Estatística (Curso de Verão) - Prova

Prof. Regis Augusto Ely

Mestrado em Economia Aplicada (PPGOM)

Universidade Federal de Pelotas (UFPel)

13 de março de 2018

1) (2.0 pontos) Para os experimentos a seguir, descreva matematicamente o espaço amostral:

a) Jogar uma moeda justa quatro vezes e observar os resultados.

b) Observar a taxa de alfabetização de uma população com n indivíduos.

c) Medir o tempo de duração de uma bateria.

d) Contar o número de carros por domicílio em uma população.

2) (2.0 pontos) Prove as seguintes afirmações (assumindo que todos os eventos condicionantes tem probabilidades positivas):

a) Se $P(B) = 1$, então $P(A|B) = P(A)$ para qualquer A.

b) Se $A \subset B$, então $P(B|A) = 1$ e $P(A|B) = P(A)/P(B)$.

c) Se A e B são mutuamente exclusivos (disjuntos), então $P(A|A \cup B) = \frac{P(A)}{P(A)+P(B)}$.

d) $P(A \cap B \cap C) = P(A|B \cap C)P(B|C)P(C)$.

3) (1.0 ponto) Um teste de tuberculose tem a seguinte confiabilidade: se uma pessoa tem tuberculose, o teste dá positivo 98% das vezes; se uma pessoa não tem tuberculose, o teste dá negativo 99% das vezes. Em uma população em que 2 de cada 10.000 pessoas tem tuberculose, uma pessoa é selecionada aleatoriamente para fazer o teste, que dá positivo. Qual a probabilidade aproximada dessa pessoa ter tuberculose de fato?

4) (2.0 pontos) Considere a seguinte função densidade de probabilidade conjunta de duas variáveis aleatórias X e Y:

$$f(x,y) = \begin{cases} 24xy & \text{se } 0 < x < 1, 0 < y < 1 \text{ e } x + y < 1 \\ 0 & \text{caso contrário} \end{cases}$$

a) Encontre $P(X + Y < 1/2)$.

b) Ache a função de distribuição marginal de X.

c) Ache a função de distribuição marginal de Y. Esperança EX^2

d) Ache a função de distribuição de X condicional à $Y = 1/2$.

5) (2.0 pontos) Considere a função densidade de probabilidade conjunta de duas variáveis aleatórias X e Y dada por:

$$f(x,y) = \begin{cases} xy/96 & \text{se } 0 < x < 4 \text{ e } 0 < y < 5 \\ 0 & \text{caso contrário} \end{cases}$$

a) Encontre $E(X)$.

b) Encontre $E(Y)$.

c) Encontre $Cov(X, Y)$.

d) As variáveis X e Y são independentes? Comprove sua reposta.

6) (1.0 ponto) Sejam X uma variável aleatória não negativa, e $X_1, X_2, \ldots$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas com $EX_i = \mu$ e $VarX_i = \sigma^2 < \infty$. Defina $\bar{X} = (1/n)\sum_{i=1}^{n} X_i$ e $\alpha > 0$. Então:

a) Prove que $P(X \geq \alpha) \leq \frac{1}{\alpha}E(X)$.

b) Prove que $\bar{X}_n$ converge em probabilidade para μ.

Estatística 2018 / Gabarito

1 a) O espaço amostral é um conjunto matemático, que neste caso, conterá $2^4 = 16$ elementos. Ele é descrito por:

$$S = \{(H,H,H,H), (H,H,H,T), (H,H,T,H), (H,T,H,H), (T,H,H,H), (H,H,T,T), (H,T,H,T), (H,T,T,H), (T,T,H,H), (T,H,T,H), (T,H,H,T), (H,T,T,T), (T,H,T,T), (T,T,H,T), (T,T,T,H), (T,T,T,T)\},$$

onde H representa Cara e T Coroa.

b) A taxa de alfabetização é o número de indivíduos alfabetizados dividido pelo número total de indivíduos, no caso, n. Assim,

$S = \{0, 1/n, 2/n, \ldots, 1\}$, sendo um conjunto enumerável.

c) Tempo de duração é um número contínuo e positivo. Assim,

$S = [0, \infty)$.

d) Novamente um conjunto enumerável, mas que não tem limite superior. Assim,

$S = \{0, 1, 2, \ldots\}$

(2) a) Pelo probabilidade condicional, $P(A/B) = \frac{P(A \cap B)}{P(B)}$. Como $P(B) = 1$, então $P(A \cap B) = P(A)$, logo

$$P(A/B) = \frac{P(A)}{1} = P(A).$$

//

b) Novamente, pelo probabilidade condicional,

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad e \quad P(B/A) = \frac{P(A \cap B)}{P(A)}.$$

Como $A \subset B$, então $A \cap B = A$. Assim,

$$P(B/A) = \frac{P(A)}{P(A)} = 1 \quad ; \ e$$

$$P(A/B) = \frac{P(A)}{P(B)}.$$

//

c) Se A e B são disjuntos então $A \cap B = \emptyset$. Pela probabilidade condicional,

$$P(A/A \cup B) = \frac{P(A \cap (A \cup B))}{P(A \cup B)}.$$

Pela propriedade distributiva dos operadores de união e interseção, podemos escrever

$$A \cap (A \cup B) = \underbrace{(A \cap A)}_{A} \cup \underbrace{(A \cap B)}_{\emptyset} = A$$

↳ pois A e B são disjuntos

Assim, $P(A / A \cup B) = \dfrac{P(A)}{P(A \cup B)}$.

Mas como A e B são disjuntos
$P(A \cup B) = P(A) + P(B) - \underbrace{P(A \cap B)}_{0} = P(A) + P(B)$

De modo que $P(A / A \cup B) = \dfrac{P(A)}{P(A) + P(B)}$

//

d) Pela probabilidade condicional, podemos escrever
$P(A \cap (B \cap C)) = P(A / B \cap C) \cdot P(B \cap C)$.

Da mesma forma, $P(B \cap C) = P(B / C) \cdot P(C)$

Logo, $P(A \cap B \cap C) = P(A / B \cap C) \cdot P(B / C) \cdot P(C)$

③ Trata-se de uma questão em que podemos aplicar a regra de Bayes para solucionar. Para isso devemos definir cada um dos eventos e suas respectivas probabilidades. Considere os seguintes eventos:

T => Pessoa tem tuberculose
NT => Pessoa não tem tuberculose
Pos => Teste dá positivo
Neg => Teste dá negativo

São informadas as seguintes probabilidades:

$P(Pos/T) = 0{,}98 \Rightarrow P(Neg/T) = 0{,}02$
$P(Neg/NT) = 0{,}99 \Rightarrow P(Pos/NT) = 0{,}01$
$P(T) = \dfrac{2}{10.000} = 0{,}0002 \Rightarrow P(NT) = 0{,}9998$

A pergunta é $P(T/Pos) = ?$

Pela regra de Bayes:

$$P(T/Pos) = \frac{P(Pos/T) \cdot P(T)}{P(Pos)} = \frac{0{,}98 \cdot 0{,}0002}{P(Pos)}$$

Como não temos $P(Pos)$ devemos calcular pela lei das probabilidades totais:

$$P(Pos) = P(Pos/T) \cdot P(T) + P(Pos/NT) \cdot P(NT)$$
$$= 0{,}98 \cdot 0{,}0002 + 0{,}01 \cdot 0{,}9998$$
$$= 0{,}010194$$

Assim, $P(T/Pos) = \dfrac{0{,}98 \cdot 0{,}0002}{0{,}010194} = 0{,}019227$

Ou seja, a probabilidade desta pessoa ter tuberculose de fato é de 1,92%.

4) a) Note que os condicionantes da função densidade de probabilidade são $0 < x < 1$, $0 < y < 1$ e $x + y < 1$.
Assim,

$$P(X+Y < 1/2) = \int_0^{1/2}\int_0^{1/2-x} 24xy\,dy\,dx$$
$$= \int_0^{1/2} 12x\left(\frac{1}{4} + x^2 - x\right)dx$$
$$= \int_0^{1/2} (3x + 12x^3 - 12x^2)\,dx$$
$$= \frac{3}{8} + \frac{3}{16} - \frac{1}{2} = \boxed{\frac{1}{16}}$$

b) Para achar a marginal de X devemos integrar a função densidade de distribuição conjunta em relação à ~~X~~. Porém, note o limitador $x+y<1$.
Assim,

$$f_X(x) = \int_0^{1-x} 24xy\,dy = \frac{24 \cdot x \cdot y^2}{2}\Big|_0^{1-x}$$
$$= 12x(1-x)^2$$

Assim, $f_X(x) = \begin{cases} 12x(1-x)^2 & p/\ 0 < x < 1 \\ 0 & \text{caso contrário.} \end{cases}$

c) Devido a simetria da fdp conjunta, a distribuição marginal de Y será semelhante a de X, de modo que

$$f_Y(y) = \begin{cases} 12y(1-y)^2 & p/\ 0 < y < 1 \\ 0 & \text{caso contrário} \end{cases}$$

d) A função de distribuição de X condicional a Y pode ser dada por:

$$f(x/y) = \frac{f(x,y)}{f_Y(y)}$$

Assim, $f(x/y=\frac{1}{2}) = \dfrac{f(x,\frac{1}{2})}{f_Y(\frac{1}{2})}$

Avaliando nos respectivas funções temos:

$$f(x/y=\tfrac{1}{2}) = \frac{24 \cdot x \cdot \frac{1}{2}}{12 \cdot \frac{1}{2}(1-\frac{1}{2})^2} = \frac{12x}{12 \cdot \frac{1}{2} \cdot \frac{1}{4}}$$

Logo, $f(x/y=\frac{1}{2}) = \begin{cases} 8x & p/\ 0 < x < \frac{1}{2} \\ 0 & \text{caso contrário} \end{cases}$

Note que o condicionante de X advém do fato de que $X+Y<1$ e $Y=\frac{1}{2}$.

5 a) O valor esperado de X é dado por $E(X) = \int_{-\infty}^{\infty} x \cdot f_x(x)\, dx$,

Já a distribuição marginal de X é dada por $f_x(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$.

Assim,

$$E(X) = \int_0^4 x \cdot \int_0^5 \frac{x.y}{96}\, dy\, dx = \int_0^4 x \cdot \left[\frac{x.y^2}{96.2}\right]_0^5 dx =$$

$$= \int_0^4 \frac{25x^2}{192}\, dx = \frac{25x^3}{192.3}\Big|_0^4 =$$

$$= \frac{1600}{576} = \frac{400}{144} = \boxed{2,777}$$

b) Utilizando o mesmo raciocínio do item anterior:

$$E(y) = \int_0^5 y \cdot \int_0^4 \frac{x.y}{96}\, dx\, dy = \int_0^5 y \cdot \left[\frac{x^2.y}{96.2}\right]_0^4 dy =$$

$$= \int_0^5 \frac{16.y^2}{192}\, dy = \frac{16y^3}{192.3}\Big|_0^5 =$$

$$= \frac{2000}{576} = \frac{125}{36} = \boxed{3,472}$$

6 a) Note que $E(X) = \int_{-\infty}^{\infty} x\, f(x)\,dx =$

$$= \int_{-\infty}^{a} x\, f(x)\,dx + \int_{a}^{\infty} x f(x)\,dx \geq \int_{a}^{\infty} x f(x)\,dx$$

$$\geq \int_{a}^{\infty} a\, f(x)\,dx = a \cdot P(X \geq a)$$

Logo, $P(X \geq a) = \dfrac{E(X)}{a}$. //

b) Pela definição de convergência em probabilidade, X_n convergirá em probabilidade para μ se

$$\lim_{n \to \infty} P(|\bar{X}_n - X| \geq \epsilon) = 0$$

Para verificar que isso é válido relembre que $E(\bar{X}) = \mu$ e $Var(\bar{X}) = \dfrac{\sigma^2}{n}$

Assim quando $n \to \infty$, $Var(\bar{X}) \to 0$ e de acordo com a desigualdade de Chebycheff:

$$P(|X - c| \geq \varepsilon) \leq \frac{1}{\varepsilon} \cdot Var(X),$$

de modo que no caso de $\bar{X}_n$:

$$P(|\bar{X}_n - X| \geq \epsilon) = 0 \text{ quando } n \to \infty$$

//