

# CE-003: Estatística II - Turma K/O

## Avaliações Semanais (2º semestre 2016)

### Semana 2 (av-01)

1. Dezesesseis equipes irão disputar um torneio de jogos eliminatórios. As equipes são sorteadas para ocupar as posições iniciais da chave de jogos pré-definida. As equipes perdedoras da penúltima etapa fazem um jogo para disputar a terceira colocação e as vencedoras disputam para definir a primeira e segunda colocações. Ao final são atribuídas medalhas (ouro, prata e bronze) às três primeiras colocadas.

- (a) As regras do torneio são “justas” em relação a possibilidade de vitórias das equipes?
- (b) Quantas possíveis configurações de equipes medalhistas podem ocorrer (sem considerar a classificação entre as três primeiras)?
- (c) E considerando a classificação?

Considerando que as equipes possuem chance igual de vitória em cada jogo.

- (d) Qual a probabilidade de uma determinada equipe vencer o torneio?
- (e) Qual a probabilidade de uma determinada equipe conseguir uma medalha?
- (f) Se um país concorre com duas equipes, qual a probabilidade de que o país tenha alguma medalhista?

Agora considerando que as equipes possuem níveis técnicos diferentes.

- (g) As probabilidades calculadas nos três itens anteriores seriam diferentes?
- (h) Sugira algum procedimento para calcular tais probabilidades.

### Solução:

- (a) Sim. A partir do sorteio todos tem a mesma probabilidade (1/16) de vencer.
- (b) Quantas possíveis configurações de equipes medalhistas podem ocorrer (sem considerar a classificação entre as três primeiras)?
- (c) E considerando a classificação?

Considerando que as equipes possuem chance igual de vitória em cada jogo.

- (d) Cada time tem que jogar e vencer quatro partidas para vencer. Sob a probabilidade de 1/2 de vencer cada partida a probabilidade de vencer o torneio fica:

$$P[V_1 \cap V_2 \cap V_3 \cap V_4] = P[V_1] \cdot P[V_2] \cdot P[V_3] \cdot P[V_4] = (1/2)^4 = 1/16 = 0,0625$$

Outro argumento é o de que, se o torneio é “justo” todos têm a mesma probabilidade de vitória e portanto como são 16 equipes a probabilidade de cada uma vencer é de 1/16.

- (e)  $P[\text{medalha}] = P[\text{ouro} \cup \text{prata} \cup \text{bronze}] \stackrel{Mut.Exc.}{=} P[\text{ouro}] + P[\text{prata}] + P[\text{bronze}] = 1/16 + 1/16 + 1/16 = 3/16 = 0,1875$
- (f) Sendo as equipes  $A$  e  $B$  e denotando por  $P[A]$  e  $P[B]$  suas respectivas probabilidades de medalha seria necessário calcular:

$$P[A \cup B] = P[A] + P[B] - P[A \cap B]$$

Note-se que os eventos não são mutuamente exclusivos nem independentes. Seriam mutuamente exclusivos se fosse uma particular medalha (ouro, prata ou bronze).

- (g) Sim, pois em cada partida a probabilidade de vitória de cada equipe não seria mais 1/2
- (h) Seria necessário conhecer as probabilidades de vitória de cada equipe em cada um dos possíveis confrontos. Tais probabilidades poderiam ser modeladas como função da posição das equipes em algum “ranking”.

---

### Semana 3 (av-02)

1. Em um teste de segurança, dois indivíduos tentam invadir um sistema. A probabilidade do primeiro conseguir é de 0,20 e do segundo 0,12. Faça suposição(ões) necessária(s) e responda:

- (a) Qual a probabilidade de ambos invadirem o sistema?
- (b) Qual a probabilidade do sistema não ser invadido?
- (c) Qual(ais) suposição(ões) foi(ram) feita(s)?

**Solução:**

Notação:

 $A$  : primeiro indivíduo invade o sistema $B$  : segundo indivíduo invade o sistema

(a)  $P[A \cap B] \stackrel{ind}{=} P[A] \cdot P[B] = 0,20 \cdot 0,12 = 0,024$

(b)  $P[\bar{A} \cap \bar{B}] \stackrel{ind}{=} P[\bar{A}] \cdot P[\bar{B}] = 0,80 \cdot 0,88 = 0,704$

Solução alternativa:

$$1 - P[A \cup B] = 1 - [P[A] + P[B] - P[A \cap B]] \stackrel{ind}{=} 1 - (P[A] + P[B] - P[A] \cdot P[B]) = 1 - (0,20 + 0,12 - 0,20 \cdot 0,12) = 0,704$$

(c) Suposição de independência entre os eventos  $A$  consegue invadir e  $B$  consegue invadir.

2. Uma determinada doença atinge um a cada 40.000 indivíduos de uma população. Um teste para detectar a doença fornece resultados corretos em 98% dos exames.

(a) Se um indivíduo é selecionado ao acaso da população para fazer o teste e o resultado é positivo, qual a probabilidade de que de fato tenha a doença?

(b) A probabilidade seria a mesma caso um indivíduo fizesse o teste por indicação de um médico, após um exame clínico no qual o médico suspeitou da doença. Justifique sua resposta.

**Solução:**

Notação e dados:

 $D$  : indivíduo possui doença  $\bar{D}$  : indivíduo não possui doença $P$  : teste positivo  $N$  : teste negativo

$$P[D] = 1/40000 \quad P[\bar{D}] = 39999/40000$$

$$P[P|D] = 0,98 \quad P[N|D] = 0,02$$

$$P[N|\bar{D}] = 0,98 \quad P[P|\bar{D}] = 0,02$$

(a)  $P[D|P] = \frac{P[P|D]P[D]}{P[P|D]P[D] + P[P|\bar{D}]P[\bar{D}]} = \frac{0,98 \cdot (1/40000)}{0,98 \cdot (1/40000) + 0,02 \cdot (39999/40000)} = 0,001224$

(b) Não pois nesta subpopulação de indivíduos com sintomas clínicos a probabilidade de doença não é mais 1/40000.

Por exemplo, supondo que registros mostrem que 40% dos indivíduos com sintomas clínicos de fato possuem a doença (e então  $P[D] = 0,4$ ), a probabilidade um indivíduo testado positivo ter a doença seria:

$$P[D|P] = \frac{P[P|D]P[D]}{P[P|D]P[D] + P[P|\bar{D}]P[\bar{D}]} = \frac{0,98 \cdot 0,4}{0,98 \cdot 0,4 + 0,02 \cdot 0,6} = 0,9703$$

**Semana 4 (av-03)**

1. Uma variável aleatória tem função de densidade de probabilidade dada por:

$$f(x) = \begin{cases} Kx & \text{para } 0 < x < 5 \\ 0 & \text{caso contrário} \end{cases}$$

(a) Obtenha o valor de  $K$ .(b) Obtenha  $P[X < 2]$ .(c) Obtenha  $P[1,5 < X \leq 3]$ .(d) Obtenha  $P[X \geq 2,5]$ .(e) Obtenha o valor médio de  $X$ .(f) Obtenha a expressão de  $P[X < x]$  para um valor qualquer de  $x$ .**Solução:**

(a)

$$\begin{aligned}\int_0^5 f(x)dx &= 1 \\ \int_0^5 Kxdx &= 1 \\ K \frac{x^2}{2} \Big|_0^5 &= 1 \\ K \frac{5^2 - 0^2}{2} &= 1 \\ K &= \frac{2}{25} = 0,08\end{aligned}$$

(b)  $P[X < 2] = \int_0^2 0,08xdx = 0,08 \frac{2^2 - 0^2}{2} = 0,16$ .

(c)  $P[1,5 < X \leq 3] = \int_{1,5}^3 0,08xdx = 0,08 \frac{3^2 - (1,5)^2}{2} = 0,27$ .

(d)  $P[X \geq 2,5] = \int_{2,5}^5 0,08xdx = 0,08 \frac{5^2 - (2,5)^2}{2} = 0,75$ .

(e)  $E[X] = \int_0^5 x \cdot (0,08x)dx = 0,08 \frac{5^3 - 0^3}{3} = 3,33$ .

(f)  $F(x) = P[X < x] = \int_0^x 0,08xdx = 0,08 \frac{x^2 - 0^2}{2} = 0,04x^2$ .

**Observação:** as probabilidades anteriores poderiam ter sido calculadas à partir da expressão de  $F(x)$ :

(b)  $P[X < 2] = F(2) = 0,042^2 = 0,16$ .

(c)  $P[1,5 < X \leq 3] = F(3) - F(1,5) = 0,043^2 - 0,041,5^2 = 0,36 - 0,09 = 0,27$ .

(d)  $P[X \geq 2,5] = 1 - F(2,5) = 1 - 0,042,5^2 = 1 - 0,25 = 0,75$ .

Gráficos das funções e solução computacional (utilizando a linguagem R):

```
> fx <- function(x) ifelse(x > 0 & x < 5, 0.08 * x, 0)
> Fx <- function(x) {y <- ifelse(x > 0 & x < 5, 0.04 * x^2, 0); y[x>5] <- 1; return(y)}
> curve(fx, from=-1, to=6, xlab="x", ylab="f(x)")
> curve(Fx, from=-1, to=6, xlab="x", ylab="F(x)")
```

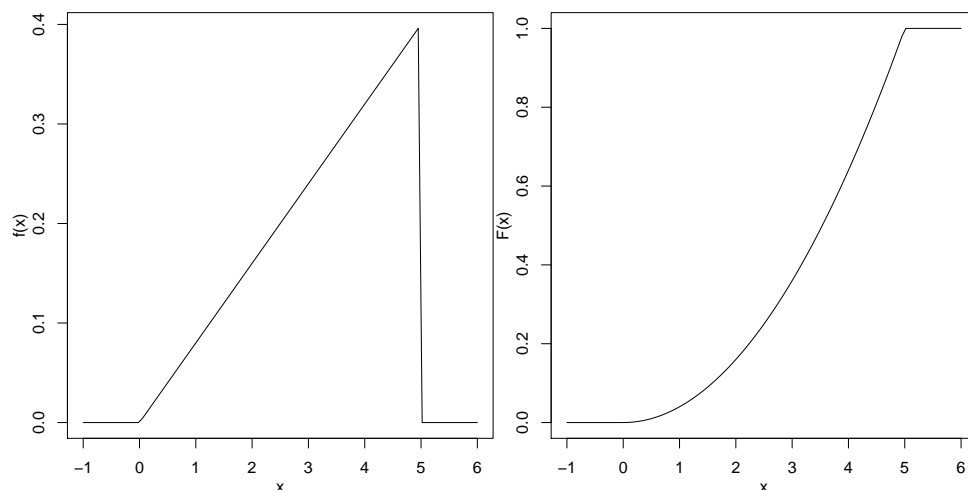


Figura 1: Funções de densidade de probabilidade  $f(x)$  (esquerda) e de probabilidade acumulada  $F(x)$  (direita)

Cálculos das probabilidade por integração de  $f(x)$  e por  $F(x)$

```
> #(b)
> integrate(fx, 0, 2)$value
[1] 0.16
> Fx(2)
[1] 0.16
> #(c)
> integrate(fx, 1.5, 3)$value
[1] 0.27
> Fx(3) - Fx(1.5)
```

```
[1] 0.27
> #(d)
> integrate(fx, 2.5, 5)$value
[1] 0.75
> 1-Fx(2.5)
[1] 0.75
```

2. Registros mostram que em um determinado site de vendas são efetuadas, em média, 3,2 transações por dia. Além disto verificou-se que a distribuição de Poisson, com função de probabilidade dada por:

$$P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!}$$

é adequada para descrever o número de vendas diárias.

- Sabe-se que para distribuição de Poisson  $E[X] = \lambda$ . Desta forma, qual a probabilidade de que em um dia o número de transações fique abaixo da média?
- Qual a probabilidade de que não haja transações em um determinado dia?
- Qual a probabilidade de que haja pelo menos duas transações em um determinado dia?
- Qual a probabilidade de que se passem dois dias consecutivos sem transações?
- Qual a probabilidade de que se tenha três dias consecutivos com número de transações abaixo da média antes que se tenha um dia com número maior que a média?

### Solução:

Notação:

$X$  : número de transações diárias no site

$X \sim P(\lambda = 3,2)$

$$P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!}$$

- $P[X \leq 3] = P[X = 0] + P[X = 1] + P[X = 2] + P[X = 3] = \sum_{i=1}^n \frac{e^{-3,2} 3,2^i}{i!} = 0.6025$
- $P[X = 0] = \frac{e^{-3,2} 3,2^0}{0!} = e^{-3,2} = 0.04076$
- $P[X \geq 2] = 1 - P[X < 2] = 1 - (P[X = 0] + P[X = 1]) = 1 - e^{-3,2}(1 + 3,2) = 0.8288$
- $P[X_1 = 0 \cap X_2 = 0] \stackrel{ind}{=} P[X_1 = 0] \cdot P[X_2 = 0] = e^{-3,2} \cdot e^{-3,2} = 0.001662$

Solução alternativa:

$Y_1$  : número de transações em dois dias no site

$Y_1 \sim P(\lambda = 2 \cdot 3,2 = 6,4)$

$$P[Y_1 = 0] = \frac{e^{-6,4} 6,4^0}{0!} = 0.001662$$

- $P[\cdot] = P[X \leq 3]^3 \cdot P[X > 3] = 0.08694$

Solução alternativa:

$Y_2$  : número de dias com transação abaixo da média até o primeiro dia com transações acima da média

$Y_2 \sim G(p = P[X > 3] = 0.3975)$

$$P[Y_2 = y] = (1 - p)^x \cdot p = (0.6025)^3 \cdot 0.3975 = 0.08694$$

Soluções computacionais com o programa R:

```
> (pa <- ppois(3, lambda=3.2))
[1] 0.6025
> (pb <- dpois(0, lambda=3.2))
[1] 0.04076
> (pc <- ppois(1, lambda = 3.2, lower=F))
[1] 0.8288
> (pd <- dpois(0, lambda=6.4))
[1] 0.001662
```

```
> (pe <- dgeom(3, prob=1-pa))
```

```
[1] 0.08694
```

---

## Semana 5 (av-04)

1. Um indivíduo vai participar de uma competição que consiste em responder questões que são lhe são apresentadas sequencialmente. Com o nível de conhecimento que possui, a chance de acertar uma questão escolhida ao acaso é de 75%. Neste contexto, para cada diferente situação apresentada a seguir, defina a variável aleatória, sua distribuição de probabilidades e calcule a probabilidade solicitada. Se preciso, faça suposições necessárias e adequadas em cada caso.

- Se for responder até errar uma pergunta, qual a probabilidade de conseguir acertar quatro ou mais questões?
- Se for responder cinco perguntas, qual a probabilidade de acertar quatro ou mais?
- Se for responder até acertar a terceira pergunta, qual a probabilidade de errar apenas uma?
- Se o candidato selecionar aleatoriamente seis questões de um banco de 40 questões das quais o candidato sabe a resposta de 30 delas (75%), qual a probabilidade de acertar ao menos cinco delas.

Ainda neste contexto considere que o candidato responde, em média, 1,8 questões por minuto.

- Qual a probabilidade de conseguir responder ao menos três questões em três minutos?
- Qual a probabilidade de que o tempo para resposta de uma questão seja superior a 40 segundos?

### Solução:

(a)

$X$  : Número de acertos até o primeiro erro

$X \sim G(0, 25)$

$$P[X \geq 4] = 1 - P[X \leq 3] = 1 - \sum_{i=0}^3 (1 - 0,25)^i (0,25) = 0.316$$

(b)

$X$  : Número de acertos em cinco perguntas

$X \sim B(n = 5, p = 0,75)$

$$P[X \geq 4] = P[X = 4] + P[X = 5] = \sum_{i=4}^5 \binom{5}{i} 0,75^i (1 - 0,75)^{5-i} = 0.633$$

(c)

$X$  : Número de erros até o terceiro acerto

$X \sim BN(r = 3, p = 0,75)$

$$P[X = 1] = \binom{3+1-1}{3-1} 0,75^3 (1 - 0,75)^1 = 0.316$$

(d)

$X$  : Número de acertos nas seis questões selecionadas

$X \sim HG(30, 10, 6)$

$$P[X \geq 5] = P[X = 5] + P[X = 6] = \sum_{i=5}^6 \frac{\binom{30}{i} \binom{10}{6-i}}{\binom{40}{6}} = 0.526$$

(e)

$X$  : Número de questões respondidas em 3 minutos

$X \sim P(3 \cdot 1,8 = 5,4)$

$$P[X \geq 3] = 1 - P[X \leq 2] = 1 - \sum_{i=0}^2 \frac{e^{-5,4} 5,4^i}{i!} = 0.905$$

(f)

$X$  : tempo (em min.) para responder uma questão

$X \sim \text{Exp}(\lambda = 1, 8)$

$$P[X \geq 40/60] = \int_{40/60}^{\infty} 1,8e^{-1,8x} dx = 0.301$$

**Soluções computacionais com o programa R:**

```
> (pa <- pgeom(3,p=0.25, lower=F))
[1] 0.3164
> (pb <- pbinom(3, size=5, prob=0.75, lower=F))
[1] 0.6328
> (pc <- dnbinom(1, size=3, prob=0.75))
[1] 0.3164
> (pd <- phyper(4, m=30, n=10, k=6, lower=F))
[1] 0.526
> (pe <- ppois(2, lam=5.4, lower=F))
[1] 0.9052
> (pf <- pexp(40/60, rate=1.8, lower=F))
[1] 0.3012
```

- 
2. Assume-se que o tempo entre acessos a um blog tem uma distribuição exponencial com média de 1,5 segundos.
- (a) Qual a probabilidade de haver duas conexões com intervalo inferior a 1,5 segundos?
  - (b) Qual a probabilidade de se passarem 5 segundos sem conexão alguma?
  - (c) Tendo havido uma conexão, qual a probabilidade da próxima conexão ocorrer entre 0,5 e 2,5 segundos?
  - (d) Se já se passou 1 segundo sem conexão, qual a probabilidade de se passar mais 0,5 segundos adicionais sem conexão?
  - (e) Qual a probabilidade do intervalo entre conexões não superar 3,5 segundos se já se passaram 2 segundos sem conexão?

**Solução:**

$X$  : intervalo de tempo entre conexões (segundos)

$X \sim \text{Exp}(\lambda = 1/1,5 = 2/3)$

$$f(x) = \frac{2}{3}e^{-2x/3} \mathbf{I}_{(0,\infty)}(x) \quad F(x) = 1 - e^{-2x/3}$$

- (a)  $P[X < 1,5] = \int_0^{1,5} f(x)dx = F(1,5) = 0.63$
- (b)  $P[X > 5] = \int_5^{\infty} f(x)dx = 1 - F(5) = 0.036$
- (c)  $P[X < 0,5] = \int_0^{2,5} f(x)dx = F(2,5) - F(0,5) = 0.53$
- (d)  $P[X > 1,5|X > 1] = \frac{\int_{1,5}^{\infty} f(x)dx}{\int_1^{\infty} f(x)dx} = {}^1P[X > 0,5] = 1 - F(0,5) = 0.72$
- (e)  $P[X < 3,5|X > 2] = \frac{\int_2^{3,5} f(x)dx}{\int_2^{\infty} f(x)dx} = \frac{F(3,5)-F(2)}{1-F(2)} = {}^1P[X < 1,5] = F(1,5) = 0.63$

**Soluções computacionais com o programa R:**

```
> (pa <- pexp(1.5, rate=2/3))
[1] 0.6321
> (pb <- pexp(5, rate=2/3, lower=F))
[1] 0.03567
> (pc <- diff(pexp(c(0.5,2.5), rate=2/3)))
[1] 0.5277
```

---

<sup>1</sup>propriedade de falta de memória da exponencial

```

> (pd <- pexp(0.5, rate=2/3, lower=F))
[1] 0.7165
> (pe <- pexp(1.5, rate=2/3))
[1] 0.6321

```

---

## Semana 7 (av-05)

- Seja uma v.a.  $X$  com distribuição normal de média  $\mu = 250$  e variância  $\sigma^2 = 225$ . Obtenha:
  - $P[X > 270]$ .
  - $P[X < 220]$ .
  - $P[|X - \mu| > 25]$ .
  - $P[|X - \mu| < 30]$ .
  - $P[X < 270 | X > 250]$ .
  - o valor  $x_1$  tal que  $P[X > x_1] = 0,80$ .
  - o valor  $x_2$  tal que  $P[X < x_2] = 0,95$ .
  - qual deveria ser um novo valor da média  $\mu$  para que  $P[X < 240] \leq 0,10$  ?
  - com  $\mu = 250$  qual deveria ser um novo valor da variância  $\sigma^2$  para que  $P[X < 240] \leq 0,10$  ?
  - qual deveria ser um novo valor da variância  $\sigma^2$  para que  $P[|X - \mu| > 15] \leq 0,10$  ?

### Solução:

$$X \sim N(250, 15^2)$$

- $P[X > 270] = P[Z > \frac{270-250}{15}] = P[Z > 1.3333] = 0.0912$
- $P[X < 220] = P[Z < \frac{220-250}{15}] = P[Z < -2] = 0.0228$
- $P[|X - \mu| > 25] = P[X < 225 \cup X > 275] = P[Z < -1.667] + P[Z > 1.667] = 0.0956$
- $P[|X - \mu| < 30] = P[220 < X < 280] = P[-2 < Z < 2] = 0.9545$
- $P[X < 270 | X > 250] = \frac{P[250 < X < 270]}{P[X > 250]} = \frac{0.4088}{0.5} = 0.8176$
- $z = \frac{x_1 - 250}{15} = -0.842 \rightarrow x_1 = 237.4$
- $z = \frac{x_2 - 250}{15} = 1.645 \rightarrow x_2 = 274.7$
- $z = \frac{240 - \mu}{15} = -1.282 \rightarrow \mu = 259.2$
- $z = \frac{240 - 250}{\sigma} = -1.282 \rightarrow \sigma = 7.8 \rightarrow \sigma^2 = 60.8$
- $P[|X - \mu| > 15] = P[X < \mu - 15 \cup X > \mu + 15] \leq 0,10 \rightarrow z = \frac{15}{\sigma} = 1.645 \rightarrow \sigma = 9.1 \rightarrow \sigma^2 = 83.1$

Comandos em R para soluções:

```

> (qa <- pnorm(270, mean=250, sd=15, lower=FALSE))
[1] 0.09121
> (qb <- pnorm(220, mean=250, sd=15))
[1] 0.02275
> (qc <- 2*pnorm(250-25, mean=250, sd=15))
[1] 0.09558
> (qd <- diff(pnorm(c(250-30,250+30), mean=250, sd=15)))
[1] 0.9545
> (qe <- diff(pnorm(c(250,270), mean=250, sd=15))/pnorm(250, mean=250, sd=15, lower=FALSE))
[1] 0.8176
> (qf <- qnorm(0.80, mean=250, sd=15, lower=FALSE))
[1] 237.4
> (qg <- qnorm(0.95, mean=250, sd=15))
[1] 274.7
> (qh <- 240 - 15 * round(qnorm(0.10), dig=3))
[1] 259.2
> (qi <- (240 - 250)/round(qnorm(0.10), dig=3))
[1] 7.8
> (qj <- 15/round(qnorm(0.95), dig=3))
[1] 9.119

```

---