n por $1/4$.

Estimadores de momentos, mínimos quadrados (EMQ) e máxima verossimilhança (EMV)

- Dado um conjunto de valores x_1, \dots, x_n extraídos de uma população com distribuição $f(x; \theta)$, o método dos momentos consiste em determinar os parâmetros $\theta = (\theta_1, \dots, \theta_k)$ resolvendo as k equações $\mu_k = m_k$ envolvendo os momentos $\mu_k = E(X^k)$ da distribuição e os momentos empíricos $m_k = \frac{1}{n} \sum_{i=1}^n x_i^k$ obtidos a partir dos dados.
- Dado um conjunto de dados $(x_1, y_1), \dots, (x_n, y_n)$ e um modelo $Y = g(X; \theta) + \varepsilon$ com $E(\varepsilon) = 0$, o EMQ dos parâmetros θ é aquele que minimiza $S(\theta) = \sum_{i=1}^n [Y_i - g(X_i; \theta)]^2$.
- Dado um conjunto de valores x_1, \dots, x_n extraídos de uma população com distribuição $f(x; \theta)$, o EMV dos parâmetros θ é aquele que minimiza

$$L(\theta; x_1, \dots, x_n) = \prod_{i=1}^n f(x_i; \theta) \text{ ou, equivalentemente, } \ell(\theta; x_1, \dots, x_n) = \sum_{i=1}^n \log f(x_i; \theta).$$

Transformação entre distribuições normais

Se $F_X(x)$ é a cdf de uma v.a. $X \sim N(\mu, \sigma^2)$ e $\Phi(z)$ é a cdf de uma v.a. padrão $Z \sim N(0, 1)$, então

$$F_X(x) = \Phi\left(\frac{x - \mu}{\sigma}\right), \quad \Phi(-z) = 1 - \Phi(z).$$

Alguns valores de $\Phi(z) = P(Z < z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{1}{2}u^2\right) du$ aparecem na tabela abaixo.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998

Tabela: Valores da distribuição cumulativa normal padrão $\Phi(z) = P(Z < z)$ para $z \geq 0$.