Preliminaries.

Sep. 18

1. Expectation, variance & covariance.

▸ Given r.v.'s X, Y, constants a, b

- Expectation $\mathbb{E}X$. $\quad \mathbb{E}(aX+b) = a\mathbb{E}X + b$.
- Variance $\quad Var X = \mathbb{E}(X-\mathbb{E}X)^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2$

$$Var(aX+b) = a^2 Var X.$$

- Covariance. $\quad Cov(X,Y) = \mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)] = \mathbb{E}(XY) - \mathbb{E}X \cdot \mathbb{E}Y$.

Theorem 3.1: Let $Y_1, \ldots, Y_n$ be uncorrelated random variables with $Var(Y_i) = \sigma^2$ for all $i = 1, \ldots, n$. Let $c_1, \ldots, c_n$ and $d_1, \ldots, d_n$ be two sets of constants. Then

$Cov(Y_i, Y_j) = 0,\ i \neq j$

$$Cov\left(\sum_{i=1}^n c_i Y_i, \sum_{i=1}^n d_i Y_i\right) = \left(\sum_{i=1}^n c_i d_i\right)\sigma^2$$

Pf: $LHS \overset{def}{=} \mathbb{E}\left[(\underbrace{\textstyle\sum c_i Y_i - \mathbb{E}\sum c_i Y_i}_{\sum c_i (Y_i - \mathbb{E}Y_i)})(\textstyle\sum d_i Y_i - \mathbb{E}\sum d_i Y_i)\right]$

$$= \mathbb{E}[(\textstyle\sum c_i \tilde{Y}_i)(\sum d_i \tilde{Y}_i)], \qquad \tilde{Y}_i := Y_i - \mathbb{E}Y_i.$$

$$= \mathbb{E}[\textstyle\sum_{i,j=1}^n c_i d_j \tilde{Y}_i \tilde{Y}_j]$$

$$= \mathbb{E}\textstyle\sum_i c_i d_i \tilde{Y}_i^2 + \mathbb{E}\sum_{i\neq j} c_i d_j \tilde{Y}_i \tilde{Y}_j$$

$$= \textstyle\sum_i c_i d_i \underbrace{\mathbb{E}\tilde{Y}_i^2} + \sum_{i \neq j} c_i d_j \underbrace{\mathbb{E}\tilde{Y}_i\tilde{Y}_j} \longrightarrow = Cov(Y_i, Y_j) = 0 \quad i \neq j$$

$$= \textstyle\sum_i c_i d_i \sigma^2 \qquad \mathbb{E}(Y_i - \mathbb{E}Y_i)^2 = Var Y_i = \sigma^2$$

▸ Given r.vec's $X, Y \in \mathbb{R}^d$, where $X = (X_1, \cdots, X_d)^T$, const vec. $a, b \in \mathbb{R}^{d\times 1}$.

- Expectation $\quad \mathbb{E}X = (\mathbb{E}X_1, \cdots, \mathbb{E}X_d)^T$.
- Covariance matrix.

Def. $\quad Cov(X,Y) = \mathbb{E}[\underbrace{(X-\mathbb{E}X)}_{d\times 1}\underbrace{(Y-\mathbb{E}Y)^T}_{1\times d}] \quad d\times d$

$$Cov(X) = Cov(X,X) = \mathbb{E}[(X-\mathbb{E}X)(X-\mathbb{E}X)^T]$$

Property $\quad Cov(a^T X, b^T Y) = a^T Cov(X,Y)\, b. \quad 1\times 1$

2. Normal, χ^2, t, F distributions

▷ Constructions / Relations

① $X_1, X_2, \cdots, X_r \overset{iid}{\sim} N(0,1)$. Then $Z = \sum_{i=1}^{r} X_i^2 \sim \chi^2(r)$

② $X \sim N(0,1) \perp\!\!\!\perp \underset{}{Z}^{=\chi^2} \sim \chi^2(r)$ Then $T = \dfrac{X}{\sqrt{Z/r}} \sim t(r)$

③ $X_1 \sim \chi^2(r_1) \perp\!\!\!\perp X_2 \sim \chi^2(r_2)$. Then $F = \dfrac{X_1/r_1}{X_2/r_2} \sim F(r_1, r_2)$

▷ Additivity of indpt χ^2.

Assume $Z_1 \sim \chi^2(r_1) \perp\!\!\!\perp Z_2 \sim \chi^2(r_2)$. Then $Z_1 + Z_2 \sim \chi^2(r_1 + r_2)$

Pf: Let $X_1, \cdots, X_{r_1+r_2} \overset{iid}{\sim} N(0,1)$.

$Z_1 \overset{d}{=} X_1 + \cdots + X_{r_1}$, $Z_2 \overset{d}{=} X_{r_1+1} + \cdots + X_{r_1+r_2}$.

Then $Z_1 + Z_2 \overset{d}{=} X_1 + \cdots + X_{r_1+r_2} \sim \chi^2(r_1+r_2)$.

▷ Thm. 3.4 Let $X \sim N(0, I_n)$. A is a symmetric idempotent matrix of rank k. Then $X^T A X \sim \chi^2(k)$.

Review

1. Regression model

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i \quad , \quad i = 1, \cdots, n$$

(matrix form) $Y = X\beta + e$.

randomness: $\underset{\Delta}{e} \longrightarrow Y$. X is deterministic.

with some assumptions e_i's $\overset{iid}{\sim} N(0, \sigma^2)$

2. Estimation.

2.1 β.

Method: LS (= MLE in Gaussian cases)

Optimization: (objective)

$$\hat{\beta} = \underset{\beta'}{\operatorname{argmin}} \sum (y_i - (\beta'_0 + \beta'_1 x_{i1} + \cdots + \beta'_p x_{ip}))^2 = \sum \hat{e}_i^2 = RSS$$

$$= \underset{\beta'}{\operatorname{argmin}} (Y - X\beta)^T (Y - X\beta).$$

Solution by first order conditions: $\hat{\beta} = (X^T X)^{-1} X^T Y$

- Case of $p=1$:

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}_1 \\ \hat{\beta}_1 = \dfrac{\sum_{i=1}^{n} (x_{i1} - \bar{x}_1)(y_i - \bar{y})}{\sum_{i=1}^{n} (x_{i1} - \bar{x}_1)^2} = \dfrac{S_{x_1 y}}{S_{x_1 x_1}} \end{cases}$$

- Centered model.

2.2. σ^2

$$\hat{\sigma}^2_{MLE} = \frac{RSS}{n}$$

$$\hat{\sigma}^2_{UE} = \frac{RSS}{n-p'}$$ (by the properties of RSS in Sec. 2.3)

- Specific forms for different cases of $p=1$, $p \geq 1$.