Probabilidade e Estatística, 2010/2

CCT - UDESC

Prof. Fernando Deeke Sasse

Prova 4

1. A contaminação de nuvens por por nitrato de prata para produzir chuva. A queda de chuva (mm/min) de 21 nuvens que foram selecionadas aleatoriamente e contaminadas com nitrato de prata é descrita pela lista: 18.0, 30.7, 19.8, 27.1, 22.3, 18.8, 31.8, 23.4, 21.2, 27.9, 31.9, 27.1, 25.0, 24.7, 26.9, 21.8, 29.2, 34.8, 26.7, 31.6.

(a) Você pode afirmar que a quantidade média de chuva excede 25 mm/min? Use $\alpha = 0.01$.
(b) Há evidência de que a queda de chuva é normalmente distribuída?
(c) Compute o poder do teste a média verdadeira é 27mm/min.
(d) Estime o tamanho da amostra requerido para detectar uma média verdadeira de 27.5mm/min, se queremos um poder do teste de ao menos 0.9.
(e) Explique como a parte (a) poderia ser respondida construindo um intervalo de confiança unilateral sobre a média.

Solução:

Estabeleçamos a média 25mm/min como um limite não tolerado:

$H_0 : \mu = 25, \quad H_1 : \mu > 25$

(a)

```
> restart:
> with(Statistics):
> L := [18.0, 30.7, 19.8, 27.1, 22.3, 18.8, 31.8, 23.4, 21.2, 27.9, 31.9, 27.1, 25.0, 24.7, 26.9,
      21.8, 29.2, 34.8, 26.7, 31.6]:
> s := StandardDeviation(L)
```

$$s := 4.78476474700800 \qquad (1)$$

```
> μ := 25; n := nops(L); α := 0.01;
```

$$\mu := 25$$
$$n := 20$$
$$\alpha := 0.01 \qquad (2)$$

```
> k := n - 1:
> T := RandomVariable(StudentT(k)):
> tc := Quantile(T, 1-alpha)
```

$$tc := 2.539483164 \qquad (3)$$

Ou seja, o intervalo de aceitação de H_0 na variável t é $(-\infty, 2.54)$. Notemos agora

```
> xobs := Mean(L)
```

$$xobs := 26.03500000 \qquad (4)$$

```
> tobs := evalf((xobs - mu)·sqrt(n)/s)
```

$$tobs := 0.967374773291495 \qquad (5)$$

Como esta valor está no interior do intervalo de aceitação, a hipótese nula é aceita e não podemos afirmar que a quantidade de chuva excede 25mm/min.

Equivalentemente, poderíamos determinar o intervalo de aceitação nas variáveis não padronizadas:

> $XC := evalf\left(\frac{tc \cdot s}{\text{sqrt}(n)}\right)$

$$XC := 2.71700807808130 \tag{6}$$

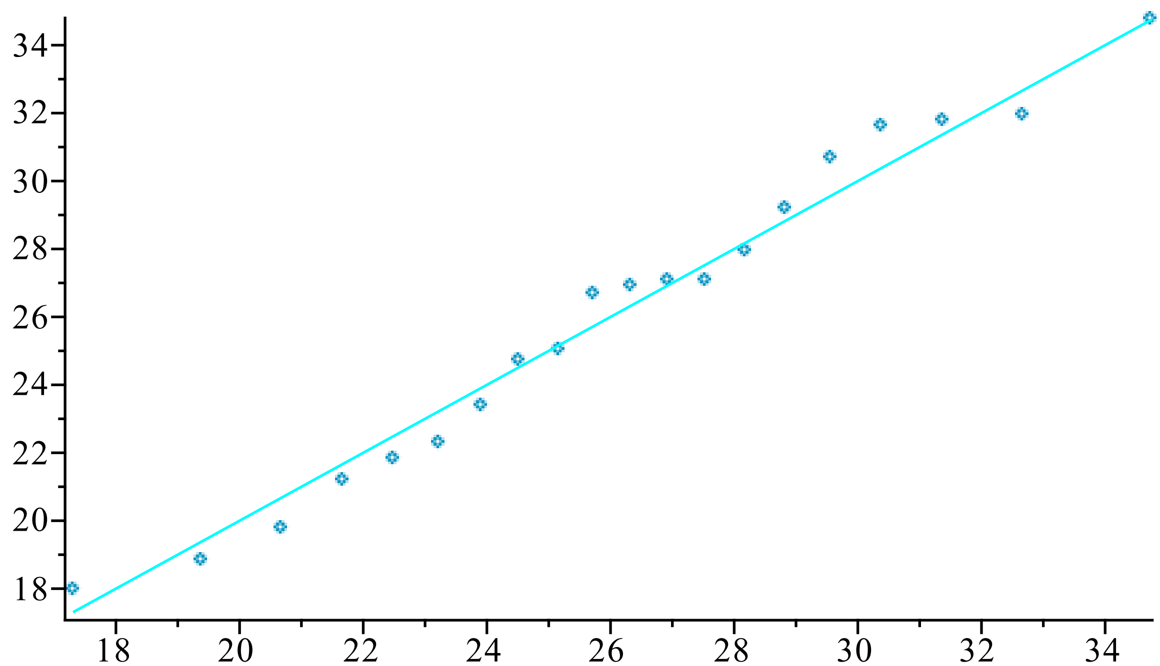
> $L2 := \mu + XC$

$$L2 := 27.7170080780813 \tag{7}$$

O intervalo de aceitação agora é (-∞, 27.72). Como o valor observado está no interior deste intervalo, a hipótese nula é aceita.

(b)

> $NormalPlot(L)$



A disposição dos pontos no gráfico corrobora a suposição de que a população é normal.

(c)

> $mu2 := 27 :$

> $t2 := evalf\left(\frac{(mu2 - \text{mu}) \cdot \text{sqrt}(n)}{s}\right)$

$$t2 := 1.86932323341352 \tag{8}$$

```
> T2 := RandomVariable(NonCentralStudentT(n-1, t2));
```

$$T2 := \_R1 \tag{9}$$

> $\text{beta} := CDF(T2, tc)$

$$\beta := 0.721916155768221213 \tag{10}$$

> $1 - \text{beta}$

$$0.2780838442 \tag{11}$$

(d)

```
> mu3 := 27.5 : n := 52 : beta := 1 :
> while 1 - beta < 0.9 do
  t3 := evalf((mu3 - mu)·sqrt(n) / s);
  T3 := RandomVariable(NonCentralStudentT(n - 1, t3));
  beta := CDF(T3, tc);
  n := n + 1 :
  end do:
  n - 1;
```

$$55 \tag{12}$$

Portanto, devemos ter $n = 55$, ao menos.

(e) Determinemos um intervalo de confiança de 99% limitado superiormente, sobre a média observada:

```
> mu :='mu': n := nops(L) :
> T := RandomVariable(StudentT(n - 1)) :
> tc := Quantile(T, 1 - alpha)
```

$$tc := 2.539483164 \tag{13}$$

```
> xc := evalf(xobs + tc·s / sqrt(n))
```

$$xc := 28.7520080780813 \tag{14}$$

O intervalo de confiança limitado inferiormente para a média observada é então $(-\infty, 28.75)$. Como $\mu = 25$ está dentro deste intervalo, a hipótese nula é aceita.

Solução usando a distribuição normal, supondo σ conhecido

Suponhamos que sabemos que $\sigma = 4.78$.

(a)

$H_0 : \mu = 25, \quad H_1 : \mu > 25$

```
> restart
> with(Statistics) :
> L := [18.0, 30.7, 19.8, 27.1, 22.3, 18.8, 31.8, 23.4, 21.2, 27.9, 31.9, 27.1, 25.0, 24.7, 26.9,
      21.8, 29.2, 34.8, 26.7, 31.6] :
> xobs := Mean(L)
```

$$xobs := 26.03500000 \tag{1.1}$$

```
> μ := 25; n := nops(L); α := 0.01; sigma := 4.78 :
```

$$\mu := 25$$

$$n := 20$$

$$\alpha := 0.01 \tag{1.2}$$

```
> X := RandomVariable(Normal(μ, σ / sqrt(n))) :
> xc := Quantile(X, 1 - alpha)
```

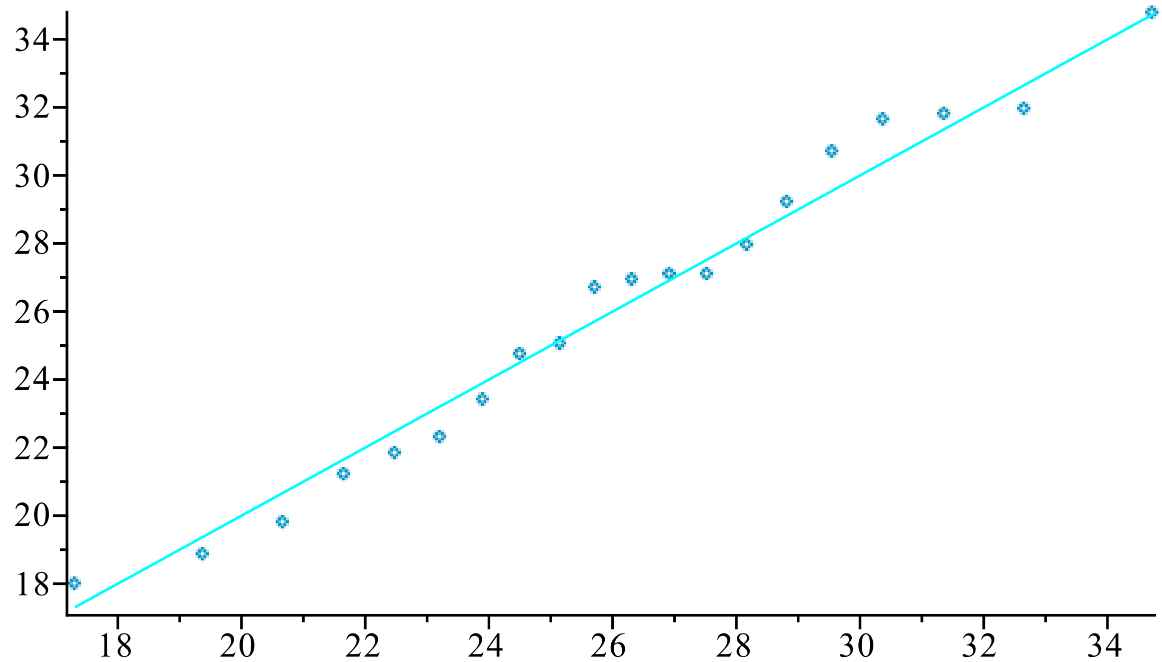
(1.3)

$$xc := 27.48649481 \quad (1.3)$$

Ou seja, o intervalo de aceitação de H_0 é (-∞, 27.50). Como $xobs := 26.035$ está no interior do intervalo de aceitação, a hipótese nula é aceita e não podemos afirmar que a quantidade de chuva excede 25mm/min.

(b)

> $NormalPlot(L)$



A disposição dos pontos no gráfico corrobora a suposição de que a população é normal.

(c)

> $mu2 := 27 :$

> $X2 := RandomVariable\left(Normal\left(\mu 2, \frac{\sigma}{\mathrm{sqrt}(n)}\right)\right) :$

> $\text{beta} := CDF(X2, xc)$

$$\beta := 0.675503393838038702 \quad (1.4)$$

> $1 - \text{beta}$

$$0.3244966062 \quad (1.5)$$

Portanto, o poder do teste é 0.32.

(d)

> $mu3 := 27.6 :$

> $\text{delta} := \text{mu} - mu3 : \text{beta} := 0.1 :$

> $Z := RandomVariable(Normal(0, 1)) :$

> $z1 := Quantile\left(Z, 1 - \frac{\text{alpha}}{2}\right) : z2 := Quantile(Z, 1 - \text{beta}) :$

> $n := \dfrac{(z1 + z2)\cdot\cdot 2\cdot \text{sigma}\cdot\cdot 2}{\text{delta}\cdot\cdot 2}$

$$n := 50.29144822 \qquad \textbf{(1.6)}$$

Portanto, necessitamos ao menos 51 elementos na amostra.

(e) Determinemos um intervalo de confiança sobre a média:

> $X := RandomVariable\left(Normal\left(xobs, \dfrac{\sigma}{\text{sqrt}(n)}\right)\right):$

> $xc := Quantile(X, 1 - \text{alpha})$

$$xc := 27.60303403 \qquad \textbf{(1.7)}$$

Ou seja, o intervalo de confiança superiormente limitado de 99% sobre a média observada é $(-\infty, 27.60)$. Como $\mu = 25$ está no interior deste intervalo, a hipótese nula é aceita.

2. A vida em horas de uma bateria é aproximadamente normalmente distribuída, com desvio-padrão $\sigma = 1.25$ *horas*. Uma amostra aleatória de 10 baterias tem uma vida média de 40.5h.
(a) Há evidências que suportem a afirmação de que vida da bateria excede 40h? Use $\alpha = 0.05$.
(b) Qual é o valor P do teste na parte (a)?
(c) Qual é o erro β para o teste na parte (a) se a média verdadeira é 42?
(d) Qual é o tamanho requerido da amostra para assegurar que β não excede 0.10 se a média verdadeira é 44h?
(e) Explique como você poderia responder a questão na parte (a) calculando um intervalo de confiança adequado.
Solução
Estabeleçamos a média 40h como um limite não tolerado:
$H_0 : \mu = 40, \quad H_1 : \mu > 40$
(a)

> $restart:$

> $with(Statistics):$

> $\mu := 40; \sigma := 1.25; n := 10; \alpha := 0.05;$

$$\mu := 40$$
$$\sigma := 1.25$$
$$n := 10$$
$$\alpha := 0.05 \qquad \textbf{(15)}$$

> $X := RandomVariable\left(Normal\left(\mu, \dfrac{\sigma}{\text{sqrt}(n)}\right)\right):$

> $xc := Quantile(X, 1 - \alpha);$

$$xc := 40.65018548 \qquad \textbf{(16)}$$

O intervalo de aceitação é $(-\infty, 40.65)$. Como a média observada 40.5 está dentro deste intervalo, a hipótese nula é consistente com a observação. Ou seja, não há evidências que suportem a afirmação de que vida da bateria excede 40h.
(b)

> $xobs := 40.5:$

> $(1 - evalf(CDF(X, \text{xobs})))$

$$0.1029516054 \tag{17}$$

Ou seja, o valor-P do teste é 0.1029.

(c)

> $X := RandomVariable\left(Normal\left(42, \frac{\sigma}{\text{sqrt}(n)}\right)\right) :$

> $\beta := evalf(CDF(X, xc))$

$$\beta := 0.0003191552664 \tag{18}$$

(d)

> delta := 4 : beta := 0.1 :

> $Z := RandomVariable(Normal(0, 1)) :$

> $z1 := Quantile\left(Z, 1 - \frac{\text{alpha}}{2}\right) : z2 := Quantile(Z, 1 - \text{beta}) :$

> $n := \frac{(z1 + z2)\cdot\cdot 2\cdot\text{sigma}\cdot\cdot 2}{\text{delta}\cdot\cdot 2}$

$$n := 1.026115534 \tag{19}$$

Ou seja, 2 elementos.

(e)

> $X2 := RandomVariable\left(Normal\left(xobs, \frac{\sigma}{\text{sqrt}(n)}\right)\right) :$

> $xc2 := Quantile(X2, \alpha);$

$$xc2 := 38.47026594 \tag{20}$$

O intervalo de confiança para a média observada, limitado inferiormente, é $(38.47, +\infty)$. Como $\mu = 40$ está dentro deste intervalo, a hipótese nula é aceita, de modo que não há evidências que suportem a afirmação de que vida da bateria excede 40h.

3. Um fabricante de baterias afirma que a vida média delas é 48h de operação, se os procedimentos corretos de carregamentos forem seguidos. Um estudo de 5000 baterias é feito e 15 param antes de 48h. Estes resultados experimentais suportam a afirmação de que menos de 0.2% das baterias fabricadas irão falhar durante 48h de operação? Faça um teste de hipóteses com $\alpha = 0.01$.

Solução

Hipóteses:

$H_0 : p = 0.002, \quad H_1 : \text{p} < 0.002$

> $restart :$

> $with(Statistics) :$

> $Z := RandomVariable(Normal(0, 1)) :$

> $zc := Quantile(Z, 0.01)$

$$zc := -2.326347874 \tag{21}$$

O intervalo de aceitação de H_0 é $(-2.33, +\infty)$, Se $z0$ estiver neste intervalo, H_0 é aceita.

> $p0 := 0.002 : n := 5000 :$

> $z0 := \dfrac{(15 - n \cdot p0)}{\mathrm{sqrt}(n \cdot p0 \cdot (1 - p0))}$

$$z0 := 1.582722344 \quad \textbf{(22)}$$

Como $z0$ de fato está no intervalo, a hipótese nula é aceita, e os dados experimentais não suportam a afirmação de que menos de 0.2% das baterias fabricadas irão falhar durante 48h de operação.

4. O conteúdo de açúcar (em mg) em pêssegos enlatados é normalmente distribuído. Uma amostra de 10 latas é selecionada e é medido um desvio-padrão amostral de $s = 4.8\ mg$.

(a) Determine um intervalo de confiança de 95% para σ.

(b) Teste a hipótese de que a variância é $\sigma^2 = 18\ mg^2$ contra a hipótese $\sigma^2 \neq 18\ mg^2$. Use $\alpha = 0.05$.

(c) Qual é o valor P do teste?

(d) Resolva a parte (b) usando intervalo de confiança bilateral de 95% para σ.

Solução

(a)

> $restart:$

> $with(Statistics):$

> $n := 10 : \mathrm{alpha} := 0.05 : s := 4.8 :$

> $X := RandomVariable(ChiSquare(n - 1));$

$$X := \_R \quad \textbf{(23)}$$

> $x2 := Quantile\left(X, 1 - \dfrac{\mathrm{alpha}}{2}\right); x1 := Quantile\left(X, \dfrac{\mathrm{alpha}}{2}\right);$

$$x2 := 19.02276775$$
$$x1 := 2.700389725 \quad \textbf{(24)}$$

> $s1 := \dfrac{(n - 1) \cdot s \cdot \cdot 2}{x1}; s2 := \dfrac{(n - 1) \cdot s \cdot \cdot 2}{x2};$

$$s1 := 76.78891609$$
$$s2 := 10.90062197 \quad \textbf{(25)}$$

>

O intervalo de confiança de 95% sobre a variância observada é, portanto, (10.90,76.79).

(b) $H_0 : \sigma^2 = 18,\ H_0 : \sigma^2 \neq 18$

> $sigma02 := 18$

$$\sigma 02 := 18 \quad \textbf{(26)}$$

Teste estatístico:

> $X02 := \dfrac{(n - 1) \cdot s \cdot \cdot 2}{sigma02}$

$$X02 := 11.52000000 \quad \textbf{(27)}$$

Como este valor está contido no intervalo $(x1, x2) = (2.70, 19.02)$, a hipótese nula é aceita.

(c)

> $2 \cdot (1 - CDF(X, X02))$

$$0.483481764 \quad \textbf{(28)}$$

>

(d) Como $s^2 = 23.04$ está dentro do intervalo de confiança de 95% para a média amostral (10.90,76.79), a hipótese nula é aceita. Ou seja, $\sigma^2 = 18$.