14.4 Hothorn-Bretz-Westfall procedure

24

Def 14.6 Max-abs-t distribution

25

Let $T = (T_1, \dots, T_m)^T \sim mvt_{m,\nu}(\Sigma)$, where Σ is a positive semidefinite matrix. The distribution of a random variable $H = \max_{j=1,\dots,m} |T_j|$

will be called the max-abs-t-distribution of dimension m with ν degrees of freedom and a scale matrix Σ and will be denoted as $h_{m,\nu}(\Sigma)$.

NOTATION:

- $0 < p < 1$: $h_{m,\nu}(p; \Sigma) = p \cdot 100\%$ quantile of distribution $h_{m,\nu}(\Sigma)$,

  i.e.

  $$p = P\left(\max_{j=1,\dots,m} |T_j| \le h_{m,\nu}(p; \Sigma)\right)$$

- $CDF_{h,m,\nu}(\cdot\,; \Sigma) \equiv$ cdf of a random variable $H \sim h_{m,\nu}(\Sigma)$.

Remarks:

- $\Sigma > 0 \Rightarrow \underbrace{f_T(t)}_{\text{density of } T} \propto |\Sigma|^{-1/2} \left(1 + \frac{t^T \Sigma^{-1} t}{\nu}\right)^{-\frac{\nu+m}{2}}$

- $CDF_{h,m,\nu}(h, \Sigma) = P\left(\max_{j=1,\dots,m} |T_j| \le h\right) =$

$$= P(\forall j: |T_j| \le h) = \underbrace{\int_{-h}^{h} \cdots \int_{-h}^{h} f_T(t)\, dt_1 \dots dt_m}$$

numerical methods (Monte Carlo integration) available to calculate such integrals and to determine quantiles

$\rightarrow$ R packages mvtnorm and mnormt

14.4.2 General multiple comparison procedure for a linear model

Consider a linear model $Y|X \sim N_n(X\beta, \sigma^2 I_n)$

$\text{rank}(X_{n\times k}) = k < n$

- $\mathbb{L} = \begin{pmatrix} l_1^T \\ \vdots \\ l_m^T \end{pmatrix}$ given matrix

- $\theta := L\beta = (\underbrace{l_1^T\beta}_{\theta_1}, \dots, \underbrace{l_m^T\beta}_{\theta_m})^T$ ≡ parameters of interest

---

We allow for
- $m > k$
- linearly dependent rows in L
- matrix $V := L(X^TX)^{-1}L^T$ neither diagonal nor invertible

---

<u>Multiple comparison problem</u>

$H_0: \quad \underbrace{\theta_1 = \theta_1^0}_{\substack{l_1^T\beta = \theta_1^0 \\ H_1}} \ \& \dots \& \ \underbrace{\theta_m = \theta_m^0}_{\substack{l_m^T\beta = \theta_m^0 \\ H_m}}$ global hyp.

element. hyp.

$H_0: \ \Theta = \Theta^0$ for chosen $\Theta^0 \in \mathbb{R}^m$

(standard) notation

- $\hat{\beta} = (X^T X)^{-1} X^T Y$
- $\hat{\theta} = \mathbb{L}\hat{\beta} = (l_1^T \hat{\beta}, \ldots, l_m^T \hat{\beta})^T = (\hat{\theta}_1, \ldots, \hat{\theta}_m)^T$
  $= \text{LSE of } \theta$
- $V = \mathbb{L}(X^T X)^{-1} \mathbb{L}^T = (v_{jl})_{\substack{j=1,\ldots,m \\ l=1,\ldots,m}}$
- $\mathbb{D} = \operatorname{diag}\left(\frac{1}{\sqrt{v_{11}}}, \ldots, \frac{1}{\sqrt{v_{mm}}}\right)$
  $\phantom{\mathbb{D} = \operatorname{diag}} \; d_1, \ldots, d_m$
- MS_e: residual mean square of the model with $\nu_e = n-k$ degrees of freedom

---

REPETITION (properties of LSE under normality)

$$\hat{\theta}_j \sim \mathcal{N}(\theta_j, \sigma^2 v_{jj}) \qquad \text{(given } X\text{)}$$

$$Z_j = \frac{\hat{\theta}_j - \theta_j}{\sqrt{\sigma^2 v_{jj}}} \sim \mathcal{N}(0,1) \qquad \text{(also uncondit.)}$$

$$Z_j = \frac{1}{\sqrt{\sigma^2}} d_j (\hat{\theta}_j - \theta_j)$$

$$T_j = \frac{\hat{\theta}_j - \theta_j}{\sqrt{MS_e v_{jj}}} \sim t_{n-k} \qquad \text{(also uncondit.)}$$

$$T_j = \frac{1}{\sqrt{MS_e}} d_j (\hat{\theta}_j - \theta_j)$$

For vectors (given X):

Would be shown easily

$$\mathbb{Z} = (Z_1, \ldots, Z_m)^T = \frac{1}{\sqrt{\sigma^2}} \mathbb{D}(\hat{\theta} - \theta) \sim \mathcal{N}_m(0_m, \mathbb{D}V\mathbb{D})$$

$$\mathbb{T} = (T_1, \ldots, T_m)^T = \frac{1}{\sqrt{MS_e}} \mathbb{D}(\hat{\theta} - \theta) \sim mvt_{m,n-k}(\mathbb{D}V\mathbb{D})$$

Theorem 6.2(ix)

NOTES:

- diagonal elements of V and DVD are >0 but V and DVD are not necessarily positive definite
- if $\operatorname{rank}(\mathbb{L}) = m \leq k$

$\Rightarrow$ $V>0$, $DVD>0$ and under $H_0: \theta = \theta^0$, we have (Theorem 6.2 (x))

$$Q_0 = \frac{1}{m}(\hat{\theta}-\theta^0)^T(MSe \cdot V)^{-1}(\hat{\theta}-\theta^0)$$

$$= \frac{1}{m} T^T (DVD)^{-1} \mathbb{T} \sim F_{m,n-k}$$

(even uncondit.)

→ elliptical confidence regions ~~for~~ for θ

→ will not be used here

→ different approach now, but also used on statistics $\mathbb{T}$

Remember (again); with given $\theta = (\theta_1, .., \theta_m)^T$: (true value of the param.)

$$T = (T_1, .., T_m)^T = \frac{1}{\sqrt{MSe}} D(\hat{\theta}-\theta) \sim mvt_{m,n-k}(DVD)$$

conditionally given X

$$T_j = \frac{\hat{\theta}_j - \theta_j}{\sqrt{MSe\, v_{jj}}}$$

Let for $\theta^0 = (\theta_1^0, \dots, \theta_m^0) \in \mathbb{R}^m$

$$T_j^0 := T_j(\theta_j^0) := \frac{\hat{\theta}_j - \theta_j^0}{\sqrt{MSe\, v_{jj}}} \quad , j = 1, \dots, m$$

Then under $H_0 : \theta = \theta^0$

$$\mathbb{T}^0 = \mathbb{T}(\theta^0) = (T_1^0, \dots, T_m^0)^T \underset{\text{given } X}{\sim} mvt_{m, n-k}(\mathbb{D}V\mathbb{D})$$

Hence $\forall \theta^0 \in \mathbb{R}^m$, $\forall\, 0 < \alpha < 1$

$$1 - \alpha = P\Big(\max_{j=1,\dots,m} |T_j^0| \underset{\text{does not matter}}{\overset{\leq}{=}} h_{m,n-k}(1-\alpha, DVD) ; \theta = \theta^0 \Big)$$

$$= P\big(\forall j = 1,\dots,m \quad |T_j^0| < h_{m,n-k}(1-\alpha, DVD) ; \theta = \theta^0 \big)$$

$$= P\Big(\forall j = 1,\dots,m \quad \Big| \frac{\hat{\theta}_j - \theta_j^0}{\sqrt{MSe\, v_{jj}}} \Big| < h_{m,n-k}(1-\alpha, DVD) ; \theta = \theta^0 \Big)$$

$$= P\big(\forall j = 1,\dots,m \quad (\theta_j^{HL}(\alpha), \theta_j^{HU}(\alpha)) \ni \theta_j^0 ; \theta = \theta^0 \big),$$

where $\theta_j^{HL}(\alpha) = \hat{\theta}_j - \sqrt{MSe\, v_{jj}}\; h_{m,n-k}(1-\alpha, DVD)$

$\theta_j^{HU}(\alpha) = \hat{\theta}_j + \sqrt{MSe\, v_{jj}}\; h_{m,n-k}(1-\alpha, DVD)$

It seems that we have just derived a set of simultaneous confidence intervals for θ_j, $j = 1, \dots, m$.

Theorem 14.5 Hothorn-Bretz-Westfall MCP for linear hypotheses in a normal lin. model

Short: • simultaneous confidence intervals for $\theta_j = l_j^T \beta$, $j = 1, .., m$

exact ⊗ coverage

$$1-\alpha = P\left(\forall j = 1, .., m \; (\theta_j^{HL}(\alpha), \theta_j^{HU}(\alpha)) \ni \theta_j^0 ; \; \theta = \theta^0\right)$$

• P-values adjusted for multiple comparisons, & elementary hypotheses $H_j : \theta_j = \theta_j^0$, $j = 1, .., m$

$$p_j^H = 1 - CDF_{h, m, n-k}(|t_j^0|; DVD), \quad j = 1, .., m$$

t_j^0 = value of $T_j^0 = T_j(\theta_j^0) = \dfrac{\hat{\theta}_j - \theta_j^0}{\sqrt{MSe \, v_{jj}}}$

attained with given data

Proof: The only issue to clarify are the p-values adjusted for multiple comparison.

We have

$$\forall j: \quad [\theta_j^{HL}(\alpha), \theta_j^{HU}(\alpha)) \not\ni \theta_j^0$$

$\Leftrightarrow$ H_j rejected by MCP

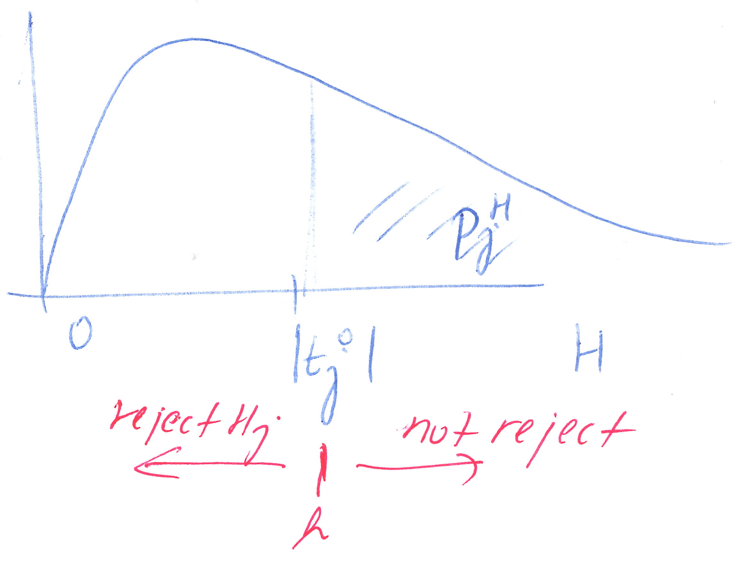
$$\Leftrightarrow \quad |T_j(\theta_j^0)| \geq h_{m,n-k}(1-\alpha, DVD)$$

Hence $p_j^H = \inf\{\alpha : [\theta_j^{HL}(\alpha), \theta_j^{HU}(\alpha)) \not\ni \theta_j^0\}$

$$= \inf\{\alpha : |T_j(\theta_j^0)| \geq h_{m,n-k}(1-\alpha, DVD)\}$$

Let t_j^0 = value of $T_j(\theta_j^0)$ attained with given data

$$p_j^H = \inf\{\alpha : |t_j^0| \geq h_{m,n-k}(1-\alpha, DVD)\}$$



$$p_j^H = 1 - CDF_{h_{m,n-k}}(|t_j^0|; DVD)$$

R package <u>multcomp</u>

Final remarks

Simultaneous conf. intervals:

$$\hat{\theta}_j \pm \sqrt{MSe\, l_j^T (X^T X)^{-1} l_j} \cdot h_{m,n-k}(1-\alpha; DVD)$$

VII

Single conf. interval:

$$\hat{\theta}_j \pm \sqrt{MSe\, l_j^T (X^T X)^{-1} l_j} \cdot t_{n-k}\left(1-\frac{\alpha}{2}\right)$$

14.5 Confidence band for the regression function

We will assume now:

$$(Y_i, Z_i^T)^T \overset{iid}{\sim} (Y, Z^T)^T, \quad i = 1, .., n$$

For a known transformation $t : \mathbb{R}^p \to \mathbb{R}^k$, linear model holds with ~~tt~~ regressors $X_i = t(Z_i)$, $i = 1, .., n$ + normality

We assume

$$Y | \mathbb{Z} \sim \mathcal{N}_n (\mathbb{X}\beta, \sigma^2 I_n), \quad \text{rank}(\mathbb{X}_{n \times k}) = k$$

$$Y_i | Z_i \sim \mathcal{N}(X_i^T \beta, \sigma^2), \quad X_i = t(Z_i)$$

$$i = 1, .., n$$

Regression function

$$\mathbb{E}(Y | X = t(z)) = \mathbb{E}(Y | Z = z) = m(z) = t^T(z)\beta,$$

$$z \in \mathbb{R}^p$$

Theorem 6.3 $\to$ confidence interval for the model based mean

$$\forall z \in \mathbb{R}^p \; \forall \beta^0 \in \mathbb{R}^k \; \forall \sigma_0^2 > 0$$

$$P\left(t^T(z)\hat{\beta} \pm t_{n-k}\left(1 - \frac{\alpha}{2}\right)\sqrt{MS_e \, t^T(z)(\mathbb{X}^T\mathbb{X})^{-1} t(z)} \ni t^T(z)\beta^0 ; \beta = \beta^0, \sigma^2 = \sigma_0^2\right)$$

$$= 1 - \alpha$$

NOW: $P(\forall z \in \mathbb{R}^p \ldots\ldots) = 1 - \alpha$

Theorem 14.6 Confidence band for the regression function

[31]

SHORT: $\forall \beta^0 \in \mathbb{R}^k \; \forall \sigma_0^2 > 0$

$$1-\alpha = P\left(\forall z \in \mathbb{R}^p,\; t^T(z)\hat{\beta} \pm \sqrt{k\, F_{k,n-k}(1-\alpha)\, MS_e\, t^T(z)(X^TX)^{-1}t(z)} \ni t^T(z)\beta^0;\; \beta=\beta^0, \sigma^2=\sigma_0^2\right)$$

Proof: see the notes (not exam)

Band FOR the regression function versus band AROUND the regres. fun. [32]

Illustration: Kojeni [33]