I

Sampling from the Normal Distribution

Thm: $X_1, \cdots, X_n$ is a random sample from Gaussian distribution $N(\mu, \sigma^2)$, sample mean $\bar{X} \equiv \frac{1}{n}\sum_{i=1}^{n} X_i$, sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Then: (a) $\bar{X} \perp\!\!\!\perp S^2$;

(b) $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$;

(c) $\frac{n-1}{\sigma^2} S^2 \sim \chi^2_{n-1}$.

proof: (a) $\mathbb{1}$ is the one vector, then

$$\bar{X} = \frac{1}{n}\mathbb{1}^T \underline{X}, \quad \underline{X}_0 = (X_1 - \bar{X}, \cdots, X_n - \bar{X})^T$$
$$= \underline{X} - \mathbb{1}\bar{X}$$
$$= \underline{X} - \frac{1}{n}\mathbb{1}\mathbb{1}^T \underline{X}$$
$$= P\underline{X}$$

where $P = I - \frac{1}{n}\mathbb{1}\mathbb{1}^T$

We can see that, $P = P^T$, $P^2 = P$.

$\therefore \mathrm{Cov}[\bar{X}, \underline{X}_0]$

$= \mathrm{Cov}[\frac{1}{n}\mathbb{1}^T\underline{X}, P\underline{X}]$

$= \frac{1}{n}\mathbb{1}^T \mathrm{Var}\underline{X}\ P^T$

$= \frac{\sigma^2}{n}\mathbb{1}^T(I - \frac{1}{n}\mathbb{1}\mathbb{1}^T)$

$= \underline{0}$

$\because \bar{X}$ and $\underline{X}_0$ are Gaussian

$\therefore \bar{X} \perp\!\!\!\perp \underline{X}_0$

$\therefore \bar{X} \perp\!\!\!\perp S^2 = g(\underline{X}_0)$

(b) This is easy to see.

(c) $$\frac{n-1}{\sigma^2}S^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

$$= \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma} - \frac{\bar{X}-\mu}{\sigma}\right)^2$$

$$= \sum_{i=1}^{n}(Z_i-\bar{Z})^2 \qquad \left(Z_i = \frac{X_i-\mu}{\sigma}\right)$$

$$= \underline{Z}_0^T\underline{Z}_0 \qquad (\underline{Z}_0 = (Z_1-\bar{Z}, \cdots, Z_n-\bar{Z})^T)$$

$$= \underline{Z}^T P^T P \underline{Z} \qquad (\text{same } P \text{ in (a)})$$

$$= \underline{Z}^T P \underline{Z} \qquad (P=P^T,\ P=P^2)$$

$$= \underline{Z}^T U^T \Lambda U \underline{Z} \qquad (P \text{ symmetric})$$

$$=_d \underline{Z}^T \Lambda \underline{Z} \qquad (U \text{ unitary, } \underline{Z} \text{ standard Gaussian})$$

$$= \sum_{i=1}^{n} \lambda_i Z_i^2$$

$$= \sum_{i=1}^{n-1} Z_i^2 \qquad \left(\begin{array}{l}\lambda_i = 0 \text{ or } 1, \text{ and } \operatorname{tr} P = n-1 \\ \therefore \lambda_1 = \cdots = \lambda_{n-1} = 1,\ \lambda_n = 0\end{array}\right)$$

$$\sim \chi^2_{n-1}$$

$\square$.

Def: $X_1, \cdots, X_n$ is a random sample from distribution $N(\mu, \sigma^2)$, the distribution of $T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ is called <u>Student's t distribution</u> with $n-1$ degrees of freedom, denoted as $T \sim t_{n-1}$.

Equivalently, r.v. $T \sim t_p$ if

$$f_T(t) = \frac{\Gamma(\frac{p+1}{2})}{\Gamma(\frac{p}{2})} \frac{1}{(p\pi)^{\frac{1}{2}}} \frac{1}{(1+\frac{t^2}{p})^{\frac{p+1}{2}}} \qquad (-\infty < t < +\infty)$$

proof of equivalence:

Let $U = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$, $V = \frac{n-1}{\sigma^2}S^2$, then $T = \frac{U}{\sqrt{\frac{V}{n-1}}}$

Let $p = n-1$, then

$U \sim N(0,1)$, $V \sim \chi_p^2$, and $U \perp\!\!\!\perp V$

$\therefore f_{U,V}(u,v) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \frac{1}{\Gamma(\frac{p}{2})\, 2^{\frac{p}{2}}} v^{\frac{p}{2}-1} e^{-\frac{v}{2}}$

Make transformation $(T, W) = \left(\frac{U}{\sqrt{\frac{V}{p}}}, V\right)$, then

$$\left| J_{(T,W)}(U,V) \right| = \sqrt{\frac{W}{p}} \quad \text{, and it's bijective}$$

$\therefore f_{T,W}(t,w) = f_{U,V}\left(t\sqrt{\frac{w}{p}}, w\right) \sqrt{\frac{w}{p}}$

$\therefore$ Marginal pdf

$$f_T(t) = \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma(\frac{p}{2})\, 2^{\frac{p}{2}}} e^{-\frac{1}{2}t^2\frac{w}{p}} w^{\frac{p}{2}-1} e^{-\frac{w}{2}} \left(\frac{w}{p}\right)^{\frac{1}{2}} dw$$

$$= \frac{1}{\sqrt{2\pi}\, \Gamma(\frac{p}{2})\, 2^{\frac{p}{2}} p^{\frac{1}{2}}} \int_0^{+\infty} e^{-\frac{1}{2}(1+\frac{t^2}{p})w} w^{\frac{p+1}{2}-1} dw$$

$$= \frac{1}{\sqrt{p\pi}\, 2^{\frac{p+1}{2}}\, \Gamma(\frac{p}{2})} \cdot \Gamma\left(\frac{p+1}{2}\right) \left(\frac{2}{1+\frac{t^2}{p}}\right)^{\frac{p+1}{2}}$$

$$= \frac{\Gamma(\frac{p+1}{2})}{\Gamma(\frac{p}{2})} \frac{1}{\sqrt{p\pi}} \frac{1}{\left(1+\frac{t^2}{p}\right)^{\frac{p+1}{2}}}$$

$\square$.

Def:

$X_1, \cdots, X_n$ is a random sample from a $N(\mu_X, \sigma_X^2)$ population, and $Y_1, \cdots, Y_m$ is a random sample from an independent $N(\mu_Y, \sigma_Y^2)$ population. The distribution of $F = \frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2}$ is called <u>Snedecor's F distribution</u> with $n-1$ and $m-1$ degrees of freedom.

Equivalently, r.v. $F \sim F_{p,q}$ if

$$f_F(x) = \frac{\Gamma(\frac{p+q}{2})}{\Gamma(\frac{p}{2})\, \Gamma(\frac{q}{2})} \left(\frac{p}{q}\right)^{\frac{p}{2}} \frac{x^{\frac{p}{2}-1}}{\left(1+\frac{p}{q}x\right)^{\frac{p+q}{2}}} \qquad (0 < x < +\infty)$$

Thm: a) If $X \sim F_{p,q}$, then $\frac{1}{X} \sim F_{q,p}$

b) If $X \sim t_q$, then $X^2 \sim F_{1,q}$

c) If $X \sim F_{p,q}$, then $\dfrac{\frac{p}{q}X}{1+\frac{p}{q}X} \sim B(\frac{p}{2}, \frac{q}{2})$