VII Chi-Square Distributions

- A random variable Y is called chi-square distribution with degree of freedom k and is denoted by $\chi^2(k)$ if one of the following equivalent conditions holds

(i) $Y \overset{d}{=} Z_1^2 + \cdots + Z_k^2$,

where $Z_i \overset{iid}{\sim} N(0,1)$.

(ii) $Y \sim \text{Gamma}(\frac{k}{2}, 2)$

(iii) The probability density function of Y is

$$P_Y(y) = \frac{1}{\Gamma(\frac{k}{2})\, 2^{\frac{k}{2}}} \times y^{\frac{k}{2}-1} e^{-\frac{y}{2}}, \quad y > 0$$

(iv) The moment generating function of Y is

$$M_Y(u) = (1-2u)^{-\frac{k}{2}}, \quad u < \frac{1}{2}$$

- A random variable Y is called non-central chi-square distribution with degree of freedom k and non-centrality parameter $\delta > 0$

and is denoted by $\chi^2(k;\underline{\delta})$ if

one of the following equivalent conditions holds

(i) $Y \overset{d}{=} Z_1^2 + \cdots + Z_k^2$,

where $Z_i \sim N(\mu_i, 1)$ and $\sum_{i=1}^{k} \mu_i^2 = \delta$

(ii) $Y \overset{d}{=} V + Z^2$,

where $V \sim \chi^2(k-1)$ and $Z \sim N(\mu, 1)$

are independent and $\mu^2 = \delta$.

(iii) The moment generating function of Y is

$$M_Y(u) = (1-2u)^{-\frac{k}{2}} \times \exp\left(\frac{\delta u}{1-2u}\right), \quad u < \frac{1}{2}.$$

* $\mathbb{E}[Y] = k + \delta$, $Var(Y) = 2k + 4\delta$

* Suppose $Z \sim N_d(\mu, I)$ and

A is a symmetric matrix satisfying

$A^2 = A$, $tr(A) = k$, and $\mu^T A \mu = \delta$.

Then, $Z^T A Z \sim \chi^2(k;\delta)$.

(pf) Consider the spectral decomposition of A:

$$A = PDP^T,$$

where P is an orthogonal matrix ($PP^T = I = PP^T$) and D is a diagonal matrix, $D = \mathrm{diag}(\lambda_i)$.

Since $A^2 = A$, $\lambda_i = 0$ or 1 for $i = 1, \cdots, d$.

Also, since $\mathrm{tr}(A) = k$,

$$|\{i : \lambda_i = 1\}| = k.$$

WLOG, let $\lambda_1 = \cdots = \lambda_k = 1$ and

$$\lambda_{k+1} = \cdots = \lambda_d = 0.$$

Note that

$$Z^T A Z = Z^T P D P^T Z$$

$$= \sum_{i=1}^{d} \lambda_i (P^T Z)_i^2 = \sum_{i=1}^{k} (P^T Z)_i^2.$$

Because

$$P^T Z \sim N_d(P^T \mu, I) \text{ and}$$

$$\sum_{i=1}^{k} (P^T \mu)_i^2 = \mu^T P D P^T \mu = \mu^T A \mu = \delta,$$

We can derive that

$$Z^T A Z \sim \chi^2(k;\delta).$$

$\boxed{\text{VL}}$

$\boxed{\text{III}}$ t distributions

- A random variable Y is called t distribution with degree of freedom k and is denoted by $t(k)$ if

$$Y \overset{d}{=} \frac{Z}{\sqrt{V/k}},$$

where $Z \sim N(0,1)$ and $V \sim \chi^2(k)$ are independent

- A random variable Y is called non-central t distribution with degree of freedom k and non-centrality parameter $\delta > 0$ and is denoted by $t(k;\delta)$ if

$$Y \overset{d}{=} \frac{Z}{\sqrt{V/k}},$$

where $Z \sim N(\delta, 1)$ and $V \sim \chi^2(k)$ are independent

[iii] F distributions

- A random variable Y is called F distribution with degree of freedom k_1 and k_2 and is denoted by $F(k_1, k_2)$ if

$$Y \overset{d}{=} \frac{V_1 / k_1}{V_2 / k_2},$$

where $V_1 \sim \chi^2(k_1)$ and $V_2 \sim \chi^2(k_2)$ are independent

- A random variable Y is called non-central F distribution with degree of freedom k_1 and k_2 and non-centrality parameter $\delta > 0$ and is denoted by $F(k_1, k_2 ; \delta)$ if

$$Y \overset{d}{=} \frac{V_1 / k_1}{V_2 / k_2},$$

where $V_1 \sim \chi^2(k_1 ; \delta)$ and

$V_2 \sim \chi^2(k_2)$ are independent.

例 Example $Y = X\beta + \varepsilon \quad \varepsilon \sim N(0, \sigma^2 I)$

$H_0: \beta_1 = \cdots = \beta_k = 0$

$H_1:$ not H_0

$$SSM = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

under H_0,

$$l(\beta, \sigma^2; Y, X) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{\|y - \beta_0 \cdot \mathbb{1}\|_2^2}{2\sigma^2}$$

is maximized when $\hat{\beta}_0 = \bar{y}$, $\hat{\sigma}^2 = \dfrac{\|y - \bar{y}\cdot\mathbb{1}\|_2^2}{n}$

$$\therefore \max_{H_0} l(\beta, \sigma^2 | Y, X) = -\frac{n}{2}\Big(1 + \log 2\pi/n + \log \underbrace{\|y - \bar{y}\mathbb{1}\|_2^2}_{SST}\Big)$$

under H_1,

$l(\beta, \sigma^2; Y, X)$ is maximized when $\hat{\beta} = \hat{\beta}^{OLS} = (X^T X)^{-1} X^T y$

$$\hat{\sigma}^2 = \frac{\|y - X\hat{\beta}\|_2^2}{n}$$

and $$\max_{H_1} l(\beta, \sigma^2 | Y, X) = -\frac{n}{2}\Big(1 + \log\frac{2\pi}{n} + \log \underbrace{\|Y - X\hat{\beta}\|_2^2}_{SSE}\Big)$$

$\therefore$ maximum likelihood ratio test

$=$ reject when $\log {}^{SST}/_{SSE}$ is large.

let $\Pi_1 = \mathbb{1}(\mathbb{1}^t\mathbb{1})^{-1}\mathbb{1}^t$: rk 1

$\Pi_X = X(X^tX)^{-1}X^t$: rk $k+1$

$$SST = \|y - \bar{y}\mathbb{1}\|_2^2 = \|(I-\Pi_1)y\|_2^2 = y^T(I-\Pi_1)y$$

$$SSE = \|y - X\hat{\beta}\|_2^2 = \|(I-\Pi_X)y\|_2^2 = y^T(I-\Pi_X)y$$

$$SSM = \|X\hat{\beta} - \bar{y}\mathbb{1}\|_2^2 = \|(\Pi_X - \Pi_1)y\|_2^2 = y^T(\Pi_X - \Pi_1)y$$

$\therefore SST = SSE + SSM$.

$\therefore$ m.l.r $\iff$ reject when $\dfrac{SSM/k}{SSE/n-k-1}$ is large

$$(I-\Pi_X)y = (I-\Pi_X)(X\beta+\varepsilon) = (I-\Pi_X)\varepsilon$$

$$\therefore \frac{SSE}{\sigma^2} \sim \chi^2(n-k-1)$$

$$(\Pi_X-\Pi_1)y = (X-\Pi_1X)\beta + (\Pi_X-\Pi_1)\varepsilon$$

$\therefore \dfrac{SSM}{\sigma^2} \sim \chi^2(k;\nu)$ for some ν ($\nu=0$ when H_0)

also $(I-\Pi_X)(\Pi_X-\Pi_1)=0$ $\therefore$ $SSE \perp SSM$

thus $$\frac{SSM/k}{SSE/n-k-1} \sim F(k, n-k-1 : \nu)$$

level α - MLR :

reject when $$\frac{SSM/k}{SSE/n-k-1} \geq F_\alpha(k, n-k-1)$$