Estimation of parameters

Fittede værdier/prædiktioner
$\hat{\mu}$ er projektion af y. Hvis X har fuld rang, så

$$\begin{aligned}\hat{\mu} &= X\hat{\beta} \\ &= X((X^TX)^{-1}X^Ty\end{aligned}$$

hvor $X(X^TX)^{-1}$ er en projektionsmatrix/hatmatrix.

Projektionsmatrix
P er en projektionsmatrix på et underrum $\Omega \subseteq \mathbb{R}^n$, hvis

(a) $P^T = P$ (symmetri)

(b) $P^2 = P$ (idempotent)

(c) $Py \in \Omega$, $\forall y \in \mathbb{R}^n$

(d) $y - Py \perp \Omega$, $\forall y \in \mathbb{R}^n$

H er en projektionsmatrix af y på Ω_0. **Residualer**

$$r(y) = y - \hat{\mu}(y) = y - X\hat{\beta} = y - Hy = (I - H)y$$

Cochrans sætning (modificeret)
Antag $Z \sim \mathcal{N}_n(0, I_n)$, $H_1, \ldots, H_m$ er projektionsmatricer, hvor $H_iH_j = 0$, når $i \neq j4$ og d_i =rang($H_i > 0$). Så er

$$\begin{aligned} H_iZ &\sim \mathcal{N}(0, H_i), \qquad && i = 1, \ldots, m \qquad \text{ufhængige} \\ \|H_iZ\|^2 = Z^TH_iZ &\sim \chi^2(d_i) && i = 1, \ldots, m \qquad \text{uafhængige}\end{aligned}$$

For $i \neq j$ er H_iZ og $\|H_jZ\|^2$ uafhængige. *Bevis*
Af spektralsætningen fås

- $H_i = A_iD_iA_i^{-1}$
- A_i er ortogonal (pga. H_i symmetrisk), dvs. $A_iA_i^T = A_i^TA_i = I$
- D_i er diagonal med d_i 1-taller i diagonalen og 0 ellers (pga. H_i er en projektionsmatrix)

Antag D_i har 0 på de første $d_1 + \dots d_{i-1}$ pladser, derefter 1 på de næste d_i pladser, og 0 derefter.
Lad

$$P = \begin{bmatrix} H_i \\ \vdots \\ H_m \end{bmatrix} \in \mathbb{R}^{mn \times n}$$

$$PZ \sim \mathcal{N}_{mn}(P0, PI_nP^T) = \mathcal{N}_{mn}(0, PP^T)$$

hvor

$$\begin{aligned} PP^T &= \begin{bmatrix} H_1H_1 & H_1H_2 & \dots & H_1H_m \\ H_2H_1 & & & \\ \vdots & & \ddots & \\ H_mH_1 & & & H_mH_n \end{bmatrix} \\ &= \begin{bmatrix} H_1 & & 0 \\ & \ddots & \\ 0 & & H_m \end{bmatrix} \end{aligned}$$

Desuden:

$$\|H_iZ\|^2 = (H_iZ)^T(H_iZ) = Z^TH_i^TH_iZ = Z^TH_iZ$$

Da

$$\begin{aligned} VertH_iZ\| &= \|A_iD_iA_i^TZ\| = \|D_iA_i^TZ\| \qquad (A_i \text{ bevarer længde) og} \\ D_iA_i^TZ &= \mathcal{N}_n(0, D_iA_i^T(D_iA_i)^T) \\ &= \mathcal{N}_n(0, D_i) \end{aligned}$$

fordelt som d_i uafhængige normalfordelinger og $n - d_i$ nuller. Dvs.

$$\|H_iZ\|^2 = \|D_iA_i^TZ\|^2 \sim \chi^2(d_i)$$

Korrolar

$$\begin{aligned} \hat{\mu}(Y) &\sim \mathcal{N}(\mu, \sigma^2H) \\ r(Y) &\sim \mathcal{N}(0, \sigma2(I - H)) \end{aligned}$$

Disse er uafhængige.

Bevis
Lad $H_1 = H$ og $H_2 = I - H$ og bemærk

$$H(I-H) = H - H^2 = H - H = 0$$

Desuden er $H\mu = \mu$ (fordi $\mu \in \Omega_O$) og $(I-H)\mu = \mu - H\mu = 0$. Så

$$\hat{\mu} = HY = H(\sigma Z + \mu) = \sigma HZ + H\mu \sim \mathcal{N}(\mu, \sigma^2 H)$$

og

$$r(Y) = (I-H)Y = (I-H)(\sigma Z + \mu) = \sigma(I-H)Z + (I-H)\mu \sim \mathcal{N}(0, (I-H)\sigma^2)$$

De to fordelinger er uafhængige ifølge Cochrans sætning.

Bemærk:

- Normalfordelingen holder for kendt σ^2.
- Residualer har forskellig varians.

Korrolar

$$\underbrace{D(Y; X\beta)}_{\sim\sigma^2\chi^2(n)} = \underbrace{\underbrace{D(Y; X\hat{\beta}(Y))}_{\sim\sigma^2\chi^2(n-k)} + \underbrace{D(X\hat{\beta}; X\beta)}_{\sim\sigma^2\chi^2(k)}}_{\text{uafhængige}}$$

Bevis
Lad $H_1 = H$ og $H_2 = I - H$. Så

$$D(Y; X\beta) = \|Y - \mu\|^2 = \|\underbrace{Y - HY}_{H_2 y} + \underbrace{HY - \mu}_{H_i((Y-\mu)}\|^2$$

Af Pythagoras:

$$D(Y, ; X\beta) = \|H_2 Y\|^2 + \|H_1(Y - \mu)\|^2$$

Dermed:

$$\begin{aligned} D(Y; X\hat{\beta}(Y)) &= \|Y - X\hat{\beta}(Y)\|^2 = \|Y - HY\|^2 = \|H_2 Y\|^2 \\ &= \sigma^2 \|H_2 Z\|^2 \sim \sigma^2\chi^2(n-k) \qquad \text{og} \\ D(X\hat{\beta}(Y); X\beta) &= \|X\hat{\beta}(Y) - X\beta\|^2 = \|HY - X\beta\|^2 = \|H_1(Y-\mu)\|^2 \\ & \sigma^2\|H_1 X\|^2 \sim \sigma^2\chi^2(k) \end{aligned}$$

Summen af to uafhængige chi^2-fordelte variable giver en chi^2-fordelt variabel med summen af de to fordelingers frihedsgrader.

Estimation af σ^2
Lad $Y \sim \mathcal{N}_n(\mu, \sigma^2 I)$ og X have fuld rang. For en observation $Y = y$, så eksisterer MLE for σ^2 hvis og kun hvis at $y \notin \Omega_0$.

$$\hat{\sigma}^2_{MLE} = \frac{D(y; X\hat{\beta}(y))}{n} = \frac{\|y - X\hat{\beta}(y)\|^2}{n} = \frac{\|y - Hy\|^2}{n}$$

Dette er ikke centralt. Der haves dog

$$\hat{\sigma}^2 = \frac{\|Y - X\hat{\beta}(Y)\|^2}{n-k} = \frac{\|Y - HY\|^2}{n-k}$$

som er centralt estimat for σ^2 og er uafhængigt af $\hat{\beta}(Y)$. Der gælder desuden

$$\hat{\sigma}^2(Y) \sim \frac{\sigma^2}{n-k}\chi^2(n-k)$$

Bevis
Indsæt $\hat{\beta}$ i likelihoodfunktionen for (β, σ^2):

$$\underbrace{L((\hat{\beta}(y), \sigma^2); y) = \frac{1}{(\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} D(y; X\hat{\beta}(y))\right)}_{Profillikelihood}$$

Kan vises, at $l((\hat{\beta}(y), \sigma^2); y)$ som funktion af σ^2

- har entydigt max i $D(y, X\hat{\beta}(y))/n$, hvis $y \notin \Omega_0$
- er $\log((\sigma^2)^{-n/2})$, hvis $y \in \Omega_0$ (dvs. intet max)

Dernæst

$$\hat{\sigma}^2(Y) = \frac{\|Y - X\hat{\beta}(Y)\|^2}{n-k} = \frac{D(Y; X\hat{\beta}(Y)}{n-k} \sim \frac{\sigma^2}{n-k}\chi^2(n-k)$$

Centralitet:

$$E[\hat{\sigma}^2(Y)] = \frac{\sigma^2}{n-k}E[\chi^2] = \sigma^2$$

Uafhængighed:

$$\hat{\sigma}^2(Y) = \frac{1}{n-k}\|H_2 Y\|^2$$
$$\hat{\beta}(Y) = H_1 Y$$

De er uafhængige pga. Cochrans sætning.