# CE067 - Estatística Descritiva e Exploratória - Prova 3

1. Sobre a distribuição normal, resolva as questões abaixo, ignorando a correção de continuidade quando for o caso:

   (a) Seja X uma variável aleatória com distribuição normal com média 50 e variância 81.

      i. $P(X > 68)$

      ```
      [1] 0.02275
      ```

      ii. $P(|X - 50| < 68)$

      ```
      [1] 1
      ```

   (b) Uma clínica de emagrecimento recebe pacientes adultos com peso seguindo uma distribuição normal de média 130Kg e desvio padrão 20Kg. Para efeito de determinar o tratamento mais adequado, os 25% de menor peso são classificados de magros. Determine o valor que delimita esta classificação.

   ```
   [1] 116.51
   ```

   (c) Na distribuição $X \sim$Normal$(\mu, \sigma^2)$ encontre o número $a$ tal que $P(\mu - a\sigma < X < \mu + a\sigma) = 0,99$

   ```
   a=2,58
   ```

   (d) Com base em experiências anteriores, uma agência de banco sabe que 40% de seus clientes pagam as contas no caixa eletrônico. Dentre 600 registros de contas pagas em uma determinada semana, qual a probabilidade de ter mais de 270 contas pagas no caixa eletrônico ?

   ```
   [1] 0.006
   ```

2. Num grupo de pacientes coronarianos, sabe-se que os níveis de colesterol sérico seguem uma distribuição normal com média $\mu$ e desvio padrão $\sigma$. Observou-se que 10% do grupo tinham níveis de colesterol abaixo de 10,0 mg/10ml enquanto que 10% tinham valores acima de 35,6 mg/10ml.

   (a) Qual é a média e o desvio padrão da distribuição?

   ```
   0.1=P(X<10) e 0.1=P(X>35.6)
   0.1=P(Z<(10-mu)/sigma) e 0.1=P(Z>(35.6-mu)/sigma)
   -1.28=(10-mu)/sigma e 1.28=(35.6-mu)/sigma
   mu=10+1.28 sigma  e 1.28=(35.6-(10+1.28 sigma))/sigma
   1.28 sigma=35.6 -10 -1.28 sigma
   1.28 sigma + 1.28 sigma=35.6 -10
   2.56 sigma=25.6
   sigma=25.6/2.56= 10mg/100ml
   mu=10 + 1.28(10)= 22.8mg/10ml
   ```

   (b) Qual é a chance de um paciente selecionado aleatoriamente deste grupo ter um nível de colesterol sérico acima de 30 mg/10ml?

```
> mu = 22.8
> sigma = 10
> pr = 1 - pnorm((30 - mu)/sigma)
> round(pr, 2)
[1] 0.24
```

(c) Qual é a chance do nível **médio** de colesterol sérico de uma amostra aleatória de 16 pacientes estar acima de 30 mg/10ml?

```
> mu = 22.8
> sigma = 10/sqrt(16)
> pr = 1 - pnorm((30 - mu)/sigma)
> round(pr, 4)
[1] 0.002
```

(d) Deseja-se verificar se um certo medicamento utilizado por pacientes coronarianos reduz os níveis de colesterol sérico. Para tanto foi tomada uma amostra de 25 pacientes que fizeram uso deste medicamento diariamente por um periodo de dois meses, e o resultado observado na amostra para o nível médio de colesterol sérico foi de 19 mg/10ml. Determine o intervalo de confiança de 95% para o nível de colesterol sérico médio. Com base no intervalo de confiança obtido, qual seria sua conclusão sobre a eficácia do medicamento?

```
> sigma = 10/sqrt(25)
> i = 20 - 1.96 * sigma
> s = 20 + 1.96 * sigma
> round(cbind(i, s), 2)
         i     s
[1,] 16.08 23.92
```

(e) De qual tamanho deverá ser a amostra, para que a amplitude do intervalo de 95% de confiança para o nível médio de colesterol sérico seja de 5 mg/10ml?

```
> sigma = 10
> n = (2 * 1.96 * sigma/5)^2
> round(n)
[1] 61
```

3. O consumo médio de gasolina num certo tipo de automóvel é de 11 km/litro, segundo informações da montadora. Uma revista especializada verificou o consumo de 36 desses veículos, escolhidos ao acaso, e constatou consumo médio de 10,3 km/litro. Admita que o desvio padrão do consumo dos automóveis é de 2 km/litro.

   (a) Você conhece a distribuição do estimador do consumo médio, $\bar{X}$? Se não, é possível fazer alguma suposição? *Resp: usando o TCL* $\bar{X} \sim N(\mu, 4/36)$

   (b) Teste, ao nível de significância de 5%, a afirmação da montadora de que a média de consumo é igual a 11 km/litro, contra a alternativa de ser igual a 10 km/litro. Estabeleça as hipóteses nula e alternativa antes de fazer o teste.
   $0.05 = P(Z < (k-11)/(1/3); -1.64 = (k-11)3; k = 11 - 1.64/3 = 10.45$
   *Resp: 10.3 pertence a RC=*$\{x \in R \mid x<10.45\}$ *portanto rejeitamos* $H_0$ *ao nível de 5%*

(c) Determine a probabilidade de erro tipo II.
$P(\bar{X} > 10.45|\mu = 10) = P(Z > (10.45 - 10)/(1/3)) = P(Z > 1.35) = 0.5 - 0.411492 = 0.088$

(d) A conclusão obtida no item (b) mudaria se o nível de significância fosse aumentado para 1%?
$0.01 = P(Z < (k - 11)/(1/3); -2.33 = (k - 11)3; k = 11 - 2.33/3 = 10.22$
*Resp: 10.3 não pertence a RC=$\{x \in R \mid x<10.22\}$ portanto não podemos rejeitar $H_0$ ao nível de 1% de significância.*

(e) Que conclusão pode ser tirada através do nível descritivo?
$P(\bar{X} < 10.3|\mu = 11) = P(Z < (10.3 - 11)/(1/3)) = P(Z < -2.1) = 0.5 - 0.4821 = 0.0179$
*Resp: O nivel descritivo é de 1.8%, indicando que a probabilidade de encontrarmos valores tão baixos ou ainda menores do que o observado é menor que 2%. Este resultado fornece forte evidência contra $H_0$.*

4. Responda as seguintes questões sobre estimação:

(a) Uma variável aleatória tem média desconhecida $\mu$ e variância $\sigma^2$. Retira-se uma amostra aleatória de tamanho $n$, ou seja, $X_1, X_2, \ldots, X_n$. Encontre o valor esperado e variância dos seguintes estimadores para $\mu$:

i. $\hat{\theta}_1 = \frac{3X_1+X_n}{4}$
*Resp:*$E(\hat{\theta}_1) = \mu$ *e* $Var(\hat{\theta}_1) = \frac{5\sigma^2}{8}$

ii. $\hat{\theta}_2 = \sum_{i=1}^{n} \frac{X_i}{n}$
*Resp:*$E(\hat{\theta}_2) = \mu$ *e* $Var(\hat{\theta}_2) = \frac{\sigma^2}{n}$

(b) Classifique os dois estimadores propostos no item a) em relação ao vício, consistência e eficiência.
*Resp: $\hat{\theta}_1$ é não viciado, não consistente e menos eficiente do que $\hat{\theta}_2$ Resp: $\hat{\theta}_2$ é não viciado, consistente e mais eficiente do que $\hat{\theta}_1$*

(c) Após observar uma amostra aleatória de tamanho 9 da variável X, ou seja, ao observar $X_1, X_2, \ldots, X_9$, foi obtida a sequência de valores `6; 7 ; 5 ; 8 ; 5 ; 9 ; 5 ; 4 ;9`. Encontre as estimativas pelo uso dos estimadores $\hat{\theta}_1$ e $\hat{\theta}_2$ apresentados nos item a.1) e a.2), respectivamente
*Resp: 6,75 e 6,4*

(d) Caso a variável X tenha distribuição normal, qual será a distribuição dos estimadores $\hat{\theta}_1$ e $\hat{\theta}_2$ ?
*Resp: terão distribuição normal com média e variância encontradas no item a)*