UFRJ - CCMN - IM - Departamento de Métodos Estatísticos
Probabilidade e Estatística - Estatística

Prova Final 29-05-2014

Atenção: Não serão aceitas respostas sem justificativa: as expressões que levaram a alguma resposta numérica devem ser indicadas nos espaços apropriados.

---

1. Uma determinada mistura de um processo químico demora no mínimo um minuto para começar a reagir e depois disto o tempo (em minutos) até o processo de mistura ser concluído pode ser modelado por uma variável aleatória X com função de densidade de probabilidade dada por $f(x) = 3/x^4$ para $x \geqslant 1$ e $f(x) = 0$ para $x < 1$.

   (a) Calcule a variância do tempo até o processo de mistura ser concluído.

   (b) Calcule a probabilidade do processo demorar menos do que um minuto e meio para ser concluído.

   Solução: **(a)**

$$\begin{aligned} \mathrm{E}(X) &= \int_1^\infty x \times \frac{3}{x^4}dx = 3\int_1^\infty \frac{1}{x^3}dx = -\frac{3}{2x^2}\Big|_1^\infty = \frac{3}{2}. \\ \mathrm{E}(X^2) &= \int_1^\infty x^2 \times \frac{3}{x^4}dx = 3\int_1^\infty \frac{1}{x^2}dx = -\frac{3}{x}\Big|_1^\infty = 3. \\ \mathrm{Var}(X) &= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = 3 - \left(\frac{3}{2}\right)^2 = \frac{3}{4} = 0,75. \end{aligned}$$

   (b)

$$\begin{aligned} \Pr(X < 3/2) &= \Pr(1 \leqslant X < 3/2) = \int_1^{3/2} \frac{3}{x^4}dx \\ &= -\frac{1}{x^3}\Big|_1^{3/2} = 1 - \frac{1}{(3/2)^3} = 1 - \frac{8}{27} = \frac{19}{27} = 0,7037. \end{aligned}$$

2. Sabemos que a esperança e a variância de uma v.a. podem ser encaradas, respectivamente, como a abscissa do centro de gravidade e o momento de inércia de uma distribuição de massa, e que isso vale tanto no caso discreto como no caso contínuo. Admita que a função de densidade da v.a. contínua X é dada por

$$f(x) = \begin{cases} 0 & \text{se} \quad x \leq -3 \\ \frac{3}{5} + \frac{x}{5} & \text{se} \quad -3 < x \leq -2 \\ \frac{1}{5} & \text{se} \quad -2 < x \leq 2 \\ \frac{3}{5} - \frac{x}{5} & \text{se} \quad 2 < x \leq 3 \\ 0 & \text{se} \quad x > 3 \end{cases}$$

   (a) Determine $E(X)$ e $Var(X)$. Dica: Fazer um gráfico da função.

   (b) Admita agora que Y é uma v.a. discreta que só pode assumir três valores: $-2, 0$ e 2. Obtenha a função de probabilidade da v.a. Y, sabendo que $E(X) = E(Y)$ e $Var(X) = Var(Y)$. Isso significa que ambas as distribuições de massa têm o mesmo centro de gravidade e o mesmo momento de inércia, embora uma seja contínua e a outra seja discreta.

   Solução: **(a)** Devido à simetria da função em relação ao eixo vertical, conclui-se que $E(X) = 0$. Por outro lado, $Var(X) = E(X^2) - (EX)^2 = E(X^2)$, ou seja

$$Var(X) = \int_{\mathbb{R}} x^2 f(x)dx = \int_{-3}^3 x^2 f(x)dx = 2 \times \int_2^3 (3/5 - x/5)x^2dx + \int_{-2}^2 x^2/5dx.$$

   Obtemos $\int_2^3 (3/5 - x/5)x^2dx = 11/20$ e por outro lado $\int_{-2}^2 x^2/5dx = 16/15$. Portanto $Var(X) = 13/6$.

(b) Y pode assumir valores $-2, 0$ e 2 portanto, $EY = -2P(Y=-2) + 2P(Y=2)$. Pelo enunciado, temos $EY = 0$. Portanto $P(Y=2) = P(Y=-2)$. Use agora que $Var(Y) = E(Y^2) - (EY)^2 = 4P(Y=-2) + 4P(Y=2) - 0$. Pelo enunciado, temos $Var(Y) = 13/6$ e concluimos que $P(Y=2) = P(Y=-2) = 13/48$. Segue que $P(Y=0) = 1 - 2 \times 13/48 = 11/24$.

3. Num estudo sobre o peso dos brasileiros, foram pesadas 1000 pessoas escolhidas de forma aleatória. Os resultados desta amostra podem ser denotados por $x_1, x_2, \ldots, x_{1000}$.

   Seja X a variável aleatória representando o peso de um brasileiro selecionado ao acaso. Admita que são conhecidos $\mathrm{E}X = \mu$, $\mathrm{Var}X = \sigma^2$ e $P(X > 100) = p$ a probabilidade de que um brasileiro pese mais que 100 kg.

   (a) Sugerir um estimador para o peso médio dos brasileiros, e um estimador para a proporção de pessoas pesando mais que 100 kg.

   (b) Expresse a média e a variância destes estimadores, em função de μ, σ^2 e p.

   (c) $\bar{X}$ teria a mesma distribuição se estivéssemos amostrando 1000 pessoas sem reposição dentro de uma cidade cuja população é de 2 000 pessoas para estudar o comportamento do peso nesta cidade? (Justificar).

   Solução: (a) $\bar{X} = \frac{1}{1000}\sum_{i=1}^{1000} X_i$ e $\hat{p} := \frac{1}{1000}\sum_{i=1}^{1000} Y_i$ onde $Y_i = 1$ se $X_i > 100$ e $Y_i = 0$ caso contrário.
   (b) Pelo TCL, $\bar{X}$ tem aproximadamente distribuição $N(\mu, \frac{\sigma^2}{1000})$ e $\hat{p}$ tem aproximadamente distribuição $N(p, \frac{p(1-p)}{1000})$ onde $p = P(X > 100)$.
   (c) Não, pois pelo fato da população ser finita, a amostra não sería mais uma amostra aleatória (ou seja, as X_i's não serão mais independentes).

4. Numa industria foi colhida uma amostra de 40 peças cilíndricas , as quais depois de examinadas apresentaram os seguintes diâmetros (em cm).

1,1	1,1	1,1	1,1	1,1	1,1	1,2	1,2	1,2	1,2	1,2	1,2	1,2	1,3	1,3	1,3	1,3	1,3	1,3	1,3
1,3	1,3	1,3	1,3	1,3	1,3	1,4	1,4	1,4	1,4	1,4	1,4	1,4	1,5	1,5	1,5	1,5	1,6	1,6	1,6

   Informamos também que $\sum_{i=1}^{40} x_i = 52,5$; $\sum_{i=1}^{40} x_i^2 = 69,71$.

   (a) Construir um intervalo de confiança (IC) para o diâmetro médio populacional ao nível de 95%.

   (b) Construir um IC não conservativo para a proporção populacional de peças que têm diâmetros entre 1,2 e 1,4 (incluindo os extremos do intervalo) ao nível de confiança de 90%.

   (c) Obtenha o p-valor num teste com $\mathrm{H}_0 : \mu \leq 1,26$ versus $\mathrm{H}_1 : \mu > 1,26$. Qual decisão deve ser tomada se o nível de significância é 2%?

   Solução: (a) $\overline{x} = \frac{52,5}{40} = 1,3125$; $\quad s^2 = \frac{1}{39}(69,71 - \frac{52,5^2}{40}) = 0,0206$ e portanto $s = \sqrt{0,0206} = 0,144$.
   I.C. para μ com 95%. Como n=40 é um tamanho de amostra amostra grande, podemos usar a distribuição Normal em vez da t de Student. Assim,
   $d = z_{1-\alpha/2}\frac{s}{\sqrt{n}} = 1,96\ \frac{0,144}{\sqrt{40}} = 0,0445$
   $l_{inf} = \overline{x} - d = 1,3125 - 0,0445 = 1,268$
   $l_{sup} = \overline{x} + d = 1,3125 + 0,0445 = 1,357$
   (b) $\hat{p} = \frac{27}{40} = 0,675$; $\quad z_{1-\alpha/2} = 1,64$;
   I.C., não conservativo, para p com 90%: $\quad d = 1,64\sqrt{\frac{0,675 \times 0,325}{40}} = 0,0741$
   $l_{inf} = \hat{p} - d = 0,675 - 0,0741 = 0,5535$
   $l_{sup} = \hat{p} + d = 0,675 + 0,0741 = 0,7965$
   (c) Sendo a hipótese H_0 considerada como simples, o teste fica
   H_0: $\mu = 1,26$ versus H_1: $\mu > 1,26$
   Assim, o p-valor $= 1 - \Phi(\frac{\overline{x}-\mu_0}{s/\sqrt{n}}) = 1 - \Phi(\frac{1,3125-1,26}{0,144/\sqrt{40}}) = 1 - \Phi(2,31) = 0,0104$
   Com o nível de significância $\alpha = 2\%$, os dados trazem evidência pela rejeição da hipótese H_0.

Boa prova!