VI. Normal Linear Model

- again more theoretical part that follows chapter II (Least Squares Estimation)
- properties of LSE related quantities under normality
- by now: no distributional assumptions were needed to derive mentioned properties (the only exception: moments of standardized residuals)

~~6.1 Norr~~ **REPETITION:**

$Y|X \sim (X\beta, \sigma^2 I)$, $\operatorname{rank}(X_{n\times k}) = r \le k$ (a.s.)

$r = k \Rightarrow \hat{\beta} = (X^T X)^{-1} X^T Y$ is BLUE of β

$\operatorname{var}(\hat{\beta}|X) = \sigma^2 (X^T X)^{-1}$

$\hat{Y} = X\hat{\beta} = H \cdot Y$, $H = X(X^T X)^- X^T = QQ^T$

even if $r<k$

Q = orthonormal basis of $\mathcal{M}(X)$

$\equiv$ BLUE of $X\beta$

$\operatorname{var}(\hat{Y}|X) = \sigma^2 H$

$U = Y - \hat{Y} = MY \quad (= M\varepsilon, \ \varepsilon = Y - X\beta)$

$M = I_n - H_X = NN^T$

N = orthonormal basis of $\mathcal{M}(X)^\perp$

$E(U|X) = 0$, $\operatorname{var}(U|X) = \sigma^2 M$

$$\|U\|^2 = U^T U = Y^T \underbrace{MM}_{M} Y = \varepsilon^T M \varepsilon =: SS_e$$

$\frac{1}{n-r} \cdot SS_e = MS_e$: unbiased estimator of σ^2

6.1 Normal linear model

1

DATA: $(Y_i, X_i^T)^T \sim (Y, X^T)^T$, $i = 1, \dots, n$

LM: $\mathbb{E}(Y|X) = X^T\beta$ for some $\beta \in \mathbb{R}^k$

$\mathrm{var}(Y|X) = \sigma^2$ for some $\sigma^2 > 0$

New assumption $Y|X \sim \mathcal{N}(X^T\beta, \sigma^2)$

if additionally $(Y_i, X_i^T)^T \overset{iid}{\sim} (Y, X)$

then the new assumption also implies

$$\mathbb{Y}|\mathbb{X} \sim \mathcal{N}(\mathbb{X}\beta, \sigma^2 I_n)$$

In general, data $(Y_i, X_i^T)^T$ do not have to be iid to satisfy $\mathbb{Y}|\mathbb{X} \sim \mathcal{N}(\mathbb{X}\beta, \sigma^2 I_n)$.

Remember (joint density of $\mathbb{Y}, \mathbb{X}$)

$$f_{\mathbb{Y},\mathbb{X}}(y, x) = f_{\mathbb{Y}|\mathbb{X}}(y|x)\, f_{\mathbb{X}}(x) =$$

$$\overset{Y|X \sim N}{=} \left\{ \prod_{i=1}^{n} \frac{1}{\sigma} \varphi\left(\frac{y_i - x_i^T\beta}{\sigma} \right) \right\} \underline{f_{\mathbb{X}}(x)}$$

→ can still be whatever

independence here

but not necessarily here

→ Further, we will not require iid data, just

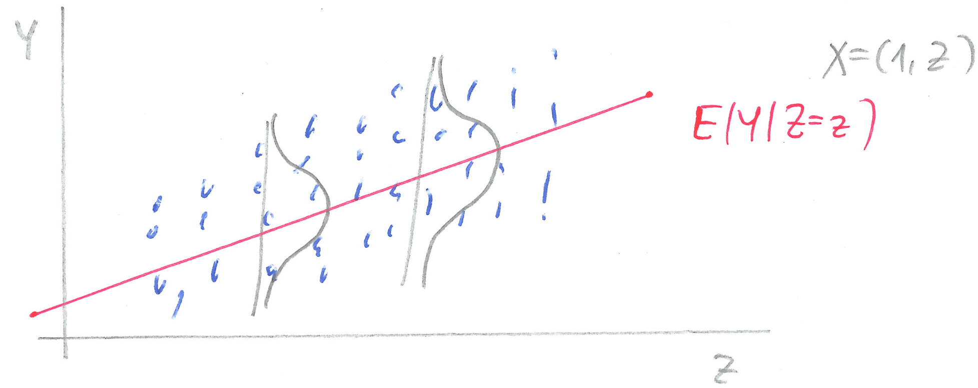
$$\mathbb{Y}|\mathbb{X} \sim \mathcal{N}_n(\mathbb{X}\beta, \sigma^2 I_n)$$

Def 6.1 Normal linear model with general data [2]

The data ~~satisfy~~ (Y, X) satisfy a normal linear model if $Y|X \sim N(X\beta, \sigma^2 I_n)$, where $\beta = (\beta_0, \dots, \beta_{k-1})^T \in \mathbb{R}^k$ and $0 < \sigma^2 < \infty$ are unknown parameters.

---

[3]



Lemma 6.1 Error terms in a normal linear model

Let $Y|X \sim N(X\beta, \sigma^2 I_n)$. The error terms

$$\varepsilon = Y - X\beta = (Y_1 - X_1^T\beta, \dots, Y_n - X_n^T\beta)^T = (\varepsilon_1, \dots, \varepsilon_n)^T$$

then satisfy

(i) $\varepsilon | X \sim N_n(0, \sigma^2 I_n)$.

(ii) $\varepsilon \sim N_n(0, \sigma^2 I_n)$.

(iii) $\varepsilon_i \overset{iid}{\sim} \varepsilon$, $i = 1, \dots, n$, $\varepsilon \sim N(0, \sigma^2)$.

Proof: $\varepsilon = Y - X\beta$

(i) $\varepsilon | X = (Y - \underbrace{X\beta}_{\text{(const / given condition)}}) | X \sim N_n(0, \sigma^2 I_n)$

(ii) $\varepsilon | X = x \sim N_n(0, \sigma^2 I_n)$ for almost all values of x

$\Rightarrow \varepsilon \sim N_n(0, \sigma^2 I_n)$

(iii) properties of normal distribution

6.2 Properties of the least squares estimators under the normality

4

Theorem 6.2 Least squares estimators under the normality

5

Let $Y|X \sim N_n(X\beta, \sigma^2 I_n)$, $\text{rank}(X_{n\times k}) = r \le k$.

Let $L_{m\times k}$ be a real matrix with non-zero rows $l_1^T, \dots, l_m^T$ and $\Theta := L\beta = (l_1^T\beta, \dots, l_m^T\beta)^T = (\theta_1, \dots, \theta_m)^T$ be a vector of linear combinations of regression parameters.

If additionally $r = k$, let $\hat\beta = (X^TX)^{-1}X^TY$ be the LSE of regression coefficients, $\hat\Theta = L\hat\beta = (l_1^T\hat\beta, \dots, l_m^T\hat\beta)^T = (\hat\theta_1, \dots, \hat\theta_m)^T$ and $V = L(X^TX)^{-1}L^T = (v_{j,t})_{j,t=1,\dots,m}$.

etc.

remember $\text{var}(\hat\Theta|X) = \sigma^2 \underbrace{L(X^TX)^{-1}L^T}_{V}$

Many properties then hold (always mentioned before proving them)

Proof: ...

(i) $\hat{Y}|X \sim \mathcal{N}_n(X\beta, \sigma^2 H)$.

We already know (Theorem 2.4) $E(\hat{Y}|X) = X\beta$

$var(\hat{Y}|X) = \sigma^2 H$

$\hat{Y} = H \cdot Y$, $Y|X \sim \mathcal{N}_n \Rightarrow \hat{Y}|X \sim \mathcal{N}_n$. $\boxed{\text{III}}$

↑ depends on X

(ii) $U|X \sim \mathcal{N}_n(0, \sigma^2 M)$.

We already know (Lemma 2.7) $E(U|X) = 0$

$var(U|X) = \sigma^2 M$

$U = MY$, $Y|X \sim \mathcal{N}_n \Rightarrow U|X \sim \mathcal{N}_n$. $\boxed{\text{III}}$

↑ depends on X

(iii) $\hat{\Theta}|X \sim \mathcal{N}_n(\Theta, \sigma^2 V)$. $\hat{\Theta} = L\hat{\beta}$

We already know (Theorem 2.5) $E(\hat{\Theta}|X) = \Theta \quad (= L\beta)$

$var(\hat{\Theta}|X) = \sigma^2 \underbrace{L(X^TX)^{-1}L^T}_{V}$

Same arguments as in proof of Theorem 2.5.

- $r = k \Rightarrow \mathcal{M}(X^T) \equiv \mathbb{R}^k$, that is $\forall j \in 1,\dots,m$ $\ell_j \in \mathcal{M}(X^T)$ $\mathbb{R}^k$
- in other words, $\mathcal{M}(L^T) \subset \mathcal{M}(X^T)$

$\Rightarrow \exists$ matrix $A_{n \times m}$ such that $L^T = X^T A$

$\Rightarrow \hat{\Theta} = L\hat{\beta} = A^T X \hat{\beta} = A^T \hat{Y}$

(i) $\hat{Y}|X \sim \mathcal{N}_n \Rightarrow \hat{\Theta}|X \sim \mathcal{N}_n$.

$\boxed{\text{III}}$

(iv) Statistics $\hat{Y}$ and U are conditionally, given X, independent

6

Jointly: $\begin{pmatrix} \hat{Y} \\ U \end{pmatrix} = \begin{pmatrix} HY \\ MY \end{pmatrix} = \begin{pmatrix} H \\ M \end{pmatrix} Y$

$Y|X \sim N_n$

$\underset{\text{jointly}}{\Longrightarrow} \begin{pmatrix} \hat{Y} \\ U \end{pmatrix} \Big| X \sim N_n$

$\operatorname{cov}(\hat{Y}, U | X) = \operatorname{cov}(HY, MY | X) =$

$= H \underbrace{\operatorname{var}(Y|X)}_{\sigma^2 I_n} \cdot M^T = \sigma^2 \underbrace{H \cdot M}_{0} \quad (= \sigma^2 Q \underbrace{Q^T N}_{0} N^T)$

uncorrelated & jointly N $\Longrightarrow$ independent. $\boxed{M}$

(v) Statistics $\hat{\theta}$ and SSe are conditionally, given X, independent

- $\hat{\theta} = L\hat{\beta} = A^T \hat{Y}$ for some matrix $A_{n \times m}$
  $=$ (linear) function of $\hat{Y}$ (see point iii)
- $SSe = \|U\|^2 =$ (measurable) function of U

$\overset{(iv)}{\Longrightarrow} \hat{\theta} \perp\!\!\!\perp SSe$ (given X) $\boxed{M}$

REMARK: Consider a model $Y | \mathbb{1} \sim N_n(\mathbb{1}\mu, \sigma^2 I_n)$
$(\equiv Y_1, \dots, Y_n \overset{iid}{\sim} N(\mu, \sigma^2))$

$\Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \overline{Y}$
$SSe = \sum (Y_i - \overline{Y})^2 \longrightarrow S^2 := \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$ } are independent using (v)

(vi) $$\frac{\|\hat{Y} - X\beta\|^2}{\sigma^2} \sim \chi^2_r$$

(vii) $$\frac{SSe}{\sigma^2} = \frac{\|U\|^2}{\sigma^2} \sim \chi^2_{n-r}$$

Remember $Y = X\beta + \varepsilon$

$$\rightarrow \quad \hat{Y} - X\beta = \underbrace{HY}_{H(X\beta+\varepsilon)} - X\beta = \underbrace{HX}_{X}\beta + H\varepsilon - X\beta =$$

$$= H \cdot \varepsilon$$

$$U = MY = \underbrace{MX}_{0}\beta + M\varepsilon = M \cdot \varepsilon$$

$$\|\hat{Y} - X\beta\|^2 = \|H\varepsilon\|^2 = \varepsilon^T H^T H \varepsilon = \varepsilon^T H \varepsilon \quad , H = QQ^T$$

$$\|U\|^2 = \|M\varepsilon\|^2 = \varepsilon^T M^T M \varepsilon = \varepsilon^T M \varepsilon \quad , M = NN^T$$

Further, only $\|U\|^2$ will be examined, the other term would be analogous.

$$\|U\|^2 = \varepsilon^T M \varepsilon = \varepsilon^T N N^T \varepsilon = \|N^T \varepsilon\|^2$$

$$\frac{1}{\sigma^2} SSe = \frac{1}{\sigma^2}\|U\|^2 = \left\| \frac{1}{\sigma} N^T \varepsilon \right\|^2$$

We know (Lemma 6.1): $\varepsilon | X \sim N_n(0, \sigma^2 I_n)$

$$\Rightarrow \quad \frac{1}{\sigma} N^T \varepsilon \,|\, X \sim N_{n-r}(\underset{?}{\bullet}, \underset{?}{\bullet})$$

$$\mathbb{E}\left(\frac{1}{\sigma} N^T \varepsilon \,/\, \mathbb{X}\right) = \frac{1}{\sigma} N^T \mathbb{E}(\varepsilon|\mathbb{X}) = \mathbb{O}_{n-r}$$

$$\mathrm{var}\left(\frac{1}{\sigma} \underset{\uparrow \text{depends on } \mathbb{X}}{N^T} \varepsilon \,/\mathbb{X}\right) = \frac{1}{\sigma^2} N^T \underbrace{\mathrm{var}(\varepsilon|\mathbb{X})}_{\sigma^2 I_n} N =$$

$$= \frac{\sigma^2}{\sigma^2} \cdot \underbrace{N^T N}_{I_{n-r}} = I_{n-r}$$

That is $\quad \frac{1}{\sigma} N^T \varepsilon \,/\mathbb{X} \sim \mathcal{N}_{n-r}(\mathbb{O}, I_{n-r})$.

$$\Longrightarrow \left\| \frac{1}{\sigma} N^T \varepsilon \right\|^2 /\mathbb{X} \sim \chi^2_{n-r}$$

For almost all values of $\mathbb{X}$

$$\underbrace{\left\| \frac{1}{\sigma} N^T \varepsilon \right\|^2}_{= \frac{1}{\sigma^2} SS_e} /\mathbb{X} \sim \chi^2_{n-r} \quad \Longrightarrow \left\| \frac{1}{\sigma} N^T \varepsilon \right\|^2 \sim \chi^2_{n-r}$$

(unconditionally)

The fact that $\quad \dfrac{\|\hat{Y} - \mathbb{X}\beta\|^2}{\sigma^2} \sim \chi^2_r$

would be shown in the same way while using matrix $H = Q \cdot Q^T$ instead of $M = NN^T$.

Insertion before we proceed with Theorem

$\theta_j = l_j^T \beta$, $j=1,\dots m$

$\hat{\theta}_j = l_j^T \hat{\beta}$ (in a full-rank model)

(iii) $\rightarrow$ $\hat{\theta}_j | X \sim \mathcal{N}(\theta_j, \sigma^2 v_{jj})$

$v_{jj} = l_j^T (X^T X)^{-1} l_j \equiv$ j-th diagonal element of $V = L(X^T X)^{-1} L^T$

$$L = \begin{pmatrix} l_1^T \\ \vdots \\ l_m^T \end{pmatrix}$$

$$\frac{\hat{\theta}_j - \theta_j}{\sqrt{\sigma^2 v_{jj}}} \Big| X \sim \mathcal{N}(0,1)$$

$\rightarrow$ if σ^2 known, it could be used to construct test of $H_0: \theta_j = \theta_j^0$ or to construct confidence intervals

replace σ^2 by MSe

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

(viii) $T_j \sim t_{n-r}$
(even unconditionally)

PROOF $\rightarrow$ next page

$\boxed{7}$

$$T_j = \frac{\hat{\theta}_j - \theta_j}{\sqrt{MSe\, v_{jj}}} = \frac{\dfrac{\hat{\theta}_j - \theta_j}{\sqrt{\sigma^2 v_{jj}}}}{\sqrt{\dfrac{SSe}{\sigma^2}\Big/(n-r)}}$$

$\sim N(0,1)$ (given X)
(point iii)

independent (given X)
(point v)

$\sim \chi^2_{n-r}$ (given X)
(point vii)

$\Rightarrow \quad T_j \mid X \sim t_{n-r}$
this for almost all values of the condition $X = x$

$\Rightarrow \quad T_j \sim t_{n-r}$

Insertion 2 (see also Appendix on Distributions)

$$\left.\begin{array}{l} U \sim N(0,1) \\ V \sim \chi^2_\nu \end{array}\right. \perp\!\!\!\perp \quad \Rightarrow \quad \frac{U}{\sqrt{\frac{V}{\nu}}} \sim t_\nu$$

---

multivariate t distribution

$$\left.\begin{array}{l} U \sim N_p(0, \Sigma) \\ V \sim \chi^2_\nu \end{array}\right. \perp\!\!\!\perp \quad \Rightarrow \sqrt{\frac{\nu}{V}} \cdot U \sim mvt_{p,\nu}(\Sigma)$$

scale matrix

Easy to see:

- $mvt_{1,\nu}(1) \equiv t_\nu$
- $\sigma_1^2, \dots, \sigma_p^2 > 0$ diagonal elements of Σ

  $T = (T_1, \dots, T_p)^T \sim mvt_{p,\nu}(\Sigma)$

  $$\Rightarrow \frac{T_j}{\sigma_j} \sim t_\nu, \quad j = 1, \dots, p$$

  (margins of mvt are (univariate) student t)

We know (iii)

$$\hat{\theta} \mid X \sim N_m(\theta, \sigma^2 V)$$

$$\theta = L\beta$$

$$\hat{\theta} = L\hat{\beta}$$

$$V = L(X^TX)^{-1}L^T = (v_{j,t})_{j,t=1,\dots,m}$$

$$D := \operatorname{diag}\left(\frac{1}{\sqrt{v_{11}}}, \dots, \frac{1}{\sqrt{v_{m,m}}}\right)$$

$$\frac{1}{\sqrt{\sigma^2}} D(\hat{\theta} - \theta) = \left(\frac{\hat{\theta}_j - \theta_j}{\sqrt{v_{jj}}} \cdot \frac{1}{\sqrt{\sigma^2}}\right)_{j=1,\dots,m}$$

as vector $\underset{(iii)}{\sim} N_m(0, DVD)$ (given X)

$$\frac{1}{\sqrt{MSe}} D(\hat{\theta} - \theta) = \Bigg(\underbrace{\frac{\hat{\theta}_j - \theta_j}{\sqrt{MSe\, v_{jj}}}}_{T_j \sim t_{n-r} \text{ (point viii)}}\Bigg)_{j=1,\dots,m} =: T$$

as vector $\sim$?

(ix) $T \mid X \sim mvt_{m,\,n-r}(DVD)$

PROOF $\rightarrow$ next page

$$T = \frac{1}{\sqrt{MSe}} \mathbb{D}(\hat{\theta} - \theta)$$

$$= \sqrt{\frac{(n-r)\,\sigma^2}{SSe}}\; \underbrace{\mathbb{D}(\hat{\theta}-\theta)\sqrt{\frac{1}{\sigma^2}}}_{\sim N_m(0,\, DVD)\ (\text{given } X)\ (\text{point iii})}$$

$\frac{(n-r)}{SSe}$: $\sim \chi^2_{n-r}$ (given X) (point vii)

point (v): $SSe \perp\!\!\!\perp \hat{\theta}$ (given X)

$$\Rightarrow T \mid X \sim mvt_{m,n-r}(DVD).$$

$\blacksquare$

(x) If $\operatorname{rank}(L_{m \times k}) = m \le r = k$, the matrix $V = L(X^TX)^{-1}L^T$ is invertible and

$$\frac{1}{m}(\hat{\theta}-\theta)^T(MSe\,V)^{-1}(\hat{\theta}-\theta) \sim F_{m,n-r}.$$

- The fact that V is invertible (under stated assumptions) was shown in the proof of Theorem 2.5 (Gauss-Markov for linear combinations).

- Further

$$Q := \frac{1}{m}(\hat{\theta}-\theta)^T(MSe\,V)^{-1}(\hat{\theta}-\theta) =$$

$$= \frac{\frac{1}{m}(\hat{\theta}-\theta)^T(\sigma^2V)^{-1}(\hat{\theta}-\theta)}{\frac{SSe}{\sigma^2(n-r)}}$$

given X: $(\hat{\theta}-\theta) \sim N_m(0, \sigma^2V)$ (point iii)

$$\Rightarrow (\hat{\theta}-\theta)^T(\sigma^2V)^{-1}(\hat{\theta}-\theta) \sim \chi^2_m$$

$\frac{SSe}{\sigma^2} \sim \chi^2_{n-r}$ (point vii)

$\hat{\theta} \perp\!\!\!\perp SSe$ (point v)

$\Rightarrow Q|X \sim F_{m,n-r}$

This for almost all values of the condition $X = x$

$\Rightarrow Q \sim F_{m,n-r}$.

WW

Consequence of Theorem 6.2:

Least squares estimator of the regression coefficients in a full-rank linear model

Let $Y|X \sim N_n(X\beta, \sigma^2 I_n)$, $\operatorname{rank}(X_{n\times k}) = k$.

Further, let $V = (X^T X)^{-1} = (v_{jt})_{j,t=0,\dots,k-1}$

$$D = \operatorname{diag}\left(\frac{1}{\sqrt{v_{00}}}, \dots, \frac{1}{\sqrt{v_{k-1,k-1}}}\right).$$

The following then holds:

(i) $\hat\beta | X \sim N_k(\beta, \sigma^2 V)$

(ii) Statistics $\hat\beta$ and SSe are conditionally, given X, independent.

(iii) For each $j = 0, \dots, k-1$, $T_j := \dfrac{\hat\beta_j - \beta_j}{\sqrt{MSe\, v_{jj}}} \sim t_{n-k}$

(iv) $T := (T_0, \dots, T_{k-1})^T = \dfrac{1}{\sqrt{MSe}} D(\hat\beta - \beta) \sim$

$$\sim mvt_{k,n-k}(DVD) \qquad (\text{given } X)$$

(v) $\dfrac{1}{k}(\hat\beta - \beta)^T MSe^{-1} X^T X (\hat\beta - \beta) \sim F_{k,n-k}$.

Proof: Theorem 6.2 with $L = I_k$.

6.2.1 Statistical inference in a full-rank lin. model

$$\hat\beta \mid X \sim N_k(\beta, \sigma^2 \underbrace{(X^T X)^{-1}}_{(v_{j,e})})$$

$$\Rightarrow \hat\beta_j \mid X \sim N(\beta, \sigma^2 v_{jj})$$

$$\Rightarrow \frac{\hat\beta_j - \beta_j}{\sqrt{\sigma^2 v_{jj}}} \cancel{\mid X} \sim N(0,1)$$

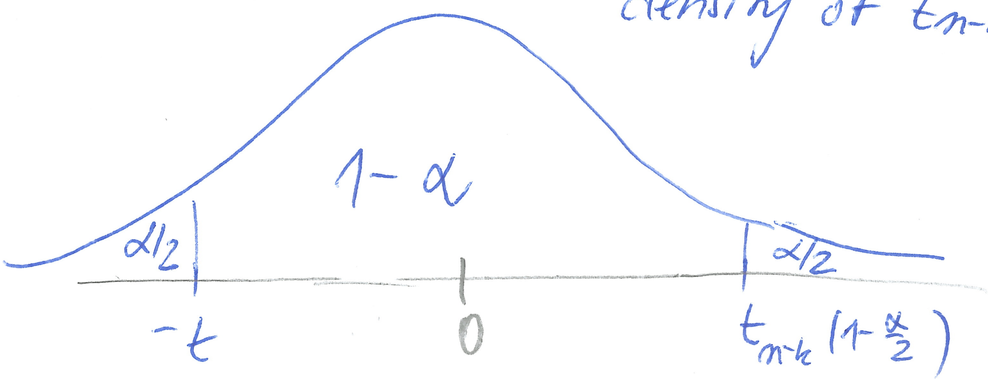
Proof is needed!

$$\Rightarrow \frac{\hat\beta_j - \beta_j}{\sqrt{MSe\, v_{jj}}} \sim t_{n-k}$$

$$S.E.(\hat\beta_j) = \sqrt{\widehat{var}(\hat\beta_j \mid X)}$$

Rest is standard (as in derivation of quantities related to one-sample t-test):

density of t_{n-k}



conf. interval:

$$\forall \beta_j^0 \in \mathbb{R} \quad P\left(\left| \frac{\hat\beta_j - \beta_j^0}{\sqrt{MSe\, v_{jj}}} \right| < t_{n-k}(1-\tfrac{\alpha}{2});\ \beta_j = \beta_j^0 \right) = 1-\alpha$$

$$\Rightarrow \forall \beta_j^0 \quad P\left(\left(\hat\beta_j \pm \sqrt{MSe\, v_{jj}}\, t_{n-k}(1-\tfrac{\alpha}{2})\right) \ni \beta_j^0;\ \beta_j = \beta_j^0 \right) = 1-\alpha.$$

test: $H_0 : \beta_j := \beta_j^0$, $\beta_j^0 \in \mathbb{R}$ chosen

under H_0 , $T_{j,0} := \dfrac{\hat{\beta}_j - \beta_j^0}{\sqrt{\text{mse}\, v_{jj}}} \sim t_{n-k}$

$$\Rightarrow P\left(|T_{j,0}| \geq t_{n-k}\left(1-\tfrac{\alpha}{2}\right) ; H_0 \right) = \alpha$$

Possible critical region is

$$C(\alpha) := \left(-\infty, -t_{n-k}\left(1-\tfrac{\alpha}{2}\right)\right] \cup \left[t_{n-k}\left(1-\tfrac{\alpha}{2}\right), \infty\right)$$

not the only one,

this one → largest power (in a certain sense)

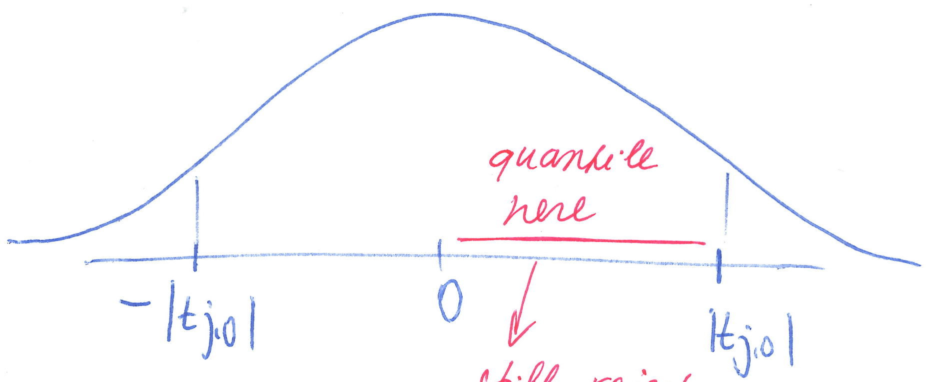
$$\Rightarrow P(\text{type I error}) = \alpha .$$

---

p-value? $t_{j,0}$ = value of $T_{j,0}$ with given data

$$p = \inf\{\alpha : t_{j,0} \in C(\alpha)\}$$

$$= \inf\{\alpha : |t_{j,0}| \geq t_{n-k}\left(1-\tfrac{\alpha}{2}\right)\}$$



still rejects H_0

„lowest" α if quantile $= |t_{j,0}|$

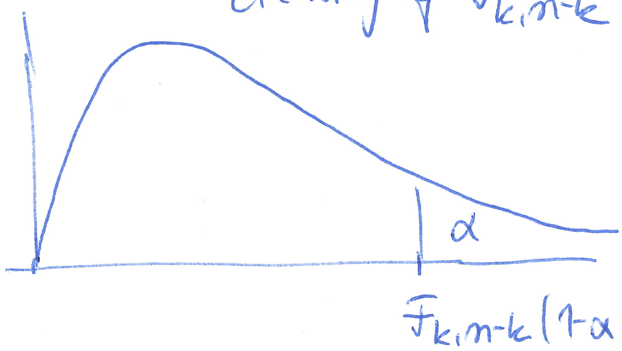
$$\Rightarrow p = 2\, CDF_{t,n-k}\left(-|t_{j,0}|\right)$$

Simultaneous inference on a vector β of regression coefficients

[10]

$$\forall \beta^0 \in \mathbb{R}^k \quad P\left(\frac{1}{k} (\hat\beta - \beta^0)^T (MS_e (X^TX)^{-1})^{-1} (\hat\beta - \beta^0) < F_{k,n-k}^{(1-\alpha)} ; \beta = \beta^0 \right) = 1-\alpha$$

density of $F_{k,n-k}$



→ Confidence region

$$S(\alpha) = \{ \beta \in \mathbb{R}^k : (\beta - \hat\beta)^T MS_e^{-1} X^T X (\beta - \hat\beta) < k\, F_{k,n-k}(1-\alpha) \}$$

<u>Test</u> $H_0: \beta = \beta^0$, $\beta^0 \in \mathbb{R}^k$ chosen

[11]

$H_1: \beta \neq \beta^0$

$$Q_0 := \frac{1}{k} (\hat\beta - \beta^0)^T MS_e^{-1} X^T X (\hat\beta - \beta^0) \underset{H_0}{\sim} F_{k,n-k}$$

<u>Hence</u> $P(Q_0 \geq F_{k,n-k}(1-\alpha); H_0) = \alpha$

→ <u>Possible</u> CRITICAL REGION

$$C(\alpha) = [F_{k,n-k}(1-\alpha); \infty)$$

→ P-value $p = 1 - CDF_{F,k,n-k}(q_0)$,

where q_0 = value of Q_0 with given data

Inference on a chosen linear combination

12

$$\theta = \ell^T \beta \quad , \quad \hat{\theta} = \ell^T \hat{\beta}$$

$$\forall \theta^0 \in \mathbb{R} \qquad T_0 := \frac{\hat{\theta} - \theta^0}{\sqrt{MSe\, \ell^T (X^T X)^{-1} \ell}} \sim t_{n-k}$$

Rest as before for β_j.

Simultaneous inference on a set of linear combinations

13 - 14

$$\theta = L\beta \quad , \quad \hat{\theta} = L\hat{\beta} \quad , \quad L_{m \times k}$$

$$\forall \theta^0 \in \mathbb{R}^m \qquad \frac{1}{m} (\hat{\theta} - \theta^0)^T (MSe\, L (X^T X)^{-1} L^T)^{-1} (\hat{\theta} - \theta^0)$$

$$\sim F_{m, n-k}$$

Rest as before for β.