Estatística Inferencial - Semana 2
Revisão de Probabilidades

Os exercícios foram extraídos dos livros:

Bussab, WO; Morettin, PA. Estatística Básica. São Paulo: Editora Saraiva, 2006 (5ª Edição).

Magalhães, MN; Lima, ACP. Noções de Probabilidade e Estatística. São Paulo: EDUSP, 2008.

1. Sendo A e B dois eventos em um mesmo espaço amostral, "traduza" para a linguagem da teoria dos conjuntos, as seguintes situações:

   a) Pelo menos um dos eventos ocorre.
   b) O evento A ocorre mas o evento B não.
   c) Nenhum deles ocorre.
   d) Exatamente um dos eventos ocorre.

2. Uma universidade tem 10000 alunos dos quais 4000 são considerados esportistas. Temos, ainda, que 500 alunos são do curso de biologia diurno, 700 da biologia noturno, 100 são esportistas e da biologia diurno e 200 são esportistas e da biologia noturno. Um aluno é escolhido, ao acaso, e pergunta-se a probabilidade de:

   a) Ser esportista.
   b) Ser esportista e aluno da biologia noturno.
   c) Não ser da biologia.
   d) Ser esportista ou aluno da biologia.
   e) Não ser esportista, nem aluno da biologia.

3. O São Paulo Futebol Clube ganha com probabilidade 0.7 se chove e com 0.8 se não chove. Em Setembro, a probabilidade de chuva é de 0.3. Se o São Paulo ganhou uma partida em Setembro, qual a probabilidade de ter chovido nesse dia?

4. As preferências de homens e mulheres por cada gênero de filme alugado em uma locadora de vídeos, estão apresentadas na próxima tabela. Sorteando-se, ao acaso, uma dessas locações de vídeo, pergunta-se a probabilidade de:

   a) Uma mulher ter alugado um filme policial?
   b) O filme alugado ser uma comédia?
   c) Um homem ter alugado ou o filme ser um romance?
   d) O filme ser policial dado que foi alugado por um homem?

Sexo/Filme	Comédia (C)	Romance (R)	Policial (P)
Homens (H)	136	92	248
Mulheres (M)	102	195	62

5. A tabela a seguir apresenta informações de alunos de uma universidade quanto às variáveis: Período, Sexo e Opinião sobre a Reforma Agrária. Determine a probabilidade de escolhermos:

   a) Uma pessoa do sexo masculino e sem opinião sobre a reforma agrária?
   b) Uma mulher contrária a reforma agrária?

c) Dentre os estudantes do noturno, um que seja a favor da reforma agrária?
d) Uma pessoa sem opinião, sabendo-se que ela é do sexo feminino?

Período	Sexo	Reforma Agrária		
		Contra	A Favor	Sem opinião
Diurno	Feminino	2	8	2
	Masculino	8	9	8
Noturno	Feminino	4	8	2
	Masculino	12	10	1

6. Num certo bairro da cidade de São Paulo, as companhias de seguro estabeleceram o seguinte modelo para número de veículos furtados por semana:

Furtos (F)	0	1	2	3	4
p_i	1/4	1/2	1/8	1/16	1/16

Calcule a média e a variância do número de furtos semanais do bairro.

7. Verifique se as expressões a seguir são funções densidade de probabilidade (assuma que elas se anulam fora dos invtervalos especificados).
   a) $f(y) = 3y$, se $0 \leq y \leq 1$.
   b) $f(y) = y^2/2$, se $y \geq 0$.
   c) $f(y) = (y-3)/2$, se $3 \leq y \leq 5$.
   d) $f(y) = 2$, se $0 \leq y \leq 2$.
   e) $f(y) = \begin{cases} (2+y)/4, & se \ -2 \leq y < 0; \\ (2-y)/4, & se \ 0 \leq y \leq 2. \end{cases}$
8. A função de probabilidade conjunta entre as variáveis aleatórias X e Y é apresentada a seguir:

$X \setminus Y$	-2	0	2	4
-1	0.1	0.2	0.1	0.2
1	0.2	0	0.1	0.1

   a) Obtenha as funções de probabilidade marginais das variáveis.
   b) X e Y são independentes?
   c) Calcule a covariância e a correlação entre X e Y.
9. Um usuário de transporte coletivo chega pontualmente às 8 horas para pegar o seu ônibus. Devido ao trânsito caótico, a demora pode ser qualquer tempo entre 1 e 20 minutos (admita que o relógio "pule" de minuto em minuto). Pergunta-se:
   a) Qual a probabilidade de demorar mais de 10 minutos?
   b) Qual a probabilidade de demorar pelo menos 5 mas não mais de 10 minutos?
   c) Qual a probabilidade da demora não chegar a 5 minutos?
   d) Se um amigo chegou 10 minutos atrasado e vai pegar o mesmo ônibus (que ainda não passou), qual a probabilidade do amigo atrasado esperar até 3 minutos?
10. Uma certa doença pode ser curada por meio de procedimento cirúrgico em 80% dos casos. Dentre os que têm essa doença, sorteamos 15 pacientes que serão submetidos à cirurgia. Fazendo alguma suposição adicional que julgar necessária, responda qual é a probabilidade de:
    a) Todos serem curados?
    b) Pelo menos dois não serem curados?
    c) Ao menos 10 ficarem livres da doença?
11. Uma indústria de tintas recebe pedidos de seus vendedores através de whatsapp, telefone e Internet. O número de pedidos que chegam por qualquer meio (no horário comercial) é uma variável aleatória discreta com distribuição Poisson com taxa de 5 pedidos por hora.
    a) Calcule a probabilidade de mais de 2 pedidos por hora.
    b) Em um dia de trabalho (8 horas), qual seria a probabilidade de haver 50 pedidos?
    c) Não haver nenhum pedido, em um dia de trabalho, é um evento raro?

12. Na comunicação entre servidores, uma mensagem é dividida em n pacotes, os quais são enviados na forma de códigos. Pelo histórico da rede, sabe-se que cada pacote tem uma probabilidade de 0.01 de não chegar corretamente a seu destino e, além disto, assume-se que o fato de um pacote chegar ou não corretamente ao destino não altera a probabilidade de chegada correta de outros pacotes. Um programa corretivo garante o envio correto da mensagem quando o número de pacotes enviados erroneamente não passar de 10% do total de pacotes da mensagem.
    a) Qual a probabilidade de uma mensagem composta de 20 pacotes ser enviada corretamente?
    b) E para uma mensagem de 200 pacotes?
13. Em um laticínio, a temperatura ideal do pasteurizador deve ser de 75°C. Se a temperatura ficar inferior a 70°C, o leite poderá ficar com bactérias indesejáveis ao organismo humano. Observações do processo mostram que, na forma de operação atual, os valores da temperatura seguem uma distribuição normal com média de 74.2°C e desvio padrão de 2.2°C.
    a) Qual a probabilidade da temperatura ficar inferior a 70°C?
    b) Qual a probabilidade da temperatura ultrapassar os 75°C desejados?
    c) Qual a probabilidade de que, em 20 pasteurização, alguma(s) dela(s) não atinja a temperatura de 70°C?
    d) Deseja-se regular equipamentos para alterar a temperatura média do processo para que a probabilidade de ficar inferior a 70°C seja de, no máximo, 0.0005. Qual deveria ser a nova média de operação?
    e) Suponha, agora, que a nova média de operação seja de 74.5°C. Deseja-se alterar o desvio padrão para satisfazer as condições do ítem anterior. Qual deve ser o novo desvio padrão de operação?

---

1. a) $A \cup B$.
   b) $A \cap B^c$.
   c) $A^c \cap B^c$.
   d) $(A \cap B^c) \cup (A^c \cap B)$.

---

2. Os cálculos ficam facilitados com a tabela abaixo:

Atividade	Biol. Noite (N)	Biol. Diuno (D)	Outros (O)	Total
Esportista (E)	200	100	3700	4000
Não esportista (E^c)	500	400	5100	6000
Total	700	500	8800	10000

a) Ser esportista: $P(E) = \frac{4000}{10000} = 0.4$.
b) Ser esportista e aluno da biologia noturno: $P(E \cap N) = \frac{200}{10000} = 0.02$.
c) Não ser da biologia: $P(O) = \frac{8800}{10000} = 0.88$.
d) Ser esportista ou aluno da biologia: $P(E \cup N \cup D) = P(E) + P(N) + P(D) - P(E \cap N) - P(E \cap D) = 0.49$.
e) Não ser esportista, nem aluno da biologia: $P(E^c \cap O) = \frac{5100}{10000} = 0.51$.

---

3. Sejam os eventos C: chove em setembro e G: o São Paulo ganha um jogo. A partir do enunciado, temos $P(G|C) = 0.7$, $P(G|C^c) = 0.8$ e $P(C) = 0.3$. Então

$$P(G) = P(G \cap C) + P(G \cap C^c) = P(G|C)P(C) + P(G|C^c)P(C^c) = 0.7 \cdot 0.3 + 0.8 \cdot (1 - 0.3) = 0.77$$

$$P(C|G) = \frac{P(C \cap G)}{P(G)} = \frac{P(G|C)P(C)}{P(G)} = 0.273.$$

4.

a) $P(P|M) = \frac{62}{136+92+248+102+195+62} = 0.074$.

b) $P(C) = \frac{136+102}{136+92+248+102+195+62} = 0.285.$

c) $P(H \cup R) = P(H) + P(R) - P(H \cap R) = \frac{(136+92+248)+(92+195)-(92)}{136+92+248+102+195+62} = 0.804.$

d) $P(P|H) = \frac{248}{136+92+248} = 0.521.$

---

5. Defina os eventos C: ser contra a reforma agrária, A: ser a favor da reforma agrária, O: ser sem opinião quanto a reforma agrária, F: ser do sexo feminino e M: ser do sexo masculino.

a) Uma pessoa do sexo masculino e sem opinião sobre a reforma agrária:

$$P(M \cap O) = \frac{8+1}{2+8+2+8+9+8+4+8+2+12+10+1} = \frac{9}{74} = 0.122.$$

b) Uma mulher contrária a reforma agrária:

$$P(F \cap C) = \frac{6}{2+8+2+8+9+8+4+8+2+12+10+1} = \frac{6}{74} = 0.081.$$

c) Dentre os estudantes do noturno, um que seja a favor da reforma agrária:

$$P(A|N) = \frac{8+10}{4+8+2+12+10+1} = \frac{18}{37} = 0.486.$$

d) Uma pessoa sem opinião, sabendo-se que ela é do sexo feminino:

$$P(O|F) = \frac{2+2}{2+8+2+4+8+2} = \frac{4}{26} = 0.154.$$

6. Num certo bairro da cidade de São Paulo, as companhias de seguro estabeleceram o seguinte modelo para número de veículos furtados por semana:

Furtos (F)	0	1	2	3	4
p_i	1/4	1/2	1/8	1/16	1/16

$E(F) = 0 \cdot 1/4 + 1 \cdot 1/2 + 2 \cdot 1/8 + 3 \cdot 1/16 + 4 \cdot 1/16 = 1.19.$

$V(F) = (0-\mu(F))^2 \cdot 1/4 + (1-\mu(F))^2 \cdot 1/2 + (2-\mu(F))^2 2 \cdot 1/8 + (3-\mu(F))^2 \cdot 1/16 + (4-\mu(F))^2 \cdot 1/16 = 1.15.$

---

7. Para ser uma função de densidade de probabilidade é necessário satisfazer duas propriedades:

I) $f(x) \geq 0$

II) $\int_{-\infty}^{+\infty} f(x)dx = 1.$

a) $f(x) \geq 0 \quad \forall x.$

$\int_0^1 3xdx = \frac{3}{2}.$

b) $f(x) \geq 0 \quad \forall x.$

$\int_0^\infty \frac{x^2}{2}dx = \text{diverge}.$

c) $f(x) \geq 0 \quad \forall x.$

$\int_3^5 \frac{x-3}{3}dx = 1.$

d) $f(x) \geq 0 \quad \forall x.$

$\int_0^2 2dx = 4.$

e) $f(x) \geq 0 \quad \forall x.$24. Na comunicação entre servidores, uma mensagem é dividida em n pacotes, os quais são enviados na forma de códigos. Pelo histórico da rede, sabe-se que cada pacote tem uma probabilidade de 0.01 de não chegar corretamente a seu destino e, além disto, assume-se que o fato de um pacote chegar ou não corretamente ao destino não altera a probabilidade de chegada correta de outros pacotes. Um programa corretivo garante o envio correto da mensagem quando o número de pacotes enviados erroneamente não passar de 10% do total de pacotes da mensagem.
   a) Qual a probabilidade de uma mensagem composta de 20 pacotes ser enviada corretamente?
   b) E para uma mensagem de 200 pacotes?

$\int_{-2}^0 \frac{2+x}{4} dx + \int_0^2 \frac{2-x}{4} dx = \frac{1}{2} + \frac{1}{2} = 1.$

Portanto, temos uma fdp apenas nas letras c) e e).

---

8. a)

X	-1	1
p_i	0.6	0.4

Y	-2	0	2	4
p_i	0.3	0.2	0.2	0.3

   b) Para que X e Y sejam independentes, todas a probabilidades conjuntas devem ser iguais ao produto das correspondentes probabilidades marginais. Este não é o caso pois $P(X = 1; Y = 0) = 0 \neq 0.08 = 0.4 \times 0.2 = P(X = 1)P(Y = 0)$. Logo, concluímos que as variáveis não são independentes.

   c) Para o cálculo de covariância, obtemos $E(X) = -0.2$; $E(Y) = 1$; e $E(XY) = -0.6$. Assim, $Cov(X,Y) = E(XY) - E(X)E(Y) = -0.4$. Para o cálculo da correlação, precisamos obter $V(X)$ e $V(Y)$. Temos $V(X) = 0.96$; $V(Y) = 5.8$, e, então, $\rho(X,Y) = -0.17$.

---

9. Um usuário de transporte coletivo chega pontualmente às 8 horas para pegar o seu ônibus. Devido ao transito caótico, a demora pode ser qualquer tempo entre 1 e 20 minutos (admita que o relógio “pule” de minuto em minuto). Pergunta-se:

$$\begin{aligned} D &: \text{tempo de espera, } D \sim U_D(1,20) \\ d &\in \{0,1,2,\cdots,20\} \\ P(D = d) &= 1/20 \end{aligned}$$

   a) Qual a probabilidade de demorar mais de 10 minutos?

$$P(D > 10) = P(D = 11) + \cdots + P(D = 20) = 10/20 = 0.5.$$

   b) Qual a probabilidade de demorar pelo menos 5 mas não mais de 10 minutos?

$$P(5 \leq D \leq 10) = P(D = 5)+P(D = 6)+P(D = 7)+P(D = 8)+P(D = 9)+P(D = 10) = 6/20 = 0.3.$$

   c) Qual a probabilidade da demora não chegar a 5 minutos?

$$P(D < 5) = P(D = 1) + \cdots + P(D = 4) = 4/20 = 0.2.$$

   d) Se um amigo chegou 10 minutos atrasado e vai pegar o mesmo ônibus (que ainda não passou), qual a probabilidade do amigo atrasado esperar até 3 minutos?

$$P(D \leq 13|D > 10) = \frac{P(10 < D \leq 13)}{P(D > 10)} = \frac{3/20}{10/20} = 0.3.$$

---

10. Uma certa doença pode ser curada através de procedimento cirúrgico em 80% dos casos. Dentre os que têm essa doença, sorteamos 15 pacientes que serão submetidos à cirurgia. Fazendo alguma suposição

adicional que julgar necessária, responda qual é a probabilidade de:

$$\begin{aligned}&Y: \text{número de pacientes curados}\\&p: \text{probabilidade de um paciente curar}\\&p = 0.8\\&n = 15\\&Y \sim b(15, 0.8).\end{aligned}$$

a) Todos serem curados?

$$P(Y = 15) = \binom{15}{15} \cdot 0.8^{15} \cdot 0.2^{15-15} = 0.035.$$

b) Pelo menos dois não serem curados?

$$P(Y \leq 13) = 1 - P(Y > 13) = 1 - (P(Y = 14) + P(Y = 15)) = 0.833.$$

c) Ao menos 10 ficarem livres da doença?

$$P(Y \geq 10) = P(Y = 10) + \cdots + P(Y = 15) = 0.939.$$

---

11. Uma indústria de tintas recebe pedidos de seus vendedores através de fax, telefone e Internet. O número de pedidos que chegam por qualquer meio (no horário comercial) é uma variável aleatória discreta com distribuição Poisson com taxa de 5 pedidos por hora.

$$\begin{aligned}&Y: \text{número de pedidos.}\\&Y \sim Po(5), y = \{0, 1, 2, \cdots\}.\\P(Y = y) &= \frac{e^{-5}5^y}{y!}.\end{aligned}$$

a) Calcule a probabilidade de mais de 2 pedidos por hora.

$$P(Y > 2) = 1 - P(Y \leq 2) = 1 - (P(Y = 0) + P(Y = 1) + P(Y = 2)) = 0.875.$$

b) Em um dia de trabalho (8 horas), qual seria a probabilidade de haver 50 pedidos? Considere $\lambda = 5 \cdot 8 = 40$.

$$P(Y = 50) = \frac{e^{-40}40^{50}}{50!} = 0.018.$$

Dica: para facilitar os cálculos use a transformação log (logarítmo natural)

$$\begin{aligned}logP(Y = 50) &= log(e^{-40}) + log(40^{50}) - log(50 \cdot 49 \cdots 1)\\&= -40 \cdot log(e) + 50 \cdot log(40) - (log50 + log49 + \cdots + log1)\\&= -40 + 50 \cdot 3.69 - 148.48\\&= -3.98.\end{aligned}$$

Então, $P(Y = 50) = e^{-3.98} = 0.0187$.

c) Não haver nenhum pedido, em um dia de trabalho, é um evento raro? Considere $\lambda = 40$.

$$P(Y = 0) = \frac{e^{-40}40^0}{0!} = 0.$$

---

12. Na comunicação entre servidores, uma mensagem é dividida em n pacotes, os quais são enviados na forma de códigos. Pelo histórico da rede sabe-se que cada pacote tem uma probabilidade de 0.01 de não chegar corretamente a seu destino, e além disto, assume-se que o fato de um pacote chegar ou não corretamente ao destino não altera a probabilidade de chegada correta de outros pacotes. Um programa corretivo, garante o envio correto da mensagem quando o número de pacotes enviados erroneamente não passar de 10% do total de pacotes da mensagem.

a) Qual a probabilidade de uma mensagem composta de 20 pacotes ser enviada corretamente?

$$\begin{aligned} Y &: \text{ número de pacotes incorretos em 20 pacotes.} \\ Y &\sim b(n = 20, p = 0.01). \\ \text{limite} &: 10\% \text{ de } 20 \text{ pacotes} = 2 \text{ pacotes.} \end{aligned}$$

$$P(Y \leq 2) = P(Y = 0) + P(Y = 1) + P(Y = 2) = \sum_{0}^{2} \binom{20}{y} (0.01)^y (1 - 0.01)^{20-y} = 0.999.$$

b) E para uma mensagem de 200 pacotes?

$$\begin{aligned} X &\sim b(n = 200, p = 0.01) \approx P(\lambda = n \cdot p = 200 \cdot 0.01 = 2). \\ &\approx N(\mu = n \cdot p = 2, \sigma^2 = n \cdot p \cdot (1-p) = 1.98). \\ &\text{A aproximação normal não é muito acurada pois } np < 5, \text{ porém conveniente} \\ \text{limite} &: 10\% \text{ de } 200 \text{ pacotes} = 20 \text{ pacotes.} \end{aligned}$$

$$\begin{aligned} P(Y \leq 20) &= P(Y = 0) + P(Y = 1) + \ldots + P(Y = 20) = \\ &= \sum_{0}^{20} \binom{200}{y} (0.01)^y (1 - 0.01)^{200-y} \approx \sum_{0}^{20} \frac{e^{-2}\, 2^y}{y!} \approx P(Y_N < 20.5) = P(Z < \frac{20.5 - 2}{\sqrt{1.98}}) \approx 1. \end{aligned}$$

---

13. Em um laticínio, a temperatura ideal do pasteurizador deve ser de 75°C. Se a temperatura ficar inferior a 70°C, o leite poderá ficar com bactérias indesejáveis ao organismo humano. Observações do processo mostram que na forma de operação atual os valores da temperatura seguem uma distribuição normal com média de 74.2°C e desvio padrão de 2.2°C.

$$\begin{aligned} Y &: \text{temperatura do pasteurizador.} \\ Y &\sim N(74.2; 2.2^2). \end{aligned}$$

a) Qual a probabilidade da temperatura ficar inferior a 70°C?

$$P(Y < 70) = P(Z < (70 - 74.2)/2.2) = 0.0281.$$

b) Qual a probabilidade da temperatura ultrapassar os 75°C desejados?

$$P(Y > 75) = P(Z < (75 - 74.2)/2.2) = 0.3581.$$

c) Qual a probabilidade de que em 20 pasteurizações, alguma(s) dela(s) não atinja(m) a temperatura de 70°C?

$$\begin{aligned} Y &: \text{número de pasteurizações que não atingem } 70°C. \\ Y &\sim b(20, p). \\ p &= P(Y < 70) = 0.0281. \\ P(Y \geq 1) &= 1 - P(Y = 0) = 0.435. \end{aligned}$$

d) Deseja-se regular equipamentos para alterar a temperatura média do processo para que a probabilidade de ficar inferior a 70°C seja de no máximo 0.0005. Qual deveria ser a nova média de operação?

$$\begin{aligned} P(Y < 70|\mu_0) &= 0.0005 \\ z_{0.0005} &= (y - \mu_0)/\sigma \\ -3.291 &= (70 - \mu_0)/2.2. \\ \mu_0 &= 70 - 2.2(-3.291) = 77.2402. \end{aligned}$$

e) Suponha agora que a nova média de operação seja de 74.5°C. Deseja-se então alterar o desvio padrão para satisfazer as condições do ítem anterior. Qual deve ser o novo desvio padrão de operação?

$$
\begin{aligned}
P(Y < 70|\sigma_0) &= 0.0005. \\
z_{0.0005} &= (y - 74.5)/\sigma_0 \\
-3.291 &= (70 - 74.5)/\sigma_0 \\
\sigma_0 &= (70 - 74.5)/(-3.291) = 1.37.
\end{aligned}
$$
