# Mathematics 3

## Special Distributions

25/11/22

Three special continuous distributions, $\chi^2$, $t$ and $F$.

Why? Important for their applications in statistics

It is customary in statistics to denote the random variable associated with these distributions by the same letter as the distribution itself.

# Special Distributions (PS-10)

① $\boxed{\chi^2}$, ② $t$ and ③ $F$ - distributions

$$\boxed{\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt}$$

## $\chi^2$-distribution

$X \equiv \chi^2$

$$f_{\chi^2}(\chi^2) = \frac{1}{2\Gamma\left(\frac{n}{2}\right)} e^{-\frac{\chi^2}{2}} \left(\frac{\chi^2}{2}\right)^{\frac{n}{2}-1}, \quad \chi^2 > 0$$

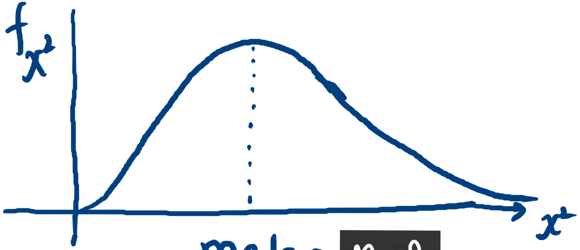
$$= 0 \quad , \quad \text{elsewhere.}$$

$n$ is a +ve integer → called the number of degrees of freedom.

Alt. expression: $X \sim \chi^2(n)$

$$f_X(x) = \frac{1}{2\Gamma\left(\frac{n}{2}\right)} e^{-\frac{x}{2}} \left(\frac{x}{2}\right)^{\frac{n}{2}-1}, \quad x > 0,$$

$$= 0, \quad \text{elsewhere,}$$




mode = $n-2$

$$f_X'(x) = 0$$

$$\Rightarrow x = n-2 \quad \text{(check!)}$$

$$f_X''(n-2) < 0$$

Theorem

① (i) If $X \sim \gamma\left(\frac{n}{2}\right)$, then $Y = 2X$ has $\chi^2(n)$-distribution

conversely (ii) If $Y \sim \chi^2(n)$, then $X = \frac{Y}{2}$ has $\gamma\left(\frac{n}{2}\right)$-distribution

$X \sim \gamma(n/2)$; $f_X(x) = \dfrac{e^{-x}\, x^{n/2-1}}{\Gamma(n/2)}$, $x > 0$

$Y = 2X \Rightarrow y = 2x$, $\dfrac{dy}{dx} = 2 > 0$ ; $f_Y(y) = f_X(x)\left|\dfrac{dx}{dy}\right|$

$$= \frac{1}{2}\,\frac{e^{-x}\, x^{n/2-1}}{\Gamma(n/2)} \Rightarrow Y \sim \chi^2(n)$$

# Mean, Variance of $\chi^2(n)$

from previous slide

$$\boxed{X \sim \chi^2(n) \Rightarrow \frac{X}{2} \sim \gamma\left(\frac{n}{2}\right)}$$ how ?

property $\boxed{E\left(\frac{X}{2}\right) = \frac{n}{2}}$ , $Var\left(\frac{X}{2}\right) = \frac{n}{2}$

$var(cX) = c^2 var(X)$

$\Rightarrow E(X) = n$ , $Var(X) = 2n$

**Theorem ②** If $x_1, x_2, \ldots, x_n$ be mutually independent standard normal variate, then

$$Y = x_1^2 + x_2^2 + \cdots + x_n^2 \text{ has } \chi^2(n) \text{ distribution.}$$

**Sol.**

**Step I:** If $x_i \sim N(0,1) \Rightarrow$ $y_i = \dfrac{x_i^2}{2}$ is $\gamma\left(\frac{1}{2}\right)$ variate

(Prob. 10, PS-⑤)

$(\forall\, i = 1, 2, \ldots, n)$

**Step II:** Since $x_1, x_2, \ldots, x_n$ are mutually independent

$\Rightarrow \dfrac{x_1^2}{2}, \dfrac{x_2^2}{2}, \ldots, \dfrac{x_n^2}{2}$ are also mutually independent

clarify

$\gamma\left(\frac{1}{2}\right)$

**Step III:** Using Reproductive property of $\gamma$-distribution

$$\frac{1}{2}\left(X_1^2+X_2^2+\cdots\cdots+X_n^2\right) \sim \gamma\left(\frac{n}{2}\right)$$

**Step IV:** Using Prop (1), $X_1^2+X_2^2+\cdots+X_n^2 \sim \chi^2(n)$ variate.

**Prob: (3)** (HW) $\boxed{\text{If } X \sim \gamma\left(\tfrac{n}{2}\right), \text{ then } Y = 2X \text{ has } \chi^2(n)\text{-distribution}}$

---

**Prob:** If $X_1, X_2, \ldots, X_n$ are mutually independent $\chi^2$-variate with degrees of freedom $k_1, k_2, \ldots, k_n$, respectively, then $X_1+X_2+\cdots+X_n$ has $\chi^2$-variate with $k_1+k_2+\cdots+k_n$ degrees of freedom.

Sol: $\psi_{X_r}(t) = \text{ch. fn of } X_r = (1-2it)^{-\frac{k_r}{2}}$ $\boxed{\text{how ?}}$

where $X_r$ is a $\chi^2$-variate

$\forall\, r = 1, 2, \ldots, n$

$$S_n = X_1 + X_2 + \cdots + X_n$$

$$\psi_{S_n}(t) = E\left\{e^{itS_n}\right\} = E\left(e^{itX_1}\right) \cdots\cdots E\left(e^{itX_n}\right)$$

$$= (1-2it)^{-\frac{1}{2}(k_1+k_2+\cdots+k_n)}.$$

$$\Rightarrow \quad S_n \sim \chi^2(k_1+k_2+\cdots+k_n).$$

PS-10

Q4 (4) If $X_1, X_2, \ldots, X_n$ are mutually independent standard normal variates, and $Y_1, Y_2, \ldots, Y_n$ are obtained by an orthogonal homogeneous linear transformation: Clarify ?

$$\boxed{Y_i = \sum_{\alpha=1}^{n} a_{i\alpha} X_\alpha} \quad (i = 1, 2, \ldots, n) \quad \text{(i)}$$

where

$$\sum_{\alpha=1}^{n} a_{i\alpha} a_{j\alpha} = \sum_{\alpha=1}^{n} a_{\alpha i} a_{\alpha j} \boxed{= \delta_{ij}} \quad (i, j = 1, 2, \ldots, n) \quad \text{(ii)}$$

then $Y_1, Y_2, \ldots, Y_n$ are mutually independent standard normal variates.

$$Y_1 = a_{11} X_1 + a_{12} X_2 + \cdots + a_{1n} X_n$$

$$Y_2 = a_{21} X_1 + a_{22} X_2 + \cdots + a_{nn} X_n$$

$$\vdots$$

$$Y_n = a_{n1} X_1 + a_{n2} X_2 + \cdots + a_{nn} X_n$$

$$\boxed{Y_i = \sum_{\alpha=1}^{n} a_{i\alpha} X_\alpha}$$

remember ?

$$\boxed{Y = AX}, \quad \boxed{AA^T = A^T A = I_n}$$

Pf: $J = \dfrac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_n)} = \det A = \pm 1$ $\boxed{\text{check!}}$

$$f_{Y_1, Y_2, \ldots, Y_n}(y_1, y_2, \ldots, y_n) = f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) \left| \frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(y_1, y_2, \ldots, y_n)} \right|$$

$$= \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum_{i=1}^{n} x_i^2}$$

[$X_r$ are standard normal variates]

$$= \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum_{i=1}^{n} y_i^2}$$

**Check!**

$$\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} x_i^2$$

[from the orthogonality relations]

$$f_{Y_1, Y_2, \ldots, Y_n}(y_1, y_2, \ldots, y_n) = f_{Y_1}(y_1) f_{Y_2}(y_2) \cdots f_{Y_n}(y_n)$$

$\delta_{ij}$

$\Rightarrow$ $Y_1, Y_2, \ldots, Y_n$ are mutually independent Std. Nor. variates.

⑤ Let, $X_1, X_2, \ldots, X_n$ be mutually independent $N(0,1)$ variates and $Y_1, Y_2, \ldots, Y_m$ $(m<n)$ be given as

will be taken up by Prof. Amit

$$Y_i = \sum_{\alpha=1}^{n} a_{i\alpha} X_\alpha \quad (i=1,2,\ldots,m) \text{ --- (i)}$$

where

$$\sum_{\alpha=1}^{n} a_{i\alpha} a_{j\alpha} = \delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i=j \end{cases} \quad (i,j=1,\ldots,m) \text{ --- (ii)}$$

Then $Q(x_1, x_2, \ldots, x_n) = \boxed{\sum_{\alpha=1}^{n} X_\alpha^2} - \sum_{\beta=1}^{m} Y_\beta^2$ is $\chi^2$-distributed with $\boxed{n}-\boxed{m}$ degrees of freedom.

Pf: Given $m \times n$ coefficients $a_{ij}$ $(i = 1, 2, \ldots, m,$ $j = 1, 2, \ldots, n)$, satisfying (ii), we can determine the rest of the $a_{ij}$ s $(i, j = 1, 2, \ldots, n)$ s.t. we'll have $Y_1, Y_2, \ldots, Y_m, Y_{m+1}, \ldots, Y_n$ with properties in 4(i) and 4(ii).

$$\sum_{\alpha=1}^{n} X_\alpha^2 = \sum_{\alpha=1}^{n} Y_\alpha^2$$

$$Q(x_1, x_2, \ldots, x_n) = \sum_{\alpha=1}^{n} Y_\alpha^2 - \sum_{\alpha=1}^{m} Y_\alpha^2 = Y_{m+1}^2 + Y_{m+2}^2 + \cdots + Y_n^2$$

where $Y_{m+1}, Y_{m+2}, \dots, Y_n$ are mutually independent standard normal variates

$$\Rightarrow \quad Q(x_1, x_2, \dots, x_n) \sim \chi^2(n-m) \qquad \blacksquare$$

# Special Distributions (Contd..) (PS-10)

W. S. Gosset (pseudonym).

## t-Distribution (Student's t-distribution)

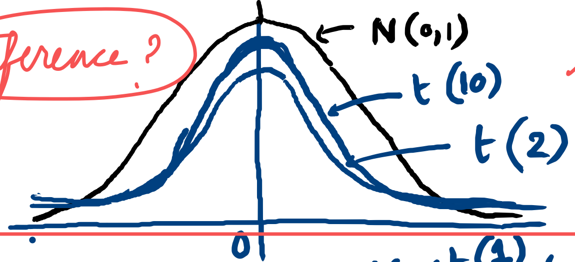
$X \sim t(n)$ ← degrees of freedom

$$f_X(x) = \frac{1}{\sqrt{n}\, B\left(\frac{1}{2}, \frac{n}{2}\right)} \frac{1}{\left(1 + \frac{x^2}{n}\right)^{\frac{n+1}{2}}}, \quad -\infty < x < \infty$$

n is a +ve integer : degrees of freedom

$$\boxed{B(\alpha, \beta) = \frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)}}$$

inference ?




t-distribution is symmetrical about the origin

$$f_X(x) = \frac{1}{\sqrt{n}\, B(\frac{1}{2}, \frac{n}{2})} \frac{1}{\left(1 + \frac{x^2}{n}\right)^{\frac{n+1}{2}}}$$

① For $\underline{n=1}$, $X \sim t(1)$, $f_T(t) = \frac{1}{\pi} \frac{1}{1+t^2}$

→ Cauchy distribution

$$f_X(x) = \frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x-\mu)^2}$$

with parameters ← $\mu = 0$, $\lambda = 1$

Proof that mean for Cauchy DNE → mean does NOT exist

⇒ For $n=1$, t-distribution has no mean

But, For $n>1$, mean exists. think about it

**Prob 4:** If (i) $X$ is a standard normal variate, (ii) $Y$ has $\chi^2(n)$ distribution, and (iii) $X, Y$ are independent, then $T = \dfrac{X}{\sqrt{\frac{Y}{n}}}$ is $t(n)$-variate.

**Sol.** (i) $X \sim N(0,1) \Rightarrow f_X(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty$

(ii) $Y \sim \chi^2(n) \Rightarrow f_Y(y) = \dfrac{e^{-\frac{y}{2}} \left(\frac{y}{2}\right)^{\frac{n}{2}-1}}{2\,\Gamma\left(\frac{n}{2}\right)}, \quad 0 < y < \infty$

(iii) $X$ and $Y$ are independent $\Rightarrow f_{X,Y}(x,y) = f_X(x)\, f_Y(y)$

$= \ldots\ldots$ , $-\infty < x < \infty$, $0 < y < \infty$

Consider $T = \dfrac{X}{\sqrt{Y/n}}$ , $U = Y$

in terms of real variable : $t = \dfrac{x}{\sqrt{y/n}}$ , $u = y$

$$\Rightarrow \quad x = t\sqrt{u/n} \ , \quad y = u$$

$$\frac{\partial(x,y)}{\partial(t,u)} = \begin{vmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial u} \end{vmatrix} = \begin{vmatrix} \sqrt{u/n} & \frac{t}{2\sqrt{nu}} \\ 0 & 1 \end{vmatrix} = \sqrt{\frac{u}{n}} > 0$$

$$f_{T,U}(t,u) = f_{X,Y}(x,y) \left| \frac{\partial(x,y)}{\partial(t,u)} \right|$$

$$= \frac{e^{-\frac{t^2 u}{2n} - \frac{u}{2}}}{\sqrt{2\pi}\ \Gamma(n/2)\sqrt{n}} \frac{1}{2^{n/2}} u^{\frac{n}{2}-\frac{1}{2}} \ , \quad \begin{matrix} -\infty < t < \infty \\ 0 < u < \infty \end{matrix}$$

$$f_T(t) = \int_0^\infty f_{T,u}(t,u)\,du$$

$$=$$

$$= \frac{1}{B\left(\frac{1}{2}, \frac{n}{2}\right)\sqrt{n}} \frac{1}{\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{2}}}, \quad -\infty < t < \infty$$

(check!)

$$f_X(x) = \frac{1}{\sqrt{n}\,B\left(\frac{1}{2}, \frac{n}{2}\right)} \frac{1}{\left(1+\frac{x^2}{n}\right)^{\frac{n+1}{2}}}, \quad -\infty < x < \infty$$

# F-distribution

why called so ?

$X \sim F(m,n)$

$$f_X(x) = \frac{m^{\frac{m}{2}} n^{\frac{n}{2}} x^{\frac{m}{2}-1}}{B\left(\frac{m}{2},\frac{n}{2}\right)(mx+n)^{\frac{m+n}{2}}} \quad , \quad x > 0$$

$f_F(f)$

$$= 0 \quad , \quad \text{elsewh}$$

$m, n$ : are +ve integers

check for $n=1$

**Check:** ① Mean exists for $n > 2$ and

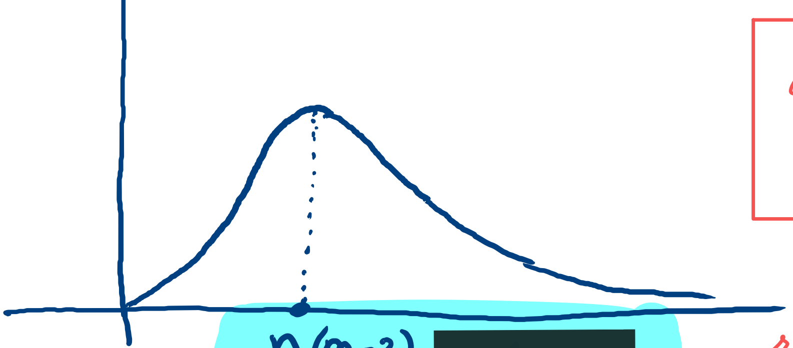
$$m_X = \frac{n}{n-2}$$

mean is independent of $m$

why ?

② For $m>2$, F-distribution has a unique mode at the point $\dfrac{n(m-2)}{m(n+2)}$. (check!)

$\boxed{m=1,2 \text{ ; check}}$



$\dfrac{n(m-2)}{m(n+2)}$ $< 1$

reason

discussed till here

**Prob 7** $X \sim \chi^2(m)$, $Y \sim \chi^2(n)$, then

$$Z = \frac{X/m}{Y/n} \text{ is } F(m,n)\text{- variate.}$$

**Sol**

$$\frac{mZ}{n} = \frac{X}{Y} = \frac{X/2}{Y/2}$$

$$\frac{X}{2} \sim \gamma\left(\frac{m}{2}\right), \quad \frac{Y}{2} \sim \gamma\left(\frac{n}{2}\right)$$

$$W = \frac{X}{Y} \sim \beta_2\left(\frac{m}{2}, \frac{n}{2}\right) \qquad (HW: \quad )$$

$$f_W(w) = \frac{1}{B\left(\frac{m}{2}, \frac{n}{2}\right)} \frac{w^{\frac{m}{2}-1}}{(1+w)^{\frac{m+n}{2}}}, \quad 0 < w < \infty$$

$$Z = \frac{nw}{m} \Rightarrow \frac{dz}{dw} = \frac{n}{m}$$

$$f_Z(z) = f_W(w) \left| \frac{dw}{dz} \right|$$

$$= \ldots$$

$$= \frac{m^{\frac{m}{2}} n^{\frac{n}{2}} z^{\frac{m}{2}-1}}{B\left(\frac{m}{2}, \frac{n}{2}\right) (n+mz)^{\frac{m+n}{2}}}$$

(check!)

**Prob:** If $X \sim F(m,n)$, then $\frac{1}{X} \sim F(n,m)$.

**Sol.** Let, $Y = \frac{1}{X}$ , $y = \frac{1}{x}$

$$\frac{dy}{dx} = -\frac{1}{x^2} < 0 \quad \forall x$$

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| = \cdots\cdots$$

$Y \sim F(n,m)$ (check!).

Prob ⑧ If $X_1, X_2, \ldots, X_n$ are mutually independent normal variates each having mean 0 and standard deviation $\sigma$. Find the distribution of

$$\boxed{X_1^2 + X_2^2 + \cdots + X_n^2}.$$

Sol. $X_i \sim N(0, \sigma) \Rightarrow \dfrac{X_i}{\sigma} \sim N(0,1)$ (check!)

$\forall\, i = 1, 2, \ldots, n.$

$$\Rightarrow \left(\frac{X_i}{\sigma}\right)^2 \sim \chi^2(1) \quad \forall\, i = 1, 2, \ldots, n$$

$$X = \frac{X_1^2 + X_2^2 + \cdots + X_n^2}{\sigma^2} \sim \chi^2(n)$$

Prob ⑦: $X_1, X_2, \ldots, X_n$ m.i. $\sim N(0,1)$

$X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi^2(n)$

Sol. $\dfrac{X_i^2}{2} \sim \gamma\left(\frac{1}{2}\right) \Rightarrow X_i^2 \sim \chi^2(1)$

$$Z = X_1^2 + \cdots + X_n^2$$

$$\Rightarrow Z = \sigma^2 X$$

$$X \sim \chi^2(n) \Rightarrow f_X(x) = \frac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1}}{2\Gamma(\frac{n}{2})}, \quad x > 0$$

$$\Rightarrow Z = \sigma^2 x$$

$$\frac{dz}{dx} = \sigma^2 > 0$$

$$f_Z(z) = f_X(x) \left|\frac{dx}{dz}\right|$$

$$= \frac{e^{-\frac{z}{2\sigma^2}} \left(\frac{z}{\sigma^2}\right)^{\frac{n}{2}-1}}{2\Gamma(\frac{n}{2})} \cdot \frac{1}{\sigma^2}, \quad z > 0$$

$\therefore$

**Prob ⑨:** If $(X,Y)$ has general bivariate normal distribution, show that

$$\boxed{\frac{\left(\frac{X-m_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{X-m_x}{\sigma_x}\right)\left(\frac{Y-m_Y}{\sigma_Y}\right) + \left(\frac{Y-m_Y}{\sigma_Y}\right)^2}{1-\rho^2}} \equiv U^2+V^2$$

has $\chi^2(2)$-distribution.

(i) $U \sim N(0,1)$

(ii) $V \sim N(0,1)$

(iii) $U, V$ are independent

**Sol:** $U = \dfrac{X-m_x}{\sigma_x}$, $V = \dfrac{1}{\sqrt{1-\rho^2}}\left\{\dfrac{Y-m_y}{\sigma_Y} - \rho\,\dfrac{X-m_x}{\sigma_x}\right\}$

$= U^2 + V^2$

?? $\boxed{f_{U,V}(u,v) = f_U(u)\, f_V(v)}$

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-m_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-m_x}{\sigma_x}\right)\left(\frac{y-m_y}{\sigma_y}\right) + \left(\frac{y-m_y}{\sigma_y}\right)^2\right\}}.$$

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} \frac{1}{\sigma_x} & 0 \\ -\frac{\rho}{\sqrt{1-\rho^2}\,\sigma_y} & \frac{1}{\sigma_x\sqrt{1-\rho^2}} \end{vmatrix} = \frac{1}{\sigma_x\sigma_y\sqrt{1-\rho^2}}$$

$$f_{U,V}(u,v) = f_{X,Y}(x,y)\left|\frac{\partial(x,y)}{\partial(u,v)}\right|$$

$$= \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\right)\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}\right) \quad *$$

$$= f_U(u)\, f_V(v)$$

(10) $X, Y$ are independent

$X \sim \chi^2(m)$ and $X+Y \sim \chi^2(m+n)$

Show that $Y \sim \chi^2(n)$.

Sol.

$$\psi_{X+Y}(t) = \text{ch.f of } X+Y$$
$$= \psi_X(t)\,\psi_Y(t)$$
$$= (1-2it)^{-\frac{m+n}{2}}$$

and $\psi_X(t) = (1-2it)^{-\frac{m}{2}}$

$\Rightarrow \psi_Y(t) = (1-2it)^{-\frac{n}{2}} \Rightarrow$ By Uniqueness Th. of Ch. fn. $Y \sim \chi^2(n)$

$\Rightarrow U, V$ are independent

$U \sim N(0,1)$

$V \sim N(0,1)$

$\Rightarrow U^2+V^2 \sim \chi^2(2)$ variate.