Some Results on Independent and Identically Distributed Normal RVs

Statistical Inference and Multivariate Analysis (MA324)

IITG: 2023

Theorem .1. *Let $X_1, X_2, \ldots, X_n$ be i.i.d. $N(0,1)$ random variables. Then*

$$\sum_{i=1}^{n} X_i^2 \sim Gamma\left(\frac{n}{2}, \frac{1}{2}\right) \equiv \chi_n^2.$$

Proof: The MGF of X_1^2 is given by

$$M_{X_1^2}(t) = E\left(e^{tX_1^2}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\left(\frac{1}{2}-t\right)x^2} dx = (1-2t)^{-\frac{1}{2}},$$

for $t < \frac{1}{2}$. Hence, the MGF of $T = \sum_{i=1}^{n} X_i^2$

$$M_T(t) = \prod_{i=1}^{n} M_{X_i^2}(t) = (1-2t)^{-\frac{n}{2}},$$

where $t < \frac{1}{2}$. Thus, $T = \sum_{i=1}^{n} X_i^2 \sim Gamma(\frac{n}{2}, \frac{1}{2})$. This distribution is also known as χ^2 distribution with degrees of freedom n. Thus, the sum of squares of n *i.i.d.* $N(0,\,1)$ RVs has a χ^2 distribution with degrees of freedom n. $\square$

Theorem .2. *Let $X_1, X_2, \ldots, X_n$ be i.i.d. $N(\mu, \sigma^2)$ random variables. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$. Then $\overline{X}$ and S^2 are independently distributed and*

$$\overline{X} \sim N(\mu,\, \sigma^2/n) \quad \text{and} \quad \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2.$$

Proof: Let A be an $n \times n$ orthogonal matrix, whose first row is

$$\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}\right).$$

Note that such a matrix exists as we can start with the row and construct a basis of $\mathbb{R}^n$. Then Gram-Schmidt orthogonalization will give us the required matrix. As A is orthogonal, its' inverse exists and $A^{-1} = A^T$, the transpose of A. Now, consider the transformation of random vector $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)'$ given by

$$\boldsymbol{Y} = A\boldsymbol{X}.$$

First, we shall try to find the distribution of $\boldsymbol{Y}$. Note that the transformation $g(\boldsymbol{x}) = A\boldsymbol{x}$ is a one-to-one transformation as A is invertible. The inverse transformation is given by $\boldsymbol{x} = A'\boldsymbol{y}$. Hence, the Jacobian of the inverse transformation is $J = det(A)$. As A is orthogonal, absolute value of $det(A)$ is one. Now, as $X_1, X_2, \ldots, X_n$ are *i.i.d.* $N(\mu, \sigma^2)$ RVs, the JPDF of $\boldsymbol{X}$, for $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)' \in \mathbb{R}^n$, is

$$\begin{aligned} f_{\boldsymbol{X}}(\boldsymbol{x}) &= \frac{1}{\left(\sigma\sqrt{2\pi}\right)^n} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right] \\ &= \frac{1}{\left(\sigma\sqrt{2\pi}\right)^n} \exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{x} - \boldsymbol{\mu})'(\boldsymbol{x} - \boldsymbol{\mu})\right], \end{aligned}$$

where $\boldsymbol{\mu} = (\mu, \mu, \ldots, \mu)'$ is a n component vector. Thus, the JPDF of $\boldsymbol{Y}$, for $\boldsymbol{y} \in \mathbb{R}^n$, is

$$\begin{aligned} f_{\boldsymbol{Y}}(\boldsymbol{y}) &= f_{\boldsymbol{X}}(A'\boldsymbol{y}) \\ &= \frac{1}{\left(\sigma\sqrt{2\pi}\right)^n} \exp\left[-\frac{1}{2\sigma^2}(A'\boldsymbol{y} - \boldsymbol{\mu})'(A'\boldsymbol{y} - \boldsymbol{\mu})\right] \\ &= \frac{1}{\left(\sigma\sqrt{2\pi}\right)^n} \exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{y} - \boldsymbol{\eta})'(\boldsymbol{y} - \boldsymbol{\eta})\right], \end{aligned}$$

where $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_n)' = A\boldsymbol{\mu}$. Note that $\eta_1 = \sqrt{n}\mu$. Moreover,

$$\boldsymbol{\eta}'\boldsymbol{\eta} = \boldsymbol{\mu}'\boldsymbol{\mu} \implies \sum_{i=1}^{n}\eta_i^2 = n\mu^2 \implies \sum_{i=2}^{n}\eta_i^2 = n\mu^2 - \eta_1^2 = 0.$$

Thus, $\eta_i = 0$ for $i = 2, 3, \ldots, n$. Hence, the JPDF of $\boldsymbol{Y}$ is

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y_1 - \sqrt{n}\mu)^2} \left\{\prod_{i=2}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{y_i^2}{2\sigma^2}}\right\} \quad \text{for } \boldsymbol{y} = (y_1, y_2, \ldots, y_n)' \in \mathbb{R}^n.$$

Therefore, $Y_1, Y_2, \ldots, Y_n$ are independent RVs and $Y_1 \sim N(\sqrt{n}\mu, \sigma^2)$ and $Y_i \sim N(0, \sigma^2)$ for $i = 2, 3, \ldots, n$, where $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)'$. Now,

$$Y_1 = \sqrt{n}\,\overline{X} \implies \sqrt{n}\,\overline{X} \sim N(\sqrt{n}\mu, \sigma^2) \implies \overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

Again,

$$\boldsymbol{Y}'\boldsymbol{Y} = \boldsymbol{X}'\boldsymbol{X} \implies \sum_{i=2}^{n} Y_i^2 = \sum_{i=1}^{n} X_i^2 - Y_1^2 = \sum_{i=1}^{n} X_i^2 - n\overline{X}^2 = (n-1)S^2.$$

For $i = 2, 3, \ldots, n$, $\frac{Y_i}{\sigma}$ are *i.i.d.* $N(0, 1)$ RVs. Thus, using the previous theorem

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{i=2}^{n}\left(\frac{Y_i}{\sigma}\right)^2 \sim \chi^2_{n-1}.$$

Notice that $\overline{X}$ is a function of Y_1 only, and S^2 is a function of $Y_2, Y_3, \ldots, Y_n$. As Y_i's are independent, $\overline{X}$ and S^2 are independent. $\square$

Definition .1 (t-distribution)**.** *A CRV X is said to have a Student's t-distribution (or simply, t-distribution) with n degrees of freedom if the PDF of X is given by*

$$f(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}\left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} \quad \text{for } t \in \mathbb{R}.$$

We will use the notation $X \sim t_n$ to denote that the RV X has a t-distribution with n degrees of freedom.

Theorem .3. *Let $X \sim N(0, 1)$ and $Y \sim \chi^2_n$ be two independent RVs. Then the RV $T = \dfrac{X}{\sqrt{Y/n}} \sim t_n$.*

Proof: This theorem can be proved using the transformation technique 2. Note that the JPDF of X and Y is

$$f_{X,\,Y}(x,\,y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \times \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} y^{\frac{n}{2}-1} e^{-\frac{y}{2}} \qquad \text{for } x \in \mathbb{R},\, y > 0.$$

Take $V = \sqrt{\frac{Y}{n}}$. Then the inverse mapping is $x = tv$ and $y = nv^2$. The Jacobian of the transformation is

$$J = \begin{vmatrix} v & t \\ 0 & 2nv \end{vmatrix} = 2nv^2 > 0.$$

Thus, the JPDF of T and V is

$$f_{T,\,V}(t,\,v) = \frac{n^{\frac{n}{2}}}{2^{\frac{n-1}{2}}\sqrt{\pi}\Gamma(\frac{n}{2})} v^n e^{-\frac{1}{2}nv^2\left(1+\frac{t^2}{n}\right)} \qquad \text{for } t \in \mathbb{R},\, v > 0.$$

Therefore, for $t \in \mathbb{R}$, the marginal PDF of T is

$$\begin{aligned} f_T(t) &= \int_0^\infty f_{T,\,V}(t,\,v)dv \\ &= \frac{n^{\frac{n}{2}}}{2^{\frac{n-1}{2}}\sqrt{n}\Gamma(\frac{n}{2})} \int_0^\infty v^n e^{-\frac{1}{2}nv^2\left(1+\frac{t^2}{n}\right)} dv \\ &= \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}. \end{aligned}$$

□

Corollary .1. *Let $X_1,\, X_2,\, \ldots,\, X_n$ be i.i.d. $N(\mu, \sigma^2)$ random variables. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$. Then*

$$\sqrt{n}\,\frac{\overline{X} - \mu}{S} \sim t_{n-1},$$

where S is the positive square root of S^2.

Proof: From Theorem .2, it is clear that $\sqrt{n}\,\frac{\overline{X}-\mu}{\sigma} \sim N(0,\,1)$. Therefore,

$$\frac{\sqrt{n}\,\frac{\overline{X}-\mu}{\sigma}}{\sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}}} = \sqrt{n}\,\frac{\overline{X} - \mu}{S} \sim t_{n-1}.$$

□

Definition .2 (F-distribution)**.** *A CRV X is said to have a F-distribution with n and m degrees of freedom if the PDF of X is given by*

$$f(x) = \frac{1}{B\left(\frac{n}{2},\, \frac{m}{2}\right)} \left(\frac{n}{m}\right)^{\frac{n}{2}} x^{\frac{n}{2}-1} \left(1 + \frac{n}{m}x\right)^{-\frac{n+m}{2}} \qquad \textit{for } x > 0.$$

We will use the notation $X \sim F_{n,m}$ to denote that the RV X has a F-distribution with n and m degrees of freedom.

Theorem .4. *Let $X \sim \chi^2_n$ and $Y \sim \chi^2_m$ are two independent RVs. Then*

$$F = \frac{X/n}{Y/m} = \frac{mX}{nY} \sim F_{n,\,m}.$$

Proof: The JPDF of X and Y is

$$f_{X,Y}(x, y) = \frac{1}{2^{\frac{m+n}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})} x^{\frac{n}{2}-1} y^{\frac{m}{2}-1} e^{-\frac{1}{2}(x+y)} \quad \text{for } x > 0,\, y > 0.$$

Taking $V = Y$, the inverse transformation is $x = \frac{n}{m} f v$ and $y = v$. The Jacobian of the inverse transformation is

$$J = \begin{vmatrix} \frac{n}{m}v & \frac{n}{m}f \\ 0 & 1 \end{vmatrix} = \frac{n}{m}v > 0.$$

Thus, the JPDF of F and V is

$$f_{F,V}(f, v) = \frac{\left(\frac{n}{m}\right)^{\frac{n}{2}}}{2^{\frac{m+n}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})} f^{\frac{n}{2}-1} v^{\frac{m+n}{2}-1} e^{-\frac{1}{2}(1+\frac{n}{m}f)v} \quad \text{for } f > 0,\, v > 0.$$

Therefore, for $f > 0$, the marginal PDF of F is

$$\begin{aligned} f_F(f) &= \frac{\left(\frac{n}{m}\right)^{\frac{n}{2}}}{2^{\frac{m+n}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})} f^{\frac{n}{2}-1} \int_0^\infty v^{\frac{m+n}{2}-1} e^{-\frac{1}{2}(1+\frac{n}{m}f)v} dv \\ &= \frac{\left(\frac{n}{m}\right)^{\frac{n}{2}}}{B\left(\frac{n}{2}, \frac{m}{2}\right)} f^{\frac{n}{2}-1} \left(1 + \frac{n}{m}f\right)^{\frac{n+m}{2}}. \end{aligned}$$

$\square$

Corollary .2. *Let $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu_1, \sigma_1^2)$ and $Y_1, Y_2, \ldots, Y_m \overset{i.i.d.}{\sim} N(\mu_2, \sigma_2^2)$. Also, assume that X_i's and Y_j's are independent. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$, $S_X^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$, $\overline{Y} = \frac{1}{m}\sum_{i=1}^m Y_i$, and $S_Y^2 = \frac{1}{m-1}\sum_{i=1}^m (Y_i - \overline{Y})^2$. Then*

$$\frac{\sigma_2^2 S_X^2}{\sigma_1^2 S_Y^2} \sim F_{n-1,m-1}.$$

Proof: The proof is straight forward from the Theorems .4 and .2. $\square$