TYPTAL

χ^2 - test

Det finns två olika varianter av dessa uppgifter:

Fackindelning med k fack

(I) Gör en tabell med N_i, np_i, i, det som är relevant

$$n = N_1 + N_2 + \ldots + N_n$$

(II) Teststorhet

Använd $T = \sum_{i=1}^{k} \frac{(N_i - np_i)^2}{np_i} \approx \chi^2(k-1)$

(III) $T > C = \chi^2(k-1)$

om $T > C \implies$ Förkasta H_0

Homogenitetstest

(I) Gör en tabell

	$\hat{p}_1 =$	$\hat{p}_2 =$		
$n_1\hat{p}_j$	N_{11}	N_{12}	...	$n_1 =$
$n_2\hat{p}_j$	N_{21}	N_{22}	...	$n_2 =$
	$N_1 =$	$N_2 =$		$n =$

(II) Teststorhet

Använd $T = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(N_{ij} - n_i\hat{p}_j)^2}{n_i\hat{p}_j} \approx \chi^2((r-1)(k-1))$

(III) $T > C = \chi^2((r-1)(k-1))$

om $T > C \implies$ Förkasta H_0 !

OBS: i uppgifterna står det "på nivån ... ", tex nivån 5%
då ska man använda $\chi^2_{0,95}$

Flera stickprov, parvisa mätningar

Det finns många olika varianter på frågor de kan ställa.

- rimligt att anta $\sigma_1 = \ldots = \sigma_4$? Motivera med test.

(I) $H_0: \sigma_1 = \sigma_4 \quad H_1: \sigma_1 \neq \sigma_4$

ställer hypoteserna mot varandra

(II) Välj teststorhet

$$V = \frac{S_1^2}{S_4^2} \sim F(m-1, n-1)$$

(III) $V > c$ eller $V < \frac{1}{c}$

där $c = F(m-1, n-1)$

om $V > c$ el. $V < \frac{1}{c}$ $\Rightarrow$ Förkasta H_0!

- Om de frågar om "någon är bättre" alt. "vilken är bäst"

(I) Hjälpvariabel

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S\sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t(m+n-2)$$

(II) Bilda $I_{\mu_1 - \mu_2}$

$$I_{\mu_1 - \mu_2} = \left(\bar{x} - \bar{y} \mp t(m+n-2) \cdot S\sqrt{\frac{1}{m} + \frac{1}{n}}\right)$$

där $$S^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{(m-1) + (n-1)}$$

(III) Slutsatser:

1. Intervallet är negativt
   $\Rightarrow$ 2 är bättre än 1
2. Intervallet är positivt
   $\Rightarrow$ 1 är bättre än 2
3. Intervallet innehåller 0, $0 \in I_{\mu_1 - \mu_2}$
   $\Rightarrow$ ingen skillnad mellan dem

• Når det gäller parvisa mätningar, tex "före-efter mätningar" eller om de frågar om "skillnad", ska endast finnas två saker man jämför

(I) ställ upp dina två s.v o skriv fördelningen

$$X_i \sim N(\mu_i, \sigma_1^2)$$
$$Y_i \sim N(\mu_i + \Delta, \sigma_2^2)$$

(II) Bilda $Z_i = Y_i - X_i$

$$\Rightarrow Z_i = Y_i - X_i \sim N(\Delta, \sigma^2)$$

(III) Konstruera I_Δ (OBS: se om det bör vara tvåsidigt el. enkelsidigt)

1. Skatta Δ

gör en tabell över skillnaderna, z_i

i	1 .... n
z_i	

$$\Rightarrow \hat{\Delta} = \frac{\sum_{i=1}^{n} z_i}{n} = \bar{z}$$

2. Skatta σ^2

σ^2 skattas med $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(z_i - \bar{z})^2$

Detta kan göras på miniräknaren:

$\boxed{\text{Stat}}$ ⇒ edit ⇒ ställ dig på "L_1" o tryck $\boxed{\text{enter}}$ för att rensa listan

sätt in talen z_i i listan

tryck $\boxed{\text{stat}}$ ⇒ calc ⇒ 1-var stats ⇒ $\boxed{\text{Enter}}$ (blir då L_1), tryck $\boxed{\text{2nd}}$ sen 2 om du vill ha L_2 ist.

3. Hjälpvariabel

$$\frac{\bar{Z} - \Delta}{S/\sqrt{n}} \sim t(n-1)$$

4. I_Δ

$I_\Delta = (\bar{z} \mp t(n-1) \cdot s/\sqrt{n})$ vid tvåsidigt

$I_\Delta = (-\infty ; \bar{z} + t(n-1) \cdot s/\sqrt{n})$ vid uppåt begränsat

$I_\Delta = (\bar{z} - t(n-1) \cdot s/\sqrt{n} ; \infty)$ vid nedåt begränsat

5. Dra slutsatser

- Konstruera intervall för σ^2

(I) Hjälpvariabel

$$\frac{(m+n-2)s^2}{\sigma^2} \sim \chi^2(m+n-2)$$

där $s^2 = \frac{(m-1)s_1^2 + (n-1)s_2^2}{(m-1)+(n-1)}$

(II) I_{σ^2}

- $I_{\sigma^2} = \left(0, \frac{(m+n-2)s^2}{\chi^2(m+n-2)}\right)$ vid uppåtbegränsat

OBS: står det tex 99% intervall i den här uppgiften ska man ta $\chi^2_{0,01}$!!

VIKTIGT!

Om det står konfidensgraden el. 99% KI
$\Rightarrow \chi^2_{0,01}$

Om det står $\alpha = 0{,}05$
$\Rightarrow \chi^2_{0,95}$

Momentmetoden

(I) Beräkna $E(X)$

$$E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx$$

(II) Momentmetoden ger...

sätt $E(X) = \bar{x}$ Lös ut för den parameter som ska skattas o sätt "hatt" på den, tex $\hat{\theta} = \ldots$

Följdfrågor

- Beräkna väntevärde och varians för skattningen

$E(\hat{\theta}) = \ldots$

$Var(\hat{\theta}) = E(X^2) - E(X)^2$, $E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f_X(x)\, dx$

Regler
- $E(aX + b) = aE(X) + b$
- $Var(aX + b) = a^2 Var(X)$

OBS: Ibland går det att se på $f_X(x)$ vilken fördelning den har o $E(X)$ hämtas lättast från F.S.

Maximum-Likelihood-metoden

- Om $f_X(x)$ är given i uppgiften

Ex.
$$f_X(x) = \begin{cases} \frac{x}{\theta^2} \cdot e^{-x/\theta} & \text{för } x \geq 0 \\ 0 & \text{annars} \end{cases}$$

(I) Teckna $L(\theta)$ ← parametern man vill skatta

$$L(\theta) = \left(\frac{1}{\theta^2}\right)^n \cdot \prod_{i=1}^{n} x_i \cdot e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i}$$ ⟸ standard

(II) Logaritmera $L(\theta)$

$$l(\theta) = \ln L(\theta) = -2n\ln(\theta) + \sum_{i=1}^{n} \ln x_i - \frac{1}{\theta}\sum_{i=1}^{n} x_i$$

Regler:
- $\ln(a \cdot b) = \ln a + \ln b$
- $\ln\left(\frac{a}{b}\right) = \ln a - \ln b$
- $\ln a^c = c \cdot \ln a$
- $\ln e^a = a$
- $e^{\ln b} = b$
- $\ln e = 1$

(III) Derivera $l(\theta)$ m.a.p θ o sätt till 0.

$\frac{dl}{d\theta} = \ldots = 0$ lös ut för parametern som ska skattas o sätt "hatt" på den.

(IV) Kolla så att det är en max-punkt:

$\frac{d^2 l}{d\theta^2} < 0$

Regel: $\sum_{i=1}^{n} c = n \cdot c$

- Om man har två observationer som är olika fördelade samt att man får två olika $f_X(x)$

(I) Teckna $L(\theta)$

Då vi har 2 olika tecknas

$L_1(\theta) = \dots$

$L_2(\theta) = \dots$

Sedan fås

$L(\theta) = L_1(\theta) \cdot L_2(\theta)$

Gör sedan som vanligt...

- Ibland får man ej med täthetsfunktionen i uppgiften

Får man ej med $f_X(x)$ i uppgiften brukar det ofta stå vilken fördelning som används.

$\Rightarrow$ gå till rätt fördelning i F.S o ta $f_X(x)$ därifrån

Gör sedan som vanligt...

- Följdfrågor

1. Är skattningen vvr?

$\Rightarrow$ se om $E(\hat{\theta}) = \theta$

2. Är skattningen konsistent?
   2 krav måste vara uppfyllda:
   - vvr
   - $Var(\hat{\theta}) \to 0$ då $n \to \infty$

3. Vilken är effektivast?

$Var(\hat{\theta}_1) < Var(\hat{\theta}_2) \Rightarrow \hat{\theta}_1$ är mer effektiv än $\hat{\theta}_2$

4. Bestäm variansen av skattningen, $Var(\hat{\theta})$

$Var(X) = E(X^2) - E(X)^2$ kan behöva användas

Regressionsanalys

• Förklaringsgraden, R^2

$R^2 = \frac{SS_R}{SS_{TOT}}$, $R^2 \longrightarrow 1$ så nära 1 som möjligt är bra $\Rightarrow$ ger hög förklaringsgrad

$SS_{TOT} = SS_E + SS_R$

" Förklaringsgraden visar hur stor andel av variationerna i Y som förklaras av $x_1, x_2 \ldots$ "

• Testa om faktorer påverkar

En variabel:

Alt 1

(I) $H_0: \beta_i = 0$ mot $H_1: \beta_i \neq 0$ på nivån ...
där β_i är "faktorn" man vill kolla på

(II) Teststorhet

$$|t| = \frac{\hat{\beta}_i}{d(\hat{\beta}_i)} \sim t(n-k-1)$$

(III) om $|t| > c = t(n-k-1)$

$\Rightarrow$ Förkasta H_0, faktorn som hör till β_i påverkar

Alt 2

Konstruera lämpligt K.I

(I) $I_{\beta_i} = (\hat{\beta}_i \mp t(n-k-1) \cdot d(\hat{\beta}_i))$

slutsats: om 0 än ingår i intervallet påverkar β_i.

Flera variabler:

(I) $H_0: \beta_1 = \beta_3 = 0$ H_1: minst en av $\beta_1 \, \text{o} \, \beta_3 \neq 0$

(II) Teststorhet

$$V = \frac{SS_R / k}{SS_E / n-k-1} \sim F(k, n-k-1)$$

(III) $V > c = F(k, n-k-1)$

$\Rightarrow$ Förkasta H_0, minst någon av förklaringsvariablerna gör nytta

Ex. på fråga:

" I modell 1, gör någon av förklaringsvariablerna nytta? "

• Jämföra två modeller

Förklarar modell i bättre än modell j?

(I) Se vilka β som skiljer modellerna åt

(II) $H_0: \beta_1 = \beta_3 = 0$ $\quad H_1$: minst en av β_1 & $\beta_3 \neq 0$

De β som skiljer

(III) Teststorhet

F.S sida 15

$$W = \frac{(SS_E^{(1)} - SS_E^{(2)})/p}{SS_E^{(2)}/(n-k-p-1)} \sim F(p, n-k-p-1)$$

(IV) om $W > C = F(p, n-k-p-1)$

$\Longrightarrow$ Förkasta H_0! (modell j är bättre)

• Hur många frihetsgrader...

$df_{REGR} = k$

$df_{RES} = n-k-1$ alt. $df_{RES} = df_{TOT} - df_{REGR}$

• Konstruera ett prediktionsintervall el. konfidensintervall, u' given

Skillnad mellan ett P.I. & K.I.:

Välj att göra ett prediktionsintervall om det gäller en enskild mätning, välj annars att göra ett konfidensintervall.

K.I: $I_{u'\beta} = (u'\hat{\beta} \mp t(n-k-1) \cdot s\sqrt{u'(X'X)^{-1}u})$

P.I.: $I_{u'\beta} = (u'\hat{\beta} \mp t(n-k-1) \cdot s\sqrt{1 + u'(X'X)^{-1}u})$

där $s^2 = \frac{SS_E}{(n-k-1)}$