CE-003: Estatística II - Turma K/O
Avaliações Semanais (2^o semestre 2015)

Semana 5 (av-01)

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.
   (a) Se for responder até errar uma pergunta, qual a probabilidade de conseguir acertar quatro ou mais questões?
   (b) Se for responder cinco perguntas, qual a probabilidade de acertar quatro ou mais?
   (c) Se for responder até acertar a terceira pergunta, qual a probabilidade de errar apenas uma?
   (d) 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.

   (e) Qual a probabilidade de conseguir responder ao menos três questões em três minutos?

   Solução:

   (a)

$$X : \text{Número de acertos até o primeiro erro}$$
$$X \sim \mathrm{G}(0,25)$$
$$P[X \geq 4] = 1 - P[X \leq 3] = 1 - \sum_{i=0}^{1} (1-0,25)^i (0,25) = 0.316$$

   (b)

$$X : \text{Número de acertos em cinco perguntas}$$
$$X \sim \mathrm{B}(\mathrm{n}=5, \mathrm{p}=0,75)$$
$$P[X \geq 4] = P[X=4] + P[X=5] = \sum_{i=4}^{5} \binom{5}{i} 0,75^i (1-0,75)^{5-i} = 0.633$$

   (c)

$$X : \text{Número de erros até o terceito acerto}$$
$$X \sim \mathrm{BN}(\mathrm{r}=3, \mathrm{p}=0,75)$$
$$P[X=1] = \binom{3+1-1}{3-1} 0,75^3 (1-0,75)^1 = 0.316$$

   (d)

$$X : \text{Número de acertos nas seis questões selecionadas}$$
$$X \sim \mathrm{HG}(30,10,6)$$
$$P[X \geq 5] = P[X=5] + P[X=6] = \sum_{i=5}^{6} \frac{\binom{30}{i}\binom{10}{6-i}}{\binom{40}{6}} = 0.526$$

   (e)

$$X : \text{Número de questões respondidas em 3 minutos}$$
$$X \sim \mathrm{P}(3 \cdot 1,8 = 5,4)$$
$$P[X \geq 3] = 1 - P[X \leq 2] = 1 - \sum_{i=0}^{2} \frac{\mathrm{e}^{-5,4} 5,4^{\mathrm{i}}}{i!} = 0.905$$

---

1. 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)

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

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

a função acumulada $F(x)$ é dada por:

$$F(x) = \int_0^x f(x)\mathrm{d}x = \frac{3}{8}\frac{x^3 - 0^3}{3} = \frac{x^3}{8}$$

(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)

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

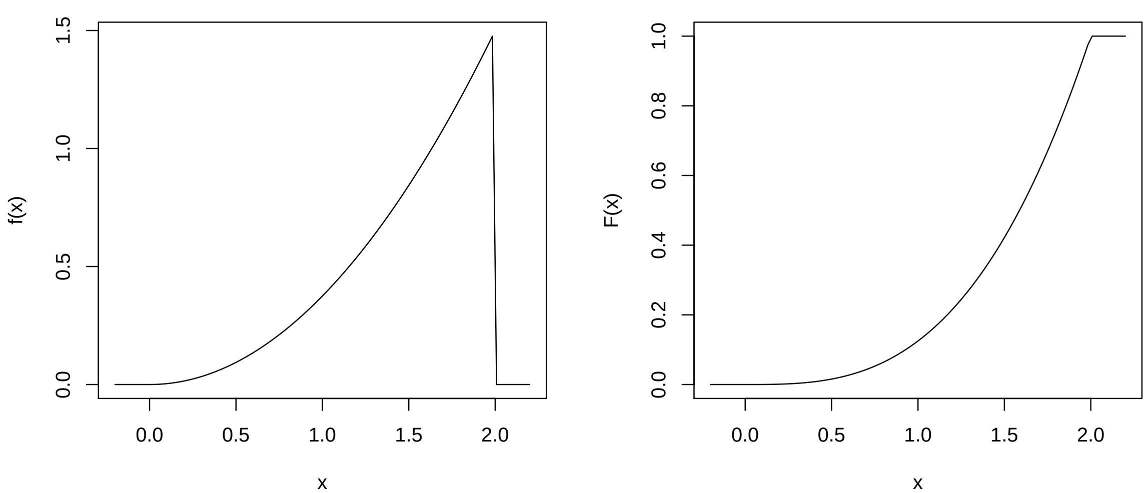

Figura 1: Função de densidade de probabilidade (esquerda) e função de distribuição (direita).

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 8 (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 9 (av-04)

1. Foi feita uma pesquisa sobre as condições salariais de professores de um certo estado. Os dados foram organizados em uma tabela. A seguir é mostrada uma pequena fração dos dados. A descrição dos atributos anotados está na Tabela abaixo.

```
  Degree Rank Sex Year YSdeg Salary
1      1    3   0   25    35  36350
2      1    3   0   13    22  35350
3      1    3   0   10    23  28200
4      1    3   1    7    27  26775
5      0    3   0   19    30  33696
6      1    3   0   16    21  28516
```

Atributo	Descrição
Degree	Formação: 1: Doutorado, 0: Mestrado
Rank	Cargo (1: Prof Assistente, 2: Prof Associado, 3: Prof Pleno)
Sex	1: feminino, 0: masculino
Year	Anos de trabalho
YSdeg	Anos desde a obtenção da maior titulação
Salary	Salário em dolares por ano

(a) Classifique cada um dos atributos (variáveis).

(b) Esboçe um gráfico adequado para resumir cada um dos atributos individualmente

(c) Como voce investigaria (por exemplo, que tipo de gráfico) se existe relação entre:

i. Sexo e Formação

ii. Sexo e Salário

iii. Anos de trabalho e salário

Solução:

(a) Sex: Qualitativa nominal
Degree, Rank: Qualitativa ordinal
Anos de trabalho*, tempo de titulação*: contínua (mas podendo ser tratada como discreta)
salário: contínua

(b) Esboçe um gráfico adequado para resumir cada um dos atributos individualmente

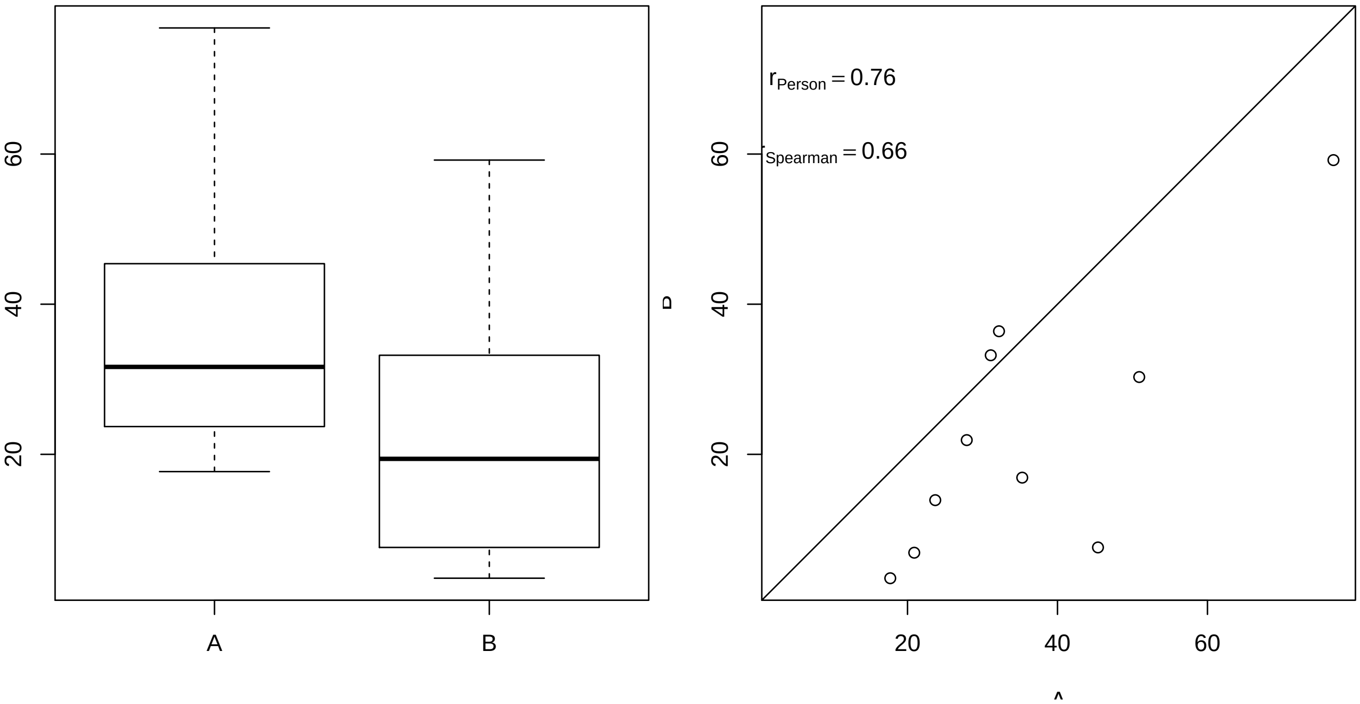

Figura 2: Box-plot e diagrama de dispersão dos tempos de processamento de dois algorítmos aplicados a um mesmo conjunto de problemas

---

Semana 11 (av-05)

1. Um conjunto de imagens foi submetido a dois algoritmos (A e B) de tratamento (filtragem, correção e classificação) e foram registrados os tempos de processamento. Alguns resumos dos dados encontram-se a seguir.

$$\overline{x}_A = 36.19 \quad \overline{x}_B = 22.98$$
$$S_A = 17.62 \quad S_B = 17.14$$

Responda as questões a seguir baseando-se nos resumos dados e justificando as respostas.

(a) Compare os desempenhos algorítmos.

(b) Existem observações discrepantes (atípicas)?

(c) Como voce descreveria a relação e correlação entre o desempenho dos algorítmos?

(d) Os algorítmos possuem variabilidades relativas, medida pelo coeficiente de variação, semelhantes?

(e) Os algorítmos possuem variabilidades, medida pela amplitude interquartílica, semelhantes?

Semana 13 (av-06)

Considere um estudo na qual deseja-se estimar a proporção de solicitações atendidas e resolvidas de uma central do usuário através de uma amostra aleatória simples.

1. Qual a população, variável aleatória, parâmetro de interesse e o estimador?
2. Se a amostra for de 4000 solicitações, qual será a margem de erro (com confiança de 95%) para a estimação da proporção de resolvidas?
3. Para este mesmo tamanho de amostra, qual seria a confiança associada a uma margem de erro de $\pm 0,01$ (1 %)?
4. Qual deveria ser o tamanho da amostra para se obter a estimativa com uma margem de erro de 2,5% com 95% de confiança?
5. E para uma margem de erro de 3% com 99% de confiança?

Solução:

1. População: solicitações atendidas, representada pela v.a.:

$$X : \text{resolução de solicitações atendidas}$$

   que é binária e pode ser codificada assumindo valores 0, para não atendidas e, 1, para atendidas. O parâmetro de interesse é a proporção de solicitações atendidas na população, que representa a probabilidade p de uma solicitação ser atendida. O estimador é a proporção de atendidas na amostra. Assume-se então que:

$$X \sim B(p)$$

   e os resultados das análises se baseiam no Teorema:

$$\text{Teo 2:} \quad \hat{p} = \overline{X} \approx N(p, \frac{p(1-p)}{n})$$

   do qual se extraem os resultados:

$$\hat{p} \pm z_{(1-\alpha)}\sqrt{\frac{p(1-p)}{n}}$$

$$\hat{p} \pm \text{ME}$$

$$\text{ME} = z_{(1-\alpha)}\sqrt{\frac{p(1-p)}{n}}$$

$$n = \frac{(z_{(1-\alpha)})^2}{\text{ME}^2}p(1-p)$$

$$p(1-p) \text{ é limitado superiormente para } p = 0,5$$

$$n = \frac{(z_{(1-\alpha)})^2}{\text{ME}^2}0,25$$

2. $M.E. = z_{(1-\alpha)}\sqrt{\frac{p(1-p)}{n}} = 1.96\sqrt{\frac{0,5(1-0,5)}{4000}} = 0.0155$
3.

$$M.E. = z_{(1-\alpha)}\sqrt{\frac{p(1-p)}{n}}$$

$$z_{(1-\alpha)} = 0,01/\sqrt{\frac{0,5(1-0,5)}{4000}} = 1.26$$

$$1-\alpha = 79.4\%$$

4. $n = \frac{(1.96)^2}{(0,025)^2}0,25 = 1537$
5. $n = \frac{(2.576)^2}{(0,03)^2}0,25 = 1844$
