5.) a) $LDL_i = \beta_0 + \beta_1 TCH_i + \beta_2 HDL_i + \beta_3 TRI_i + \varepsilon_i \qquad \varepsilon_i \overset{NEP}{\sim} N(0,\sigma^2)$

(i) $\beta_0, \beta_1, \beta_2, \beta_3$ po MNK

Vemo: $\hat\beta = (Z^T Z)' Z^T \vec X$, kjer je $(Z^T Z)'$ posplošeni inverz

Definiramo: $Z = [1 \ TCH \ HDL \ TRI] = \begin{bmatrix} 1 & TCH_1 & HDL_1 & TRI_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & TCH_n & HDL_n & TRI_n \end{bmatrix}$ in $\vec X = \begin{bmatrix} LDL_1 \\ \vdots \\ LDL_n \end{bmatrix}$ in

$\hat\beta = (\hat\beta_0, \hat\beta_1, \hat\beta_2, \hat\beta_3)$ izračunamo v R na kot: $\qquad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$

- $\hat\beta = (Z^T Z)' Z^T \vec X$ (prek G-S algoritma)
- $\hat\beta = (Z^T Z)^{-1} Z^T \vec X$ (z običajnim inverzom)
- z uporabo funkcije lm vgrajene v R

Po vseh treh metodah dobimo enak rezultat za ocene in sicer:

$\hat\beta_0 = 0.3773332 \quad \hat\beta_1 = 0.8534945 \quad \hat\beta_2 = -0.8311743 \quad \hat\beta_3 = -0.2738845$

(ii) $H_0: \beta_1 = \beta_2 = \beta_3 = 0$ na standarden način pri st. znač. $\alpha = 0.05$

1. način: ponovimo kot pri Ekonometriji 1: Najprej zapišemo oba modela

Unrestricted model: $LDL_i = \beta_0 + \beta_1 TCH_i + \beta_2 HDL_i + \beta_3 TRI_i + \varepsilon_i$

Restricted model: $LDL_i = \beta_0 + \varepsilon_i$

Za vsak model moramo poračunati vsoto kvadratov residualov:

$$VKR = \sum_{i=1}^{n} (\hat\varepsilon_i)^2 = \sum_{i=1}^{n} (y_i - (\hat\beta_0 + \hat\beta X_i))^2$$

Primer to pomeni:

- ~~vsota~~ VKR za unrestricted model: $VKR_U = \sum_{i=1}^{300} (\underset{\vec X}{LDL_i} - \hat\beta Z)^2$, kjer sta β in Z kot v (i)
- VKR za restricted model: $VKR_R = \sum_{i=1}^{300} (LDL_i - \hat\beta_0)^2$

Testiramo z F testom:

$$F = \frac{(VKR_R - VKR_U)/(k_u - k_R)}{VKR_U/(n - k_U)} \qquad test = \begin{cases} H_0 \text{ zavrnemo}; & F > F_{q, n-k_u; \alpha} \\ H_0 \text{ ne zavrnemo}; & \text{sicer} \end{cases}$$

Tu je $q = k_u - k_R$, kjer sta k_U in k_R število prostostnih stopenj v unrestricted in restricted modelu. q je torej število restrikcij, omejitev. $\quad n = 300$

$F_{q, n-k_U; \alpha}$ je vrednost Fisherjeve porazdelitve ~~pri~~ ~~pri~~ za restrikcijami modela in $\alpha = 0.05$

Poračunamo v R in dobimo:

$$F = 426.1932 > 2.6351 = F_{3, 296; 0.05}$$

torej ničelno hipotezo zavrnemo.

2. način:

H_0 zapišemo v matrični obliki: $H_0: L\beta = c$. Da dobimo našo hipotezo moramo vzeti

$$L = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{in } c = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

testna statistika je tu:

$$F = \frac{\dfrac{(L\beta - c)^T [L (Z^T Z)^{-1} L^T]^{-1} (L\beta - c)}{q}}{VKR/(n-d)}$$

test za enak kot v prvem načinu

Poračunamo v R in dobimo: $F = 426 > 2.6351 = F_{3,296;0.05}$

Tudi po tem načinu H_0 ne zavrnemo.

samo prvi stolpec matrike Z: $\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \Rightarrow \beta = (Z^T Z)^{-1} Z^T X$

$Z^T Z = [1 \dots 1] \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = n$ $\quad = n^{-1} \cdot [1 \dots 1] X = \frac{1}{n} \sum_{i=1} X_i = \bar{X}$

3. način: (iz STAT1 vemo)

Postavimo oceno z $\hat{\beta}_{H_0} = \begin{bmatrix} \bar{X} \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} LDL \\ 0 \\ 0 \\ 0 \end{bmatrix}$ število in ~~stopnja~~ omejitev $r = 3$

Vemo: $F(X) = \dfrac{\| Z\hat{\beta}(X) - Z\hat{\beta}_H(X) \|^2 / r}{VKR(X)/(n-d)} \sim F_{r, n-d; \alpha}$ (test torej enak kot pri prejšnjih 2 načinih)

$VKR(X) = \| X - Z\hat{\beta}(X) \|^2$, $\hat{\beta}(X)$ je vektor, s komponentami označenimi v a) (i).

Vstavimo v R, poračunamo in tudi tu dobimo: $F = 426 > 2.6351 = F_{3,296;0.05}$

(iii) Preizkusite $H_0: \beta_0 = 0$ na standarden način

1. način:

Če računamo z R v lm(..) modelu, R nam poračuna t-statistiko.

Moramo le še primerjati s testom:
$$\begin{cases} H_0 \text{ zavrnemo}; & |t| > t_{n-d;\alpha/2} \\ H_0 \text{ ne zavrnemo}; & \text{sicer} \end{cases},$$

kjer je $t_{n-d,\alpha/2}$ studentova porazdelitev. Poračunamo vrednosti v R in dobimo:

$|t| = 1.92036 < 1.9680 = t_{296;0.025}$, torej H_0 ne zavrnemo.

2. način:

Uporabimo enak postopek kot v 2. načinu (ii) točke (Fisher lahko deluje tudi za 1 parameter).

Nastavimo $L = [1\ 0\ 0\ 0]$ in $c = 0$, poračunamo v R in dobimo

~~$F = 3.697 < 3.87$~~ $F = 3.6878 < 3.8731 = F_{1,296;0.05}$ torej H_0 ne zavrnemo

3. način: Vzemimo še način z Studentu iz 1. načina (vsebinsko enak preizkus)

Definiramo $t(X) = \dfrac{\langle \hat{\beta}(X), v \rangle}{\sqrt{\langle (Z^T Z)^{-1} v, v \rangle} \cdot \sqrt{\frac{VKR}{(n-d)}}}$, kjer je $v = (1,0,0,0)$ (vektor)

Poračunamo v R in dobimo: $|t(X)| = 1.92036 < 1.98680 = t_{296;0.025}$, torej H_0 ne zavrnemo.

(po STAT1)

b) ~~Ex~~ $LDL_i = \beta_1 \cdot TCH_i + \beta_2 \cdot HDL_i + \beta_3 \cdot TRI_i + \varepsilon_i$

(i) $\beta_1, \beta_2, \beta_3$ po MNK

Postopamo identično kot v a)(i), le brez prostega člena.
Spet po vseh ~~ske~~ treh metodah dobimo enak rezultat in sicer:

$$\beta_1 = 0.8954892 \qquad \beta_2 = -0.7956929 \qquad \beta_3 = -0.2563894$$

(ii) ~~[illegible]~~ $H_0: (\beta_1, \beta_2, \beta_3) = (1, -1, -0.45)$

1. način iz a)(ii) → ne bo primeren

2. način iz a)(ii):

Zapišemo $H_0: L\beta = c$, kjer vzamemo $L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ in $c = \begin{bmatrix} 1 \\ -1 \\ -0.45 \end{bmatrix}$ $\quad n = 300, \; d = 3$

Kot v a)(ii) poračunamo

$$F = \frac{(L\beta - c)^T [L \cdot (Z^T Z)^{-1} L^T]^{-1} (L\beta - c) / q}{VKR / (n-d)} \qquad \text{test} = \begin{cases} H_0 \text{ zavrnemo}; & F > F_{q, n-d; \alpha} \\ H_0 \text{ ne zavrnemo}; & \text{sicer} \end{cases}$$

Poračunamo v R in dobimo: $F = 20.4 > 2.6350 = F_{3, 297; 0.05}$

~~H~~ Hipotezo H_0 torej zavrnemo.

3. način iz a)(ii):

Postavimo oceno $\hat{\beta}_H = \begin{bmatrix} 1 \\ -1 \\ -0.45 \end{bmatrix}$ in število omejitev = 3

Vemo: $$F(X) = \frac{\| Z\hat{\beta}(X) - Z\hat{\beta}_H(X) \|^2 / r}{VKR(X) / (n-d)} \sim F_{r, n-d; \alpha}$$

Test je seveda enak kot v prejšnjem načinu.

Poračunamo: $VKR(X) = \| X - Z\hat{\beta}(X) \|^2$

Vstavimo v R, izračunamo in dobimo

$$F = 20.4 > 2.6350 = F_{3, 297; 0.05}$$

In seveda tudi po tem načinu ~~ničelno hipotezo~~ H_0 zavrnemo.

(iii) Območje zaupanja za $(\beta_1, \beta_2, \beta_3)$ st. zaupanja 0.95

Postopamo kot na predavanjih:

Imamo $X \sim N(Z\beta, \sigma^2 I)$ $\quad (X = Z\beta + \varepsilon,\ \varepsilon \sim N(0, \sigma^2 I))$

Vemo, da ima Z poln rang, torej sledi:

$$\hat{\beta}(X) = (Z^T Z)^{-1} Z^T X \sim N(\beta, \sigma^2 (Z^T Z)^{-1})$$

$$\hat{\sigma}^2(X) = \frac{1}{n-d} VKR = \frac{1}{n-d} \|X - \hat{X}\|^2, \text{ kjer je } \frac{\|X - \hat{X}\|^2}{\sigma^2} = \frac{\|X - Z\hat{\beta}\|^2}{\sigma^2} \sim \chi^2_{n-d}$$

Dalje sta $Z^T X$ in $VKR(X)$ neodvisni (torej tudi $\hat{\beta}(X)$ in $VKR(X)$ neodvisni)

Standardiziramo $\hat{\beta}$: $\frac{1}{\sigma} \sqrt{Z^T Z} (\hat{\beta}(X) - \beta) \sim N(0, I_{d \times d})$

in sledi: $\|\frac{1}{\sigma} \sqrt{Z^T Z} (\hat{\beta}(X) - \beta)\|^2 \sim \chi^2_d$

Sledi:

$$F(X; \beta) = \frac{\frac{1}{\sigma^2} \langle \sqrt{Z^T Z} (\hat{\beta}(X) - \beta), \sqrt{Z^T Z} (\hat{\beta}(X) - \beta) \rangle / d}{\frac{1}{\sigma^2} \|X - Z\hat{\beta}\|^2 / (n-d)} \sim \frac{\chi^2_d / d}{\chi^2_{n-d} / (n-d)}$$

Vemo namreč, da če sta $H \sim \chi^2_n$ in $K \sim \chi^2_k$ neodv. ima $\frac{H/n}{K/k}$ porazdelitev $F_{n,k}$

Sledi torej

$$F(X; \beta) = \frac{\langle Z^T Z (\hat{\beta}(X) - \beta), (\hat{\beta}(X) - \beta) \rangle / d}{\|X - Z\hat{\beta}\|^2 / (n-d)} \sim F_{d, n-d, \alpha}$$

Nastavimo $D = (0, F_{d, n-d; \alpha}]$. Sedaj je $P(F_{d, n-d} \in D) = 1 - \alpha$.

Pridruženo območje zaupanja je $C(X) = \{\beta \mid F(X; \beta) \leq F_{d, n-d; \alpha}\}$

Diagonaliziramo $Z^T Z = Q \cdot \Lambda \cdot Q^T$, kjer je $Q \in O(d)$ ortogonalna in Λ diagonalna z $\lambda_i > 0$.

Sledi: $\langle (Z^T Z)(\hat{\beta} - \beta), \hat{\beta} - \beta \rangle = \langle \Lambda Q^T (\beta - \hat{\beta}), Q^T (\beta - \hat{\beta}) \rangle$

Rešujemo $\langle \Lambda Q^T (\beta - \hat{\beta}), Q^T (\beta - \hat{\beta}) \rangle \leq \frac{d}{n-d} VKR \cdot F_{d, n-d; \alpha} =: \eta^2 = \eta^2(X)$

Rešitev ~~je~~ ko območje zaupanja, ki je elipsoid:

$$C(X) = \left\{ \hat{\beta} + Q \cdot \begin{bmatrix} 1/\sqrt{\lambda_1} \cdot u_1 \\ \vdots \\ 1/\sqrt{\lambda_d} \cdot u_d \end{bmatrix} \;\middle|\; u_1^2 + \ldots + u_d^2 \leq \eta^2 \right\}$$

Kaj je torej treba narediti?

- diagonaliziramo $Z^T Z = Q \cdot \Lambda \cdot Q^T$, kjer je Q ~~vektor~~ matrika lastnih ~~vrednosti~~ vektorjev $Z^T Z$, in Λ diagonalna matrika z $\lambda_i > 0$ ~~a~~, ki so lastne vrednosti $Z^T Z$, na diagonali.
- η^2 izračunamo po formuli $\eta = \frac{d}{n-d} \cdot VKR \cdot F_{d, n-d; \alpha}$

Poračunamo v R in dobimo:

$\lambda_1 = 25739.25 \qquad \lambda_2 = 368.3234 \qquad \lambda_3 = 118.1484 \qquad \eta^2 = 2.057032$

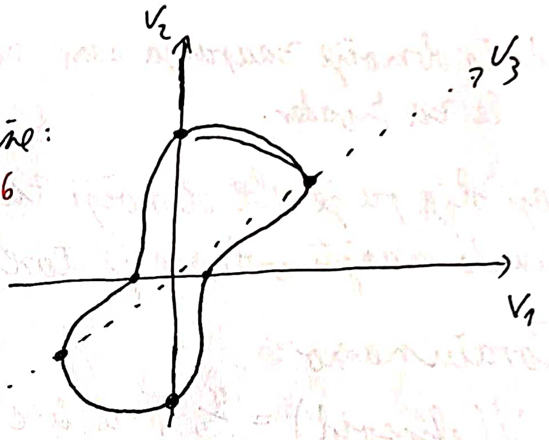
$$Q = \text{matrika last. vekt.} = [v_1 \ v_2 \ v_3] = \left[\begin{bmatrix} 0.9340480 \\ 0.1982784 \\ 0.2970521 \end{bmatrix} \begin{bmatrix} -0.2341740 \\ -0.2879862 \\ 0.9285615 \end{bmatrix} \begin{bmatrix} 0.2696606 \\ -0.9368824 \\ -0.2225613 \end{bmatrix} \right]$$

Dobimo torej območje zaupanja, ki je elipsoid:

$$\beta \in \left\{ \hat\beta + Q \begin{bmatrix} 1/\sqrt{\lambda_1}\, u_1 \\ 1/\sqrt{\lambda_2}\, u_2 \\ 1/\sqrt{\lambda_3}\, u_3 \end{bmatrix} \;\middle|\; u_1^2 + u_2^2 + u_3^2 \leq \eta^2 \right\}$$

s polosmi $\eta/\sqrt{\lambda_1}$, $\eta/\sqrt{\lambda_2}$, $\eta/\sqrt{\lambda_3}$ na stolpcih matrike Q.

To je torej elipsoid v koordinatnem sistemu $v_1 \times v_2 \times v_3$ s središčem v $\hat\beta$ in z polosmi dolžine:



$a = \eta^2/\sqrt{\lambda_1} = 0.01639642$ | $a = \sqrt{\eta^2/\lambda_1} = 0.01143216$

$b = \eta^2/\sqrt{\lambda_2} = 0.1071831$ | $b = \sqrt{\eta^2/\lambda_2} = 0.07473186$

$c = \eta^2/\sqrt{\lambda_3} = 0.1892461$ | $c = \sqrt{\eta^2/\lambda_3} = 0.1319491$

Ker slika pove več kot 1000 besed sem raje območje zaupanja narisal v programu R kot 3D diagram. Koda za izris diagrama je v R datoteki ~~ng~~ in je tu ne bom natančneje opisoval.

(iv) Hkratni intervali zaupanja za $\beta_1, \beta_2, \beta_3$ z Bonferronijevim popravkom.

Iz predavanj vemo: Ker $\hat\beta \sim N(\beta, \sigma^2(Z^TZ)^{-1}) \Rightarrow \hat\beta_j \sim N(\beta_j, \sigma^2(Z^TZ)^{-1}_{jj})$

Sledi

$$T_{n-d} \sim \frac{\frac{\hat\beta_j - \beta_j}{\sigma\sqrt{(Z^TZ)^{-1}_{jj}}}}{\sqrt{\frac{1}{\sigma^2}\|X - Z\hat\beta\|^2/(n-d)}}$$

kar je pivotna funkcija za $e(\beta, \sigma^2) = \beta_j$

Če vzamemo $D = [-t_{n-d;\alpha/2}, t_{n-d;\alpha/2}]$ dobimo interval zaupanja:

$$\left[\hat\beta_j - t_{n-d;\alpha/2} \cdot \sqrt{\frac{(Z^TZ)^{-1}_{jj}\|X - Z\hat\beta\|^2}{n-d}},\ \hat\beta_j + t_{n-d;\alpha/2}\sqrt{\frac{(Z^TZ)^{-1}_{jj}\|X - Z\hat\beta\|^2}{n-d}} \right]$$

Da dobimo območje zaupanja iz intervalov zaupanja za posamezen β_j ~~za~~ lahko po Bonferroniju zamenjamo α s popravljeno $\frac{\alpha}{d}$ ($d = 3$ pri nas).

Nato iz dobljenih I.Z. s kartezičnim produktom sestavimo območje zaupanja.

Poračunamo v R in dobimo:

$$\underbrace{\hat{\beta}_j \pm t_{n-d;\frac{\alpha}{2d}} \cdot \sqrt{\frac{(Z^TZ)^{-1}_{jj}}{(n-d)} \cdot VKR}}_{\text{meja}}$$

I.Z. za β_1: $[0.8603245, 0.9306539]$
I.Z. za β_2: $[-0.9031617, -0.6882240]$
I.Z. za β_3: $[-0.3209717, -0.1917910]$

Tudi to območje zaupanja sem narisal v R, vendar ni zelo aktualno, saj gre le za kvader.

Najbolje pa je obe območji zaupanja primerjati po volumnu - tisto, ki bo imelo manjši volumen je torej bolj natančno in posledično boljše.

Poračunamo:

$V(\text{elipsoid}) = \frac{4}{3} \cdot \pi \cdot a \cdot b \cdot c = 0.001\cancel{39}37 = 0.0004722037$

~~V(Bonferonijev~~

$V(\text{Bonferronijev kvader}) = d(IZ1) \cdot d(IZ2) \cdot d(IZ3) = 0.0019528$

($d(\cdot)$ je dolžina pripadajočega intervala zaupanja).

Volumen elipsoida je za ~~občutno~~ precej manjši, kar pomeni da je to O.Z. bolj natančno.