Lista 5: Probabilidade III

Igor Lacerda

11 de maio de 2023

Questões discursivas

1. Considere X como uma variável aleatória no espaço amostral S. Definimos a **variância** de X, indicada por $V(x)$, como

$$\boxed{V(x) = \sum_{s \in S} (X(S) - E(X))^2 \cdot p(s)}$$

   O **desvio padrão** de X, indicado por $\sigma(x)$, é definido como $\sqrt{V(x)}$.

   Note que calcular a **variância** pela definição é um processo meio chatinho. Então podemos usar o seguinte teorema:

$$\boxed{V(x) = E(X^2) - E(X)^2}$$

   De forma menos abstrata, a variância indica o "o quão bem" as probabilidades estão distribuídas. Quanto menor (maior) é a variância, mais próximas (distantes) estão as probabilidades.

2. As variáveis aleatórias X, Y do espaço amostral S são **independentes** se

$$P(X = x \wedge Y = y) = P(X = x) \cdot P(Y = y)$$

   [1] Ou, discursivamente, se a probabilidade de $X(s) = r_1$ e $Y(s) = r_2$ for igual ao produto das probabilidades de $X(s) = r_1$ e $Y(s) = r_2$, para todos os números reais r_1 e r_2.

   O legal de se ter variáveis aleatórias X, Y de um mesmo espaço amostral, é o seguinte teorema: $E(XY) = E(X) \cdot E(Y)$.

3. O **valor esperado** de uma distribuição normal corresponde ao ponto que é o *pico do gráfico*, que é chamado de μ. Seu **desvio-padrão** é σ, que representa os pontos de inflexão (quanto menor, mais próximos os valores estão).

[1] Usando a notação do prof, que é menos carregada, mas com o discurso na do livro.

4. Na distribuição normal, o “volume” de probabilidade que corresponde ao intervalo $[\mu - \sigma, \mu + \sigma]$ é $\approx 68\%$, e no intervalo $[\mu - 2\sigma, \mu + 2\sigma]$ é $\approx 95\%$.

5. A distribuição normal é uma aproximação da distribuição binomial, para n suficientemente grande.

6. Na nova abordagem do **Teorema de Bayes**, nós usamos a *chance* ao invés da probabilidade. A **chance** difere da probabilidade no seguinte sentido: enquanto a probabilidade é um número que varia entre 0 e 1 (incluindo os extremos), a chance é uma proporção do *total de possibilidades que um evento ocorre e as que não ocorre.* Nisso, simplificamos um pouco as contas do Teorema de Bayes, pois não precisamos calcular $p(E)$ (calculamos, por exemplo $\frac{p(F|E)}{p(\overline{F}|E)}$), que pode ser trabalhoso às vezes.

7. No **Silogismo do Policial**, um policial *infere* que uma certa pessoa cometeu um crime, dado que ela é suspeita e que o policial possui certa *experiência* para averiguar essa questão. Não é uma *dedução lógica* “clássica”, de fato, esse raciocínio segue de uma extensão da lógica tradicional que usa probabilidade para concluir novas informações.

8. Essa pergunta apareceu numa lista anterior. Veja a questão 10, da lista 3.

Exercícios

Dessa vez, alguns estão anexados como fotos, mas outros estão feitos aqui mesmo, usando o LaTeX.

1. Dê um exemplo de um algoritmo probabilístico não mencionado na aula.

   R: O **Algoritmo de Karger**, que faz *alguma coisa* com grafos têm uma versão probabilística.

Refazendo exercícios 6.3.4 e 6.3.10 usando o novo Bayes

6.3.4. C := Caixa, L := Laranja, P := Preta;

$$C_1 = [3L, 4P], C_2 = [5L, 6P]$$

Quero saber a **probabilidade de C_2 dado L**.

	$E = L$	$\overline{E} = P$	
$F = C_1$	3	4	$Pr(F) = \frac{1}{2}$
$\overline{F} = C_2$	5	6	$Pr(\overline{F}) = \frac{1}{2}$

Ou seja, quero saber $p(\overline{F} \mid E)$. Por Bayes:

$$\frac{p(\overline{F} \mid E)}{p(F \mid E)} = \frac{p(E \mid \overline{F}) \cdot p(\overline{F})}{p(E \mid F) \cdot p(F)}$$

$$\Rightarrow \frac{p(\overline{F} \mid E)}{p(F \mid E)} = \frac{5/11 \cdot 1/2}{3/7 \cdot 1/2}$$

$$\Rightarrow \frac{p(\overline{F} \mid E)}{p(F \mid E)} = \frac{5/11}{3/7}$$

$$\Rightarrow p(\overline{F} \mid E) = \frac{5/11}{3/7 + 5/11}$$

$$\Rightarrow p(\overline{F} \mid E) = \frac{5/11}{68/77}$$

$$\Rightarrow p(\overline{F} \mid E) \approx 0.5147$$

Uau, eu inclusive *errei* a última passagem quando fiz essa problema pela primeira vez.

6.3.10. População: 4% gripada; Teste: 97% positivos verdadeiros e 2% falsos positivos.

	F = Gripado	$\overline{F}$ = Saudável	Total
$E = +$	388	192	580
$\overline{E} = -$	12	9408	9420
Total	400	9600	10000

1. $p(F \mid E)$

$$\frac{p(\overline{F} \mid E)}{p(F \mid E)} = \frac{192/9600 \cdot 96\%}{388/400 \cdot 4\%}$$

$$\Rightarrow \frac{p(\overline{F} \mid E)}{p(F \mid E)} = \frac{192/9600 \cdot 24}{388/400}$$

$$\Rightarrow p(F \mid E) = \frac{388/400}{192/9600 \cdot 24 + 388/400}$$

$$\Rightarrow p(F \mid E) \approx 0.66$$

2. $p(\overline{F} \mid E)$

   Pelo exercício anterior:

$$\Rightarrow p(F \mid E) = \frac{24 \cdot 192/9600}{192/9600 \cdot 24 + 388/400}$$

$$\Rightarrow p(F \mid E) \approx 0.33$$

   Poderia também argumentar que é o complementar.

3. $p(F \mid \overline{E})$

$$\frac{p(F \mid \overline{E})}{p(\overline{F} \mid \overline{E})} = \frac{3\% \cdot 4\%}{9408/9600 \cdot 96\%}$$

$$\Rightarrow \frac{p(F \mid \overline{E})}{p(\overline{F} \mid \overline{E})} = \frac{3\%}{24 \cdot 9408/9600}$$

$$\Rightarrow p(F \mid \overline{E}) = \frac{3\%}{24 \cdot 9408/9600 + 3\%}$$

$$\Rightarrow p(F \mid \overline{E}) \approx 0.001$$

4. $p(\overline{F} \mid \overline{E})$

   Pelo complementar do anterior: 1 - 0.001 $\approx$ 0.999.

Correção

D. Como João escolhe 13 cartas das *51 restantes*, e há uma garantia que todas elas são do mesmo naipe (que não é o desejado), então Maria só pode ter escolhido uma entre opções de naipes, que são igualmente prováveis. Logo, a probabilidade é $\frac{1}{3}$.

E. A probabilidade de o naipe das 13 cartas que João separou é $\frac{1}{4}$, e nesse caso, é impossível Maria ter escolhido a carta do naipe desejado. Por outro lado, nos $\frac{3}{4}$ restantes (João separou outro naipe), existe $\frac{1}{3}$ de chance de a carta de Maria ser do naipe desejado. Assim, temos:

$$\frac{1}{4} \cdot 0 + \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{4}$$

Assim, em princípio, é a mesma probabilidade que no caso base. O fato de João ter separado 13 cartas *quaisquer* (no sentido que elas não agregam informação ao caso), não influencia no resultado final. Em contraste, com a **D**, em que é dada uma informação mais "concreta".