Contrastos d'hipòtesis: Una mostra

Contrastos d'hipòtesis d'una mostra

Anem a estudiar els principals contrastos per a una mostra

Un cop fixades les hipòtesi nul·la i alternativa, hem d'estudiar les condicions de la mostra per saber el test que podem aplicar

Teniu una Taula de Contrastos exhaustiva a Campus Extens, aquí nomes veurem alguns dels més comuns

Contrast per a μ de normal amb σ coneguda: Z-test

Condicions: m.a.s. de mida n d'una població $N(\mu, \sigma)$ amb σ coneguda

Empram l'estadístic de contrast

$$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Calculam el seu valor z_0 sobre la mostra

Contrast per a μ de normal amb σ coneguda: Z-test

Condicions: m.a.s. de mida n d'una població $N(\mu, \sigma)$ amb σ coneguda

Empram l'estadístic de contrast

$$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Calculam el seu valor z_0 sobre la mostra

Contrast: $\left\{ \begin{array}{l} H_0 : \mu = \mu_0 \quad (\text{ o } H_0 : \mu \leqslant \mu_0) \\ H_1 : \mu > \mu_0 \end{array} \right.$

p-valor: $P(Z \geqslant z_0)$

Contrast per a μ de normal amb σ coneguda: Z-test

Condicions: m.a.s. de mida n d'una població $N(\mu,\sigma)$ amb σ coneguda

Empram l'estadístic de contrast

$$Z=\frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Calculam el seu valor z_0 sobre la mostra

Contrast: $\left\{\begin{array}{l} H_0:\mu=\mu_0 \quad (\text{ o } H_0:\mu\geqslant\mu_0) \\ H_1:\mu<\mu_0\end{array}\right.$

p-valor: $P(Z\leqslant z_0)$

Contrast per a μ de normal amb σ coneguda: Z-test

Condicions: m.a.s. de mida n d'una població $N(\mu, \sigma)$ amb σ coneguda

Empram l'estadístic de contrast

$$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Calculam el seu valor z_0 sobre la mostra

Contrast: $\left\{ \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{array} \right.$

p-valor: $2P(Z \geqslant |z_0|)$

Contrast per a μ quan n és gran: Z-test

Si la mida n de la mostra és gran (posem, si $n \geqslant 40$), podem aplicar les regles anteriors encara que la població no sigui normal

Si, a més, σ és desconeguda, en aquest cas es pot substituir a Z per la desviació típica mostral $\widetilde{S}_X$:

$$Z = \frac{\overline{X} - \mu_0}{\frac{\widetilde{S}_X}{\sqrt{n}}}$$

Mateixos p-valors que al cas anterior

Exemple

Una organització ecologista afirma que el pes mitjà dels individus adults d'una espècie marina ha disminuït dràsticament

Se sap per les dades històriques que el pes mitjà poblacional era de 460 g

Una mostra aleatòria de 40 individus d'aquesta espècie ha donat una mitjana mostral de 420 g i una desviació típica mostral de 119 g

Amb aquestes dades, podem afirmar amb un nivell de significació del 5% que el pes mitjà és inferior a 460 g?

Exemple

(1) Hipòtesis:

$$\begin{cases} H_0 : \mu = 460 \\ H_1 : \mu < 460 \end{cases}$$

Estam en el cas VIII de la Taula I

Exemple

(1) Hipòtesis:

$$\begin{cases} H_0 : \mu = 460 \\ H_1 : \mu < 460 \end{cases}$$

Estam en el cas VIII de la Taula I

(2) $\alpha = 0.05$

Exemple

(1) Hipòtesis:

$$\begin{cases} H_0 : \mu = 460 \\ H_1 : \mu < 460 \end{cases}$$

Estam en el cas VIII de la Taula I

(2) $\alpha = 0.05$

(3) Estadístic: Com que $n = 40$ és gran, emprarem

$$Z = \frac{\overline{X} - \mu_0}{\widetilde{S}_X / \sqrt{n}}$$

Exemple

(1) Hipòtesis:

$$\begin{cases} H_0 : \mu = 460 \\ H_1 : \mu < 460 \end{cases}$$

Estam en el cas VIII de la Taula I

(2) $\alpha = 0.05$

(3) Estadístic: Com que $n = 40$ és gran, emprarem

$$Z = \frac{\overline{X} - \mu_0}{\widetilde{S}_X / \sqrt{n}}$$

(4) Valor: $z = \dfrac{420 - 460}{119/\sqrt{40}} = -2.126$

Exemple

(5) *p*-valor:

$$P(Z \leqslant -2.126) = 0.017$$

Exemple

(5) *p*-valor:

$$P(Z \leqslant -2.126) = 0.017$$

(6) Decisió: Com que $\alpha >$ *p*-valor, rebutjam (al nivell de significació $\alpha = 0.05$) que el pes mitjà és de 460 g (H_0) contra que és menor de 460 g (H_1).

Concloem que el pes mitjà és < 460 g i que per tant ha minvat en els darrers anys

Exemple

Interval de confiança:

$$\left]-\infty, \overline{X} - z_{\alpha} \cdot \frac{\widetilde{S}_X}{\sqrt{n}}\right] =]-\infty, 450.95]$$

Informe: El p-valor d'aquest contrast és 0.017, i l'interval de confiança al nivell de significació $\alpha = 0.05$ per a la mitjana poblacional μ és $]-\infty, 450.95]$. Per tant, hi ha evidència significativa per rebutjar la hipòtesi nul·la en favor que $\mu < 460$.

Contrast per a μ de normal amb σ desconeguda: T-test

Les regles de decisió són similars al cas amb σ coneguda, excepte que substituïm σ per $\widetilde{S}_X$ i empram la distribució t de Student

Recordau que si $X_1, \ldots, X_n$ és una m.a.s. d'una població normal X amb mitjana μ_0,

$$T = \frac{\overline{X} - \mu_0}{\frac{\widetilde{S}_X}{\sqrt{n}}}$$

segueix una distribució t de Student amb $n - 1$ graus de llibertat

Els p-valors es calculen amb aquesta distribució

Contrast de μ de normal amb σ desconeguda: T-test

Condicions: m.a.s. de mida n d'una població $N(\mu, \sigma)$ amb μ i σ desconegudes

Empram l'estadístic de contrast

$$T = \frac{\overline{X} - \mu_0}{\frac{\widetilde{S}_X}{\sqrt{n}}}$$

Calculam el seu valor t_0 sobre la mostra

Contrast de μ de normal amb σ desconeguda: T-test

Condicions: m.a.s. de mida n d'una població $N(\mu, \sigma)$ amb μ i σ desconegudes

Empram l'estadístic de contrast

$$T = \frac{\overline{X} - \mu_0}{\frac{\widetilde{S}_X}{\sqrt{n}}}$$

Calculam el seu valor t_0 sobre la mostra

Contrast: $\begin{cases} H_0 : \mu = \mu_0 & (\text{ o } H_0 : \mu \leqslant \mu_0) \\ H_1 : \mu > \mu_0 \end{cases}$

p-valor: $P(t_{n-1} \geqslant t_0)$

Contrast de μ de normal amb σ desconeguda: T-test

Condicions: m.a.s. de mida n d'una població $N(\mu, \sigma)$ amb μ i σ desconegudes

Empram l'estadístic de contrast

$$T = \frac{\overline{X} - \mu_0}{\frac{\widetilde{S}_X}{\sqrt{n}}}$$

Calculam el seu valor t_0 sobre la mostra

Contrast: $\left\{ \begin{array}{l} H_0 : \mu = \mu_0 \quad (\text{ o } H_0 : \mu \geqslant \mu_0) \\ H_1 : \mu < \mu_0 \end{array} \right.$

p-valor: $P(t_{n-1} \leqslant t_0)$

Contrast de μ de normal amb σ desconeguda: T-test

Condicions: m.a.s. de mida n d'una població $N(\mu, \sigma)$ amb μ i σ desconegudes

Empram l'estadístic de contrast

$$T = \frac{\overline{X} - \mu_0}{\frac{\widetilde{S}_X}{\sqrt{n}}}$$

Calculam el seu valor t_0 sobre la mostra

Contrast: $\left\{ \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{array} \right.$

p-valor: $2P(t_{n-1} \geqslant |t_0|)$

Exemple

S'espera que el nivell de colesterol en plasma d'uns malalts sota un determinant tractament es distribueixi normalment amb mitjana 220 mg/dl

Es pren una mostra de 9 malalts, i es mesuren els seus nivells:

$$203, 229, 215, 220, 223, 233, 208, 228, 209$$

Contrastau la hipòtesi que aquesta mostra efectivament prové d'una població amb mitjana 220 mg/dl

Exemple

(1) Hipòtesis:

$$\begin{cases} H_0 : \mu = 220 \\ H_1 : \mu \neq 220 \end{cases}$$

Estam en el cas IV de la Taula I

Exemple

(1) Hipòtesis:

$$\begin{cases} H_0 : \mu = 220 \\ H_1 : \mu \neq 220 \end{cases}$$

Estam en el cas IV de la Taula I

(2) Estadístic: Sota aquestes condicions (població normal, σ desconeguda, mostra petita de $n = 9$) emprarem

$$T = \frac{\overline{X} - \mu_0}{\widetilde{S}_X / \sqrt{9}}$$

que segueix una distribució t_8

Exemple

(3) Valor: $\mu_0 = 220$, $n = 9$,

```
> colesterol=c(203,229,215,220,223,233,
 208,228,209)
> x.bar=round(mean(colesterol),4)
> x.bar
[1] 218.6667
> s.tilde=round(sd(colesterol),4)
> s.tilde
[1] 10.5238
```

Exemple

(3) Valor: $\mu_0 = 220$, $n = 9$,

```
> colesterol=c(203,229,215,220,223,233,
 208,228,209)
> x.bar=round(mean(colesterol),4)
> x.bar
[1] 218.6667
> s.tilde=round(sd(colesterol),4)
> s.tilde
[1] 10.5238
```

$$t_0 = \frac{218.6667 - 220}{10.5238/\sqrt{9}} = -0.3801$$

Exemple

(4) *p*-valor:

Exemple

(4) *p*-valor:

$$2P(t_8 \geqslant |-0.3801|) = 2P(t_8 \geqslant 0.3801) = 0.7138$$

Exemple

(4) *p*-valor:

$$2P(t_8 \geqslant |-0.3801|) = 2P(t_8 \geqslant 0.3801) = 0.7138$$

Amb la taula:

$$\begin{aligned} 2P(t_8 \geqslant 0.3801) &> 2P(t_8 \geqslant 1.3968) \\ &= 2(1 - P(t_8 \leqslant 1.3968)) = 2(1-0.9) = 0.2 \end{aligned}$$

Exemple

(4) *p*-valor:

$$2P(t_8 \geqslant |-0.3801|) = 2P(t_8 \geqslant 0.3801) = 0.7138$$

Amb la taula:

$$\begin{aligned} 2P(t_8 \geqslant 0.3801) &> 2P(t_8 \geqslant 1.3968) \\ &= 2(1 - P(t_8 \leqslant 1.3968)) = 2(1-0.9) = 0.2 \end{aligned}$$

(5) Decisió: Com que el *p*-valor és molt gran, no podem rebutjar que el nivell mitjà de colesterol en plasma sigui igual a 220 mg/dl

Acceptam que el nivell de colesterol en plasma en aquesta població té mitjana 220 mg/dl

Exemple

Interval de confiança al 95%:

Exemple

Interval de confiança al 95%:

$$\left[\overline{X} - t_{8,0.975}\frac{\widetilde{S}_X}{\sqrt{n}},\ \overline{X} + t_{8,0.975}\frac{\widetilde{S}_X}{\sqrt{n}}\right] = [210.577, 226.756]$$

Informe: El p-valor d'aquest contrast és 0.7138, i l'interval de confiança del 95% per al nivell mitjà de colesterol μ és $[210.577, 226.756]$. Per tant, no hi ha evidència que ens permeti rebutjar que $\mu = 220$.

Amb R

El T-test està implementat en R en la funció `t.test`.

$$\begin{cases} H_0 : \mu = 220 \\ H_1 : \mu \neq 220 \end{cases}$$

```
> t.test(colesterol,mu=220,alternative="two.sided",
  conf.level=0.95)

        One Sample t-test
data:  colesterol
t = -0.3801, df = 8, p-value = 0.7138
alternative hypothesis: true mean is not equal
   to 220
95 percent confidence interval:
 210.5774 226.7560
sample estimates:
mean of x
 218.6667
```

Z-test contra T-test

En el cas d'una població amb σ desconeguda:

- Si la mostra és petita i la població és normal, hem d'emprar el T-test
- Si la mostra és gran i la població qualsevol, podem emprar el Z-test
- Si la mostra és gran i la població és normal, els podem emprar tots dos
- En aquest darrer cas, us recomanam que empreu el T-test (que és més exacte)

Contrastos per a p: binom-test

Condicions: m.a.s. de mida n d'una població $Be(p)$

Obtenim x_0 èxits, de manera que $\widehat{p}_X = x_0/n$

Considerem un contrast amb $H_0 : p = p_0$

Si H_0 és vertadera, el nombre d'èxits segueix una distribució $B(n, p_0)$

Contrastos per a p: binom-test

Condicions: m.a.s. de mida n d'una població $Be(p)$

Obtenim x_0 èxits, de manera que $\widehat{p}_X = x_0/n$

Considerem un contrast amb $H_0 : p = p_0$

Si H_0 és vertadera, el nombre d'èxits segueix una distribució $B(n, p_0)$

Contrast: $\left\{\begin{array}{l} H_0 : p = p_0 \quad (\text{ o } H_0 : p \leqslant p_0) \\ H_1 : p > p_0 \end{array}\right.$

p-valor: $P(B(n, p_0) \geqslant x_0)$

Contrastos per a p: binom-test

Condicions: m.a.s. de mida n d'una població $Be(p)$

Obtenim x_0 èxits, de manera que $\widehat{p}_X = x_0/n$

Considerem un contrast amb $H_0 : p = p_0$

Si H_0 és vertadera, el nombre d'èxits segueix una distribució $B(n, p_0)$

Contrast: $\left\{\begin{array}{l} H_0 : p = p_0 \quad (\text{ o } H_0 : p \geqslant p_0) \\ H_1 : p < p_0 \end{array}\right.$

p-valor: $P(B(n, p_0) \leqslant x_0)$

Contrastos per a p: binom-test

Condicions: m.a.s. de mida n d'una població $Be(p)$

Obtenim x_0 èxits, de manera que $\widehat{p}_X = x_0/n$

Considerem un contrast amb $H_0 : p = p_0$

Si H_0 és vertadera, el nombre d'èxits segueix una distribució $B(n, p_0)$

Contrast: $\left\{ \begin{array}{l} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{array} \right.$

p-valor: $2 \min\{P(B(n, p_0) \leqslant x_0), P(B(n, p_0) \geqslant x_0)\}$

Exemple

Tenim un test per detectar un determinat microorganisme. En una mostra de 25 cultius amb aquest microorganisme, el test el detectà en 21 casos. Hi ha evidència que la sensibilitat del test sigui superior al 80%?

Exemple

Tenim un test per detectar un determinat microorganisme. En una mostra de 25 cultius amb aquest microorganisme, el test el detectà en 21 casos. Hi ha evidència que la sensibilitat del test sigui superior al 80%?

(1) Hipòtesis:

$$\begin{cases} H_0 : p = 0.8 \\ H_1 : p > 0.8 \end{cases}$$

Exemple

Tenim un test per detectar un determinat microorganisme. En una mostra de 25 cultius amb aquest microorganisme, el test el detectà en 21 casos. Hi ha evidència que la sensibilitat del test sigui superior al 80%?

(1) Hipòtesis:

$$\begin{cases} H_0 : p = 0.8 \\ H_1 : p > 0.8 \end{cases}$$

(2) Estadístic: Emprarem el nombre d'èxits, que sota H_0 és $B(25, 0.8)$

Exemple

(3) Valor: $x_0 = 21$

(4) p-valor:

$$\begin{aligned} P(B(25, 0.8) \geqslant 21) &= \texttt{1-pbinom(20,25,0.8)} \\ &= 0.421 \end{aligned}$$

Exemple

(3) Valor: $x_0 = 21$

(4) p-valor:

$$P(B(25, 0.8) \geqslant 21) = \texttt{1-pbinom(20,25,0.8)} = 0.421$$

(5) Decisió: Com que el p-valor és molt gran, no podem rebutjar la hipòtesi nul·la

No hi ha evidència que la sensibilitat del test sigui superior al 80%

Exemple

Alerta als p-valors dels C.H. bilaterals no simètrics

$n = 25$, $x_0 = 21$

$$\left\{ \begin{array}{l} H_0 : p = 0.8 \\ H_1 : p > 0.8 \end{array} \right.$$

El p-valor és

$$P(B(25, 0.8) \geqslant 21) = \texttt{1-pbinom(20,25,0.8)} = 0.421$$

Exemple

Alerta als p-valors dels C.H. bilaterals no simètrics

$n = 25$, $x_0 = 21$

$$\begin{cases} H_0 : p = 0.8 \\ H_1 : p < 0.8 \end{cases}$$

El p-valor és

$$P(B(25, 0.8) \leqslant 21) = \texttt{pbinom(21,25,0.8)} = 0.766$$

Exemple

Alerta als p-valors dels C.H. bilaterals no simètrics

$n = 25,\ x_0 = 21$

$$\left\{\begin{array}{l} H_0 : p = 0.8 \\ H_1 : p \neq 0.8 \end{array}\right.$$

El p-valor és:

$$\begin{aligned} 2 \cdot P(B(25, 0.8) \geqslant 21) &= 2 \cdot 0.421 = 0.842 \\ 2 \cdot P(B(25, 0.8) \leqslant 21) &= 2 \cdot 0.766 = 1.532 \end{aligned}$$

Prenem com a p-valor el més petit (l'únic $\leqslant 1$): 0.842

Amb R

El binom-test està implementat en R en la funció `binom.test`.

```
> binom.test(21,25,p=0.8,alternative="greater",
   conf.level=0.95)

     Exact binomial test
data:  21 and 25
number of successes = 21, number of trials = 25,
  p-value = 0.4207
alternative hypothesis: true probability of success
   is greater than 0.8
95 percent confidence interval:
 0.6703917 1.0000000
sample estimates:
probability of success
                  0.84
```

Contrastos per a p quan n és gran: prop-test

Si indicam amb p la proporció poblacional i $\widehat{p}_X$ la proporció mostral, sabem que si la mostra és gran ($n \geqslant 40$)

$$Z = \frac{\widehat{p}_X - p}{\sqrt{\frac{p(1-p)}{n}}} \approx N(0,1)$$

Si la hipòtesi nul·la $H_0 : p = p_0$ és vertadera,

$$Z = \frac{\widehat{p}_X - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \approx N(0,1)$$

Mateixos p-valors que al Z-test. S'ha d'anar alerta amb l'interval de confiança. Si tenim $n \geqslant 100$, $n\hat{p}_X \geqslant 10$ i $n(1-\hat{p}_X) \geqslant 10$, es pot emprar el de Laplace. En cas contrari, s'ha d'emprar el de Wilson.

Exemple

Una associació ramadera afirma que, a les matances casolanes a les Balears, com a mínim el 70% dels porcs han estat analitzats de triquinosi

En una investigació, es visita una mostra aleatòria de 100 matances i resulta que en 53 d'aquestes s'ha realitzat l'anàlisi de triquinosi

Hem d'acceptar l'afirmació dels ramaders?

Exemple

(1) Contrast:

$$\begin{cases} H_0 : p \geqslant 0.7 \\ H_1 : p < 0.7 \end{cases}$$

(2) Estadístic: Sota aquestes condicions podem emprar

$$Z = \frac{\widehat{p}_X - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Estam en el cas XI de la Taula I

Exemple

(1) Contrast:

$$\begin{cases} H_0 : p \geqslant 0.7 \\ H_1 : p < 0.7 \end{cases}$$

(2) Estadístic: Sota aquestes condicions podem emprar

$$Z = \frac{\widehat{p}_X - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Estam en el cas XI de la Taula I

(3) Valor:

$$\widehat{p}_X = \frac{53}{100} = 0.53 \Longrightarrow z_0 = \frac{0.53 - 0.7}{\sqrt{\frac{0.7 \cdot 0.3}{100}}} = -3.71$$

Exemple

(4) *p*-valor:

$$P(Z \leqslant -3.71) = 0.0001$$

(5) Decisió: El *p*-valor és molt petit, per tant rebutjam la hipòtesi nul·la en favor de l'alternativa. Podem afirmar amb contundència que l'afirmació dels ramaders és falsa

Exemple

Interval de confiança al 95%:

Exemple

Interval de confiança al 95%:(estam en les condicions necessàries)

$$\left]-\infty, \widehat{p}_X - z_{0.05}\sqrt{\frac{\widehat{p}_X(1-\widehat{p}_X)}{n}}\right] =]-\infty, 0.6121]$$

Informe: El p-valor d'aquest contrast és 0.0001, i l'interval de confiança del 95% per a la proporció p de matances on s'han fet anàlisis de triquinosi és $]-\infty, 0.6121]$. Per tant, hi ha evidència molt significativa per rebutjar que $p = 0.7$.

Amb R

El prop-test està (millor) implementat en R en la funció `prop.test`.

```
> prop.test(53,100,p=0.7,alternative="less",
  conf.level=0.95)

      1-sample proportions test with continuity
      correction

data:  53 out of 100, null probability 0.7
X-squared = 12.9643, df = 1, p-value = 0.0001587
alternative hypothesis: true p is less than 0.7
95 percent confidence interval:
 0.0000000 0.6150364
sample estimates:
   p
0.53
```

Amb R

El prop-test està (millor) implementat en R en la funció `prop.test`.

```
> prop.test(53,100,p=0.7,alternative="less",
  conf.level=0.95, correct=FALSE)

      1-sample proportions test without
      continuity correction

data:  53 out of 100, null probability 0.7
X-squared = 13.7619, df = 1, p-value = 0.0001038
alternative hypothesis: true p is less than 0.7
95 percent confidence interval:
 0.0000000 0.6102196
sample estimates:
   p
0.53
```

Contrast per a σ de normal: χ^2-test

Recordem que si $X_1, \ldots, X_n$ és una m.a.s. d'una v.a. $X \sim N(\mu, \sigma)$, aleshores l'estadístic

$$\chi^2_{n-1} = \frac{(n-1)\widetilde{S}_X^2}{\sigma^2}$$

segueix una distribució χ^2 amb $n-1$ graus de llibertat

Per tant, si la hipòtesi nul·la $H_0 : \sigma = \sigma_0$ és vertadera,

$$\chi^2_{n-1} = \frac{(n-1)\widetilde{S}_X^2}{\sigma_0^2}$$

tindrà una distribució χ^2 amb $n-1$ graus de llibertat

Contrast per a σ de normal: χ^2-test

Condicions: m.a.s. de mida n d'una població $N(\mu, \sigma)$ amb σ desconeguda

Empram l'estadístic de contrast

$$\chi^2_{n-1} = \frac{(n-1)\widetilde{S}_X^2}{\sigma_0^2}$$

Calculam el seu valor χ^2_0 sobre la mostra

Contrast per a σ de normal: χ^2-test

Condicions: m.a.s. de mida n d'una població $N(\mu, \sigma)$ amb σ desconeguda

Empram l'estadístic de contrast

$$\chi^2_{n-1} = \frac{(n-1)\widetilde{S}_X^2}{\sigma_0^2}$$

Calculam el seu valor χ_0^2 sobre la mostra

Contrast: $\left\{ \begin{array}{l} H_0 : \sigma = \sigma_0 \quad (\text{ o } H_0 : \sigma \leqslant \sigma_0) \\ H_1 : \sigma > \sigma_0 \end{array} \right.$

p-valor: $P(\chi^2_{n-1} \geqslant \chi_0^2)$

Contrast per a σ de normal: χ^2-test

Condicions: m.a.s. de mida n d'una població $N(\mu, \sigma)$ amb σ desconeguda

Empram l'estadístic de contrast

$$\chi^2_{n-1} = \frac{(n-1)\widetilde{S}_X^2}{\sigma_0^2}$$

Calculam el seu valor χ^2_0 sobre la mostra

Contrast: $\left\{ \begin{array}{l} H_0 : \sigma = \sigma_0 \quad (\text{ o } H_0 : \sigma \geqslant \sigma_0) \\ H_1 : \sigma < \sigma_0 \end{array} \right.$

p-valor: $P(\chi^2_{n-1} \leqslant \chi^2_0)$

Contrast per a σ de normal: χ^2-test

Condicions: m.a.s. de mida n d'una població $N(\mu, \sigma)$ amb σ desconeguda

Empram l'estadístic de contrast

$$\chi^2_{n-1} = \frac{(n-1)\widetilde{S}_X^2}{\sigma_0^2}$$

Calculam el seu valor χ_0^2 sobre la mostra

Contrast: $\left\{ \begin{array}{l} H_0 : \sigma = \sigma_0 \\ H_1 : \sigma \neq \sigma_0 \end{array} \right.$

p-valor: $2\min\left\{P(\chi^2_{n-1} \leqslant \chi_0^2), P(\chi^2_{n-1} \geqslant \chi_0^2)\right\}$

Exemple

S'han mesurat els següents valors en milers de persones per a l'audiència d'un programa de ràdio en $n = 10$ dies:

$$521, 742, 593, 635, 788, 717, 606, 639, 666, 624$$

Contrastau si la variància de l'audiència és 6400 al nivell de significació del 5%, suposant que la població és normal

Exemple

(1) Hipòtesis:

$$\begin{cases} H_0 : \sigma = \sqrt{6400} = 80 \\ H_1 : \sigma \neq 80 \end{cases}$$

(no s'especifica quina alternativa es demana)

(2) $\alpha = 0.05$

(3) Estadístic: Sota aquestes condicions podem emprar

$$\chi^2_{n-1} = \frac{(n-1)\widetilde{S}_X^2}{\sigma_0^2}$$

Exemple

(4) Valor:

```
> x=c(521,742,593,635,788,717,
 606,639,666,624)
> var(x)
[1] 6111.656
```

$$\chi_0^2 = \frac{9 \cdot 6111.656}{6400} = 8.5945$$

(5) *p*-valor:

$$\begin{aligned} 2 \cdot P(\chi_9^2 \geqslant 8.5945) &= 2 \cdot 0.4755 = 0.951 \\ 2 \cdot P(\chi_9^2 \leqslant 8.5945) &= 2 \cdot 0.5245 = 1.049 \end{aligned}$$

Prenem com a *p*-valor el més petit: 0.951

Exemple

(4) Valor:

```
> x=c(521,742,593,635,788,717,
 606,639,666,624)
> var(x)
[1] 6111.656
```

$$\chi_0^2 = \frac{9 \cdot 6111.656}{6400} = 8.5945$$

(5) p-valor: 0.951

(6) Decisió: No podem rebutjar la hipòtesi que la variància sigui 6400 al nivell de significació del 5%

Exemple

Interval de confiança del 95% (Taula I, cas XIII):

$$\left[\frac{(n-1)\widetilde{S}_X^2}{\chi^2_{n-1,0.975}}, \frac{(n-1)\widetilde{S}_X^2}{\chi^2_{n-1,0.025}}\right] = [2891.53, 20369.25]$$

Informe: El p-valor d'aquest contrast és 0.951, i l'interval de confiança del 95% per a la variància σ^2 de l'audiència és $[2891.53, 20369.25]$. Per tant, no hi ha evidència que ens permeti rebutjar que $\sigma^2 = 6400$.

Amb R

El χ^2-test està implementat en R en la funció `sigma.test` del paquet `TeachingDemos`

```
> library("TeachingDemos")
> sigma.test(x,sigma=80,alternative="two.sided",
conf.level=0.95)

     One sample Chi-squared test for variance
data:  x
X-squared = 8.5945, df = 9, p-value = 0.951
alternative hypothesis: true variance is not equal
 to 6400
95 percent confidence interval:
  2891.53 20369.25
sample estimates:
var of x
6111.656
```