CE-003: Estatística II - Turma K/O
Avaliações Semanais (1° semestre 2015)

Semana 3 (av-01)

1. Considere um jogo com um baralho (52 cartas) no qual em uma primeira rodada retira-se duas cartas e em uma segunda rodada retira-se uma carta. O interesse é se as cartas são figuras (valete, dama ou rei) de qualquer naipe. Temos interesse em:
   - obter o espaço amostral;
   - obter a probabilidade de cada ponto amostral;
   - obter a distribuição de probabilidades do número de figuras obtidas nas três cartas.

   Deve-se considerar duas situações, com e sem reposição das cartas entre a primeira e a segunda rodada.

   Solução:
   Notação:

$$F : \text{a carta é uma figura}$$
$$N = \overline{F} : \text{a carta não é uma figura}$$

- O espaço amostral para as duas situações (com e sem reposição) é o mesmo.

$$\Omega = \{(FF, F); (FF, N); (FN, F); (NF, F); (FN, N); (NF, N); (NN, F); (NN, N)\}$$

- Já as probabilidades são afetadas por repor ou não as cartas

Ponto amostral	(FF,F)	(FF,N)	(FN,F)	(NF,F)	(FN,N)	(NF,N)	(NN,F)	(NN,N)
Com reposição	$\frac{12}{52}\frac{11}{51}\frac{12}{52}$	$\frac{12}{52}\frac{11}{51}\frac{40}{52}$	$\frac{12}{52}\frac{40}{51}\frac{12}{52}$	$\frac{40}{52}\frac{12}{51}\frac{12}{52}$	$\frac{12}{52}\frac{40}{51}\frac{40}{52}$	$\frac{40}{52}\frac{12}{51}\frac{40}{52}$	$\frac{40}{52}\frac{39}{51}\frac{12}{52}$	$\frac{40}{52}\frac{39}{51}\frac{40}{52}$
Sem reposição	$\frac{12}{52}\frac{11}{51}\frac{10}{50}$	$\frac{12}{52}\frac{11}{51}\frac{40}{50}$	$\frac{12}{52}\frac{40}{51}\frac{11}{50}$	$\frac{40}{52}\frac{12}{51}\frac{11}{50}$	$\frac{12}{52}\frac{40}{51}\frac{39}{50}$	$\frac{40}{52}\frac{12}{51}\frac{39}{50}$	$\frac{40}{52}\frac{39}{51}\frac{12}{50}$	$\frac{40}{52}\frac{39}{51}\frac{38}{50}$

-

$$X : \text{número de figuras obtidas nas três cartas}$$
$$x \in \{0, 1, 2, 3\}$$

Com reposição

x	0	1	2	3
P[X=x]	P[(NN,N)]	P[(FN,N)] + P[(NF,N)] + P[(NN,F)]	P[(FF,N)] + P[(FN,F)] + P[(NF,F)]	P[(FFF)]
	$\frac{40}{52}\frac{39}{51}\frac{40}{52}$	$\frac{12}{52}\frac{40}{51}\frac{40}{52} + \frac{40}{52}\frac{12}{51}\frac{40}{52} + \frac{40}{52}\frac{39}{51}\frac{12}{52}$	$\frac{12}{52}\frac{11}{51}\frac{40}{52} + \frac{12}{52}\frac{40}{51}\frac{12}{52} + \frac{40}{52}\frac{12}{51}\frac{12}{52}$	$\frac{12}{52}\frac{11}{51}\frac{12}{52}$

Sem reposição

x	0	1	2	3
P[X=x]	P[(NN,N)]	P[(FN,N)] + P[(NF,N)] + P[(NN,F)]	P[(FF,N)] + P[(FN,F)] + P[(NF,F)]	P[(FFF)]
	$\frac{40}{52}\frac{39}{51}\frac{38}{50}$	$\frac{12}{52}\frac{40}{51}\frac{39}{50} + \frac{40}{52}\frac{12}{51}\frac{39}{50} + \frac{40}{52}\frac{39}{51}\frac{12}{50}$	$\frac{12}{52}\frac{11}{51}\frac{40}{50} + \frac{12}{52}\frac{40}{51}\frac{11}{50} + \frac{40}{52}\frac{12}{51}\frac{11}{50}$	$\frac{12}{52}\frac{11}{51}\frac{10}{50}$
	0.4471	0.4235	0.1195	0.009955

OBS: no caso sem reposição a v.a. X segue uma distribuição hipergeométrica e as probabilidades podem ser obtidas pela função de probabilidade desta distribuição.

$$X \sim \text{HG}(N = 52, n = 3, k = 12)$$
$$P[X = x] \sim \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$
$$P[X = 0] = \frac{\binom{12}{0}\binom{52-12}{3-0}}{\binom{52}{3}} = 0.4471$$
$$P[X = 1] = \frac{\binom{12}{1}\binom{52-12}{3-1}}{\binom{52}{3}} = 0.4235$$
$$P[X = 2] = \frac{\binom{12}{2}\binom{52-12}{3-2}}{\binom{52}{3}} = 0.1195$$
$$P[X = 3] = \frac{\binom{12}{3}\binom{52-12}{3-3}}{\binom{52}{3}} = 0.009955$$

Semana 4 (av-02)

1. Considere que indivíduos vão fazer um teste online no qual questões serão apresentadas sequencialmente ao candidato. Calcule a probabilidade pedidas nos contextos de cada um dos itens a seguir. Procure identificar: a variável aleatória em questão e sua distribuição de probabilidades.
   (a) Suponha que oito (8) questões são retiradas com reposição (ou seja uma mesma questão pode ser retirada mais de uma vez) de um *banco* de 40 questões dos quais o candidato sabe responder a 25 delas. Qual a probabilidade de acertar três ou mais questões?
   (b) Idem anterior porém supondo agora que as questões não podem se repetir.
   (c) Supondo novamente reposição das questões, o candidato responde até errar pela primeira vez. Qual a probabilidade de acertar pelo menos três questões?
   (d) Idem anterior supondo que responde até errar pela terceira vez.

   Solução:

   (a)
   $$X : \text{número de questões certas entre oito questões selecionadas ao acaso (com repetição)}$$
   $$X \sim \text{B}(n = 8, p = 25/40)$$
   $$x \in \{0, 1, 2, \ldots 8\}$$
   $$P[X \geq 3] = 1 - P[X = 0] - P[X = 1] - P[X = 2] = 0.964$$

   (b)
   $$X : \text{número de questões certas entre oito questões selecionadas ao acaso (sem repetição)}$$
   $$X \sim \text{HG}(N = 40, K = 25, n = 8)$$
   $$x \in \{0, 1, 2, \ldots 8\}$$
   $$P[X \geq 3] = 1 - P[X = 0] - P[X = 1] - P[X = 2] = 0.9783$$

   (c)
   $$X : \text{número de acertos até o primeiro erro}$$
   $$X \sim \text{G}(p = 15/50)$$
   $$x \in \{0, 1, 2, \ldots\}$$
   $$P[X \geq 3] = 1 - P[X = 0] - P[X = 1] - P[X = 2] = 0.2441$$

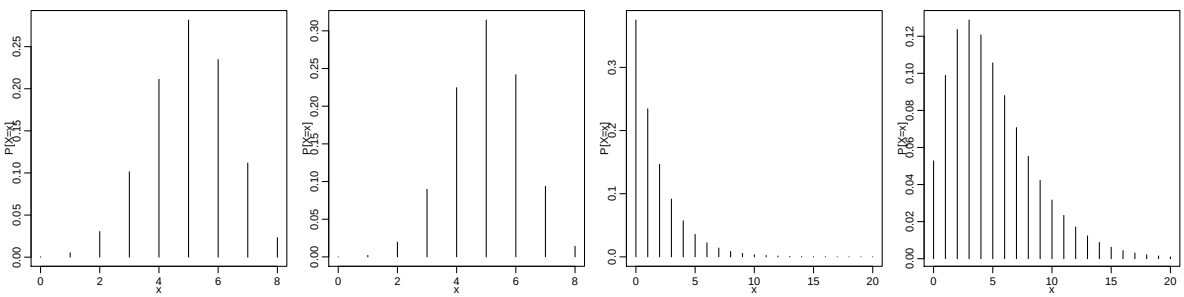
   (d)
   $$X : \text{número de acertos até o terceiro erro}$$
   $$X \sim \text{BN}(k = 3, p = 15/40)$$
   $$x \in \{0, 1, 2, \ldots\}$$
   $$P[X \geq 3] = 1 - P[X = 0] - P[X = 1] - P[X = 2] = 0.7248$$

Soluções computacionais (linguagem R):

```
> (q1 <- pbinom(2, size=8, prob=25/40, lower=FALSE))
[1] 0.964
> (q2 <- phyper(2, m=25, n=15, k=8, lower=FALSE))
[1] 0.9783
> (q3 <- pgeom(2, prob=15/40, lower=FALSE))
[1] 0.2441
> (q4 <- pnbinom(2, size=3, prob=15/40, lower=FALSE))
[1] 0.7248
```

Gráficos das distribuições de probabilidades.

```
> par(mfrow=c(1,4))
> par(mar=c(3,3,0.2, 0.2), mgp=c(1.2, 0.6, 0))
> plot(0:8, dbinom(0:8, size=8, prob=25/40), xlab="x", ylab="P[X=x]", type="h")
> plot(0:8, dhyper(0:8, m=25, n=15, k=8), xlab="x", ylab="P[X=x]", type="h")
> plot(0:20, dgeom(0:20, prob=15/40), xlab="x", ylab="P[X=x]", type="h")
> plot(0:20, dnbinom(0:20, size=3, prob=15/40), xlab="x", ylab="P[X=x]", type="h")
```



2. Um vendedor consegue vender, em média, 0,5 unidades de um produto por dia. Calcule as probabilidades de:
   (a) vender alguma unidade em um particular dia;
   (b) não efetuar nenhuma venda em uma semana (considere a semana tendo cinco dias úteis);
   (c) em uma semana (cinco dias úteis) efetuar vendas em ao menos três dias.

Solução:

(a)

$$\begin{aligned} X_1 &: \text{número de vendas em um dia} \\ x_1 &\in \{0, 1, 2, \ldots\} \\ X_1 &\sim \mathrm{P}(\lambda = 0,5) \\ P[X_1 = 0] &= \frac{\mathrm{e}^{0,5} 0,5^0}{0!} = 0.3935 \end{aligned}$$

(b)

$$\begin{aligned} X_2 &: \text{número de vendas em uma semana (cinco dias)} \\ x_2 &\in \{0, 1, 2, \ldots\} \\ X_2 &\sim \mathrm{P}(\lambda = 2,5) \\ P[X_2 = 0] &= \frac{\mathrm{e}^{2,5} 2,5^0}{0!} = 0.08208 \end{aligned}$$

(c)

$$\begin{aligned} X_3 &: \text{número de dias com vendas em uma semana (cinco dias)} \\ x_3 &\in \{0, 1, 2, 3, 4, 5\} \\ X_3 &\sim \mathrm{B}(n = 5, p = P[X_1 = 0] \\ P[X_3 \geq 3] &= P[X_3 = 3] + P[X_3 = 4] + P[X_3 = 5] = 0.6938 \end{aligned}$$

Soluções computacionais (linguagem R):

```
> (q1 <- ppois(0, lambda=0.5, lower=FALSE))
[1] 0.3935
> (q2 <- ppois(0, lambda=0.5*5))
[1] 0.08208
> (q3 <- pbinom(2, size=5, prob=dpois(0, lambda=0.5), lower=FALSE))
[1] 0.6938
```

---

3. Seja a função:

$$f(x) = \begin{cases} 3x^2/8 & 0 < x \leq 2 \\ 0 & \text{caso contrário} \end{cases}$$

(a) Mostre que $f(x)$ é uma função de densidade de probabilidade válida.
(b) Obtenha $P[0,5 < X < 1,5]$.
(c) Obtenha $P[X > 1,2]$.
(d) Obtenha $P[X > 1,2|X > 0,5]$.
(e) Obtenha o valor esperado de X.

Solução:

(a)

$$\text{Mostrar que: } f(x) \geq 0 \; \forall x \quad \text{e} \quad \int_0^2 f(x)\mathrm{d}x = 1$$

$$\frac{3}{8}\frac{2^3 - 0^3}{3} = 1$$

$$\text{a função acumulada } F(x) \text{ é dada por:} F(x) \qquad = \int_0^x f(x)\mathrm{d}x = \frac{3}{8}\frac{x^3 - 0^3}{3} = \frac{x^3}{8}$$

(b) $P[0,5 < X < 1,5] = \int_{0,5}^{1,5} f(x)\mathrm{d}x = F(1,5) - F(0,5) = 0.406$

(c) $P[X > 1,2] = \int_{1,2}^{2} f(x)\mathrm{d}x = 1 - F(1,2) = 0.784$

(d) $P[X > 1,2|X > 0,5] = \frac{\int_{1,2}^{2} f(x)\mathrm{d}x}{\int_{0,5}^{2} f(x)\mathrm{d}x} = \frac{1-F(1,2)}{1-F(0,5)} = 0.796$

(e)

$$E[X] = \int_0^2 x \cdot f(x)\mathrm{d}x = \frac{3}{8}[\frac{2^4 - 0^4}{4}] = \frac{3}{2} = 1,5$$

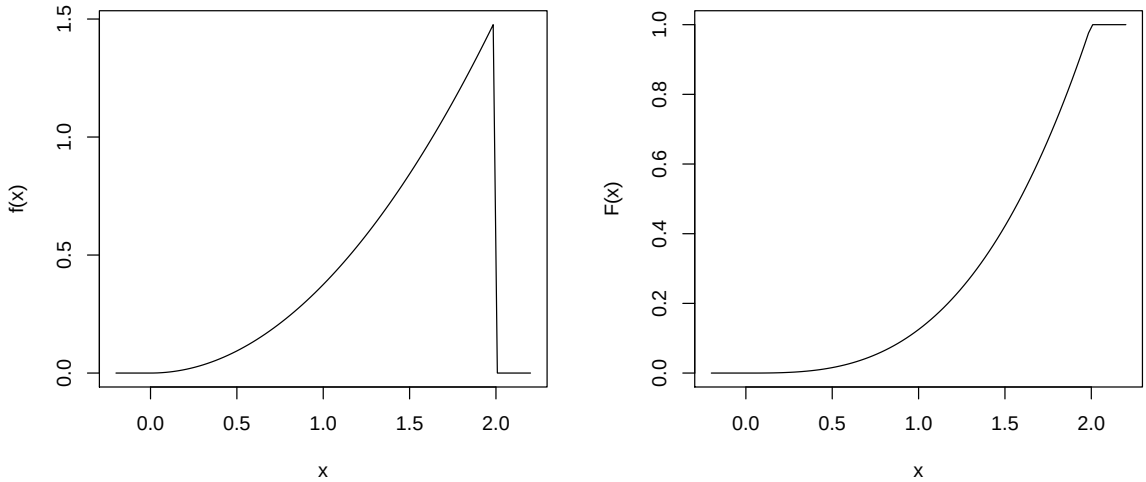


Figura 1: Função de densidade de probabilidade (esquerda) e função de distribuição (direita).

Soluções computacionais (linguagem R):

```
> require(MASS)
> ## a)
> fx <- function(x) ifelse(x > 0 & x <= 2, (3*x^2)/8, 0)
> integrate(fx, 0, 2)$value
[1] 1
> Fx <- function(x) ifelse(x>0, ifelse(x<=2, (x^3)/8,1), 0)
> Fx(2)
[1] 1
> ## b)
> integrate(fx, 0.5, 1.5)$value
[1] 0.4062
> Fx(1.5)-Fx(0.5)
[1] 0.4062
> ##c)
> integrate(fx, 1.2, 2)$value
[1] 0.784
> 1-Fx(1.2)
[1] 0.784
> ## d)
> integrate(fx, 1.2, 2)$value/integrate(fx, 0.5, 2)$value
[1] 0.7964
> (1-Fx(1.2))/(1-Fx(0.5))
[1] 0.7964
> ## e)
> efx <- function(x) ifelse(x > 0 & x <= 2, x*(3*x^2)/8, 0)
> integrate(efx, 0, 2)$value
[1] 1.5
```

Semana 5 (av-03)

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

   Solução:

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

   (a) $P[X > 270] = P[Z > \frac{270-250}{15}] = P[Z > 1.3333] = 0.0912$
   (b) $P[X < 220] = P[Z < \frac{220-250}{15}] = P[Z < -2] = 0.0228$
   (c) $P[|X - \mu| > 25] = P[X < 225 \cup X > 275] = P[X < -1.667] + P[X > 1.667] = 0.0956$
   (d) $P[|X - \mu| < 30] = P[220 < X < 280] = P[-2 < X < 2] = 0.9545$
   (e) $P[X < 270|X > 250] = \frac{P[250<X<270]}{P[X>250]} = \frac{0.4088}{0.5} = 0.8176$
   (f) $z = \frac{x_1-250}{15} = -0.842 \longrightarrow x_1 = 237.4$
   (g) $z = \frac{x_2-250}{15} = 1.645 \longrightarrow x_2 = 274.7$

(h) $z = \frac{240-\mu}{15} = -1.282 \longrightarrow \mu = 259.2$

(i) $z = \frac{240-250}{\sigma} = -1.282 \longrightarrow \sigma = 7.8 \longrightarrow \sigma^2 = 60.8$

(j) $P[|X < \mu| > 15] = P[X < \mu - 15 \cup X > \mu + 15] \leq 0,10 \longrightarrow z = \frac{15}{\sigma} = 1.645 \longrightarrow \sigma = 9.1 \longrightarrow \sigma^2 = 83.1$

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
```

Semana 6 (av-04)

1. Suponha que os escores obtidos por estudantes em um teste *online* possam ser bem modelados por uma distribuição normal com média $\mu = 120$ e variância $\sigma^2 = 12^2$.
   (a) Considera-se como estudante de alta performance os que atingem um escore a partir de 135. Qual o percentual esperado de estudantes de alta performance entre todos os que fazem o teste?
   (b) Estudantes com escore abaixo de 100 devem se reinscrever e só podem voltar a fazer o teste após seis meses e os com escore entre 100 e 125 são convidados a refazer o teste após um mês. Quais as proporções de estudantes que deverá se reinscrever e que deverá refazer o teste após um mês?
   (c) Define-se como *quartis* os escores abaixo dos quais espera-se encontrar 25, 50 e 75% dos estudantes. Quais os valores dos escores que definem os quartis?
   (d) Quanto deveria ser o valor μ da média dos escores para que ao menos 30% dos escores fossem de alta performance?
   (e) Há um outro teste que possui média $\mu = 125$ e variância $\sigma^2 = 6^2$. Em qual deles espera-se a maior proporção de estudantes de alta performance?

Solução:

$$X \sim \mathrm{N}(120, 12^2)$$

(a) $P[X > 135] = P[Z > \frac{135-120}{12}] = P[Z > 1.25] = 0.1056$

(b)

$$P[X < 100] = P[Z < \frac{100-120}{12}] = P[Z < -1.6667] = 0.0478$$
$$P[100 < X < 125] = P[\frac{100-120}{12} < Z < \frac{125-120}{12}] = P[-1.67 < Z < 0.417]$$

(c)

$$P[X < Q_1] = 0,25$$
$$z_1 = -0.674 = \frac{Q_1 - 120}{12}$$
$$Q_1 = 120 - 8.09 = 112$$

Usando o fato de que a distribuição é simétrica temos ainda que:

$$Q_2 = \mu = 120$$
$$Q_3 = 120 + 8.09 = 128$$

(d) $z = \frac{135-\mu}{15} = 0.524 \longrightarrow \mu = 128.7$

(e)

$$X_1 \sim \mathrm{N}(120, 12^2)$$
$$X_2 \sim \mathrm{N}(125, 6^2)$$
$$P[X_1 \geq 135] = P[Z_1 > \frac{135-120}{12}] = P[Z_1 > 1.25] = 0.106$$
$$P[X_2 \geq 135] = P[Z_2 > \frac{135-120}{12}] = P[Z_2 > 1.67] = 0.0478$$

Comandos em R para soluções:

```
> (qa <- pnorm(135, mean=120, sd=12, lower=FALSE))
[1] 0.1056
> (qb <- diff(pnorm(c(-Inf, 100, 125), mean=120, sd=12)))
[1] 0.04779 0.61375
> (qc <- qnorm(c(.25, .50, .75), mean=120, sd=12))
[1] 111.9 120.0 128.1
> (qd <- 135 - 12 * round(qnorm(0.70), dig=3))
[1] 128.7
> (qez <- (135 - c(120, 125))/c(12, 6))
[1] 1.250 1.667
> (qep <- pnorm(135, m=c(120, 125), sd=c(12, 6), lower=FALSE))
[1] 0.10565 0.04779
```