1. Random Functions Associated with Normal Distribution

Let $X_1, X_2, \dots, X_n$ be a random sample of size n.

Sample mean (sample average) $\boxed{\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i}$ $E(\bar{X}) = \mu$

Sample variance $\boxed{S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$ $E(s^2) = \sigma^2$

<u>Sampling Distribution of Sample Mean and Sample Variance:</u>

Let $X_1, X_2, \dots, X_n$ be a random sample of size n from a normal distribution $N(\mu, \sigma^2)$, then:

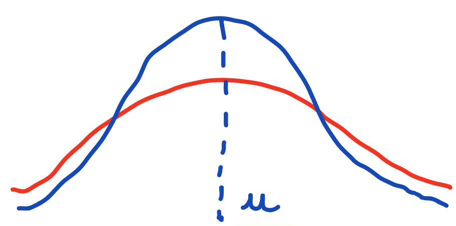
- $\bar{X}$ and S^2 are independent

$\because \bar{X}$ is linear comb of normal X_i $\Rightarrow$ normal

- $\bar{X}$ has normal distribution with mean μ and variance $\frac{\sigma^2}{n}$ $\Rightarrow$ $\boxed{\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)}$

- $\boxed{\frac{(n-1)S^2}{\sigma^2}}$ has a chi-square distribution with $n-1$ degrees of freedom

$\boxed{\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)}$ iid $X_i \sim N(\mu, \sigma^2)$



$\boxed{\chi_r^2 \sim \text{Gamma}\left(\alpha = \frac{r}{2}, \lambda = \frac{1}{2}\right)}$

<u>t-distribution (Student t-distribution)</u>

If $Z \sim N(0,1)$ and $U \sim \chi_r^2$ are independent, then

$\boxed{T = \frac{Z}{\sqrt{U/r}}}$ has t-distribution with r degrees of freedom

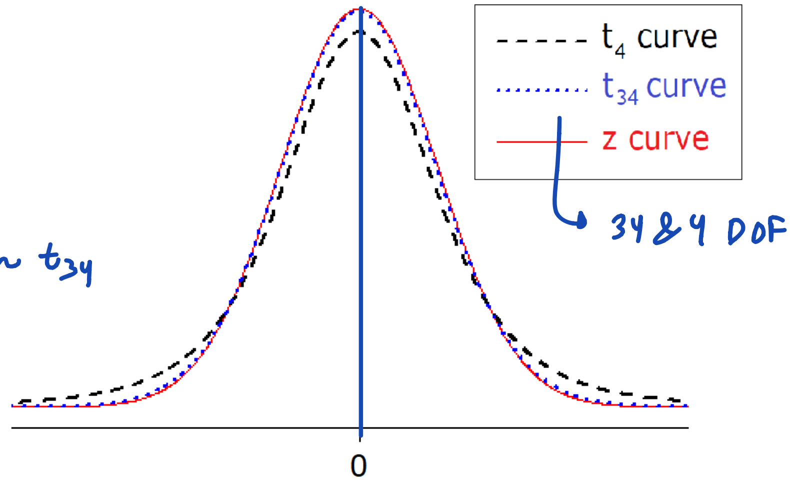
The pdf of t_r distribution $f(x) = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi r}\,\Gamma\left(\frac{r}{2}\right)}\left(1 + \frac{x^2}{r}\right)^{-(r+1)/2}$ $-\infty < x < \infty$

$\boxed{T \sim t_r \text{ or } t(r)}$

The pdf of t-distribution

- bell-shaped and symmetric around 0
- heavier tails than tails of standard normal like t_4 has more than t_{34}
- as the degrees of freedom increase, t-curve gets closer to the z-curve

$t_{34} \approx z$



So for large DOF, we can approximate t distribution to z dist. $(n \geq 30)$

Where is t distribution used?

Suppose $X_1, X_2, \dots, X_n$ is a random sample from $N(\mu, \sigma^2)$.

$\boxed{\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)}$; $\boxed{Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)}$

$\boxed{U = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)}$

Independent

$$\boxed{T = \frac{Z}{\sqrt{U/(n-1)}} \sim t_{n-1}} \Rightarrow \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)s^2}{\sigma^2}/(n-1)}} = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}} \cdot \frac{s}{\sigma}}$$

$$\Rightarrow \boxed{\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}}$$

2. Bivariate Normal Distribution

X and Y have a **bivaraite normal distribution** if their joint pdf is

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right\},$$

$-\infty < x < \infty, -\infty < y < \infty,$ where $-\infty < \mu_X < \infty,\ -\infty < \mu_Y < \infty,\ \sigma_X > 0,\ \sigma_Y > 0,\ |\rho| < 1$

If X and Y have a bivaraite normal distribution, then

- marginally, $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$
- conditionally, $Y|X = x \sim N\left(\underbrace{\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)}_{E(Y|X=x)},\ \underbrace{\sigma_Y^2(1-\rho^2)}_{Var(Y|X=x)}\right)$, similarly you can write $X|Y=y$
- the correlation between X and Y, $Corr(X,Y) = \rho$

$\boxed{aX+bY \text{ will also be normal}}$

Example: Let X be the height (in in) and Y be the weight (in lb) of male college students. Assume X and Y have a bivariate normal distribution with parameters $\mu_X = 69.3$, $\sigma_X^2 = 15.2$, $\mu_Y = 185.2$, $\sigma_Y^2 = 310.8$, and $\rho = 0.6$. If a randomly selected male college student's height is 6', what is the probability that his weight is between 190 and 210 lb?

$P(190 < Y < 210 \mid X = 72)$

$\Rightarrow Y|X=72 \sim N\left(185.2 + (0.6)\left(\sqrt{\frac{310.8}{15.2}}\right)(72-69.3),\ 310.8(1-0.6^2)\right)$

$= N(192.53, 198.91)$

$= P(-0.18 < Z < 1.24) = \Phi(1.24) - \Phi(-0.18)$

$= 0.8925 - (1-0.5714) = 0.4634$

If correlation $\rho = 0$

$$Y|X=x \sim N\left(\mu_y + 0 \cdot \frac{\sigma_y}{\sigma_x}(x-\mu_x),\ \sigma_y^2(1-0^2)\right) = N(\mu_y, \sigma_y^2)$$

$$\boxed{Y \sim N(\mu_y, \sigma_y^2)}$$

* If X, Y are bivariate normal & ρ for them is 0 $\Rightarrow$ X, Y independent $\Leftrightarrow$ $\boxed{\rho_{X,Y} = 0}$

So, $\rho \neq 0 \not\Rightarrow$ independent ✓

But no true for others as we know, only for bivariate

3. Law of Iterated Expectation

Recall: The conditional expectation of X given $Y = y$ is

$$E(X|Y=y) = \begin{cases} \sum_x x\, p_{X|Y}(x|y) & \text{if } X \text{ is discrete with conditional PMF } p_{X|Y}(x|y) \\ \int_{-\infty}^{\infty} x\, f_{X|Y}(x|y)\, dx & \text{if } X \text{ is continuous with conditional PDF } f_{X|Y}(x|y) \end{cases}$$

$= P(x,y)/P_Y(y)$

function of y

E(X|Y) is a rv

$$E[E(X|Y)] = \begin{cases} \sum_y E(X|Y=y)\, p_Y(y) & \text{if } Y \text{ is discrete with the PMF } p_Y(y) \\ \int_{-\infty}^{\infty} E(X|Y=y)\, f_Y(y)\, dy & \text{if } Y \text{ is continuous with the PDF } f_Y(y) \end{cases}$$

⇒ Use this instead of direct

rv

Law of Total Expectation

Law of Iterated Expectation: $\boxed{E(X) = E[E(X|Y)]}$ ⇒ $E(g(X)) = E[E(g(X)|Y)]$

$\boxed{\text{L.I.E}}$

when finding it is tough

Expectation of cond. expectation

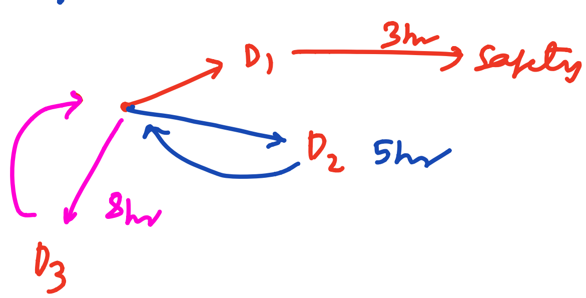
Example: A miner is trapped in a mine containing 3 doors. The first door leads to a tunnel that will take him to safety after 3 hours of travel. The second door leads to a tunnel that will return him to the mine after 5 hours of travel. The third door leads to a tunnel that will return him to the mine after 8 hours. It is completely dark in the mine and the miner cannot mark which door(s) he used in any of his previous attempts. At all times he is equally likely to choose any one of the three doors. What is the expected length of time until he reaches safety?

T

Let T = time till safety

D = door he chose in 1st attempt

$E(T) = E[E(T|D)]$



$$E(T) = E(T \mid D=1) \cdot P_D(1) + E(T \mid D=2)\, P_D(2) + E(T \mid D=3) \cdot P_D(3)$$

$$\Rightarrow E(T) = E(T \mid D=1) \cdot \frac{1}{3} + E(T \mid D=2) \cdot \frac{1}{3} + E(T \mid D=3) \cdot \frac{1}{3}$$

whole process starting again $\Rightarrow \therefore E(T)$

$$\Rightarrow E(T) = 3 \cdot \frac{1}{3} + (5 + E(T)) \cdot \frac{1}{3} + (8 + E(T)) \cdot \frac{1}{3} = \frac{2}{3} E(T) + \frac{16}{3}$$

$$\Rightarrow \quad \frac{1}{3} E(T) = \frac{16}{3} \Rightarrow \boxed{E(T) = 16}$$

Example: Joe's tank holds 15.3 gallons of gas, and he always refills his tank when he gets down to 5 gallons. Y, the amount of gas in Joe's tank, is uniformly distributed between 5 and 15.3 gallons. For y gallons of gas in the tank, the total number of miles he will run is normally distributed with mean $29.7y$ and variance 120. What is the (overall) expected number of miles he will run?

$E(X) = ?$

$X =$ # miles he will run

$y =$ amount of gas in tank

$\boxed{Y \sim \text{Unif}(5, 15.3)}$, $\boxed{X \mid Y=y \sim N(29.7y, 120)}$

$\Rightarrow E(X) = E[E(X|Y)]$

$\Rightarrow E(X|Y=y) = 29.7y$

$\Rightarrow E(X|Y) = 29.7y$

$\Rightarrow E(X) = E(29.7y) = 29.7\,E(y)$

$= (29.7) \cdot \left(\underbrace{\frac{5+15.3}{2}}_{\frac{a+b}{2}}\right) = \boxed{301.46}$

4*. Transformations of Two Random Variables

$\Rightarrow$ find Jacobian & then joint density

X_1 and X_2 have joint pdf $f_{X_1,X_2}(x_1,x_2)$. Let $Y_1 = g_1(X_1,X_2)$ and $Y_2 = g_2(X_1,X_2)$.

Suppose there exist unique functions h_1 and h_2 such that $X_1 = h_1(Y_1,Y_2)$ and $X_2 = h_2(Y_1,Y_2)$.

=> the Jacobian $J = \det\begin{pmatrix} \frac{\partial h_1(y_1,y_2)}{\partial y_1} & \frac{\partial h_1(y_1,y_2)}{\partial y_2} \\ \frac{\partial h_2(y_1,y_2)}{\partial y_1} & \frac{\partial h_2(y_1,y_2)}{\partial y_2} \end{pmatrix} = \frac{\partial h_1(y_1,y_2)}{\partial y_1} * \frac{\partial h_2(y_1,y_2)}{\partial y_2} - \frac{\partial h_1(y_1,y_2)}{\partial y_2} * \frac{\partial h_2(y_1,y_2)}{\partial y_1}$

=> joint pdf of Y_1 and Y_2 $\quad f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}(h_1(y_1,y_2), h_2(y_1,y_2)) * |J|, \quad (y_1,y_2) \in S_{Y_1,Y_2}$

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

Example $X_1 \sim Exp(1)$ | $X_2 \sim Exp(1)$

Independent

$f_{Y_1,Y_2}(y_1,y_2) = ?$

Consider $\overbrace{Y_1 = X_1 - X_2}^{g_1(X_1,X_2)}$ | $\underbrace{Y_2 = X_1 + X_2}_{g_2(X_2,Y_2)}$ $\Rightarrow$ find the joint distribution of Y_1 & Y_2

Two continuous rv are indep. only if this is true

$$\Rightarrow f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) = e^{-x_1} \cdot e^{-x_2}, \quad x_1>0, \; x_2>0$$

$$x_1 = \frac{y_1+y_2}{2} = h_1(y_1,y_2) \quad \Big| \quad x_2 \; \frac{y_2-y_1}{2} = h_2(y_1,y_2)$$

$$\therefore f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}(h_1(y_1,y_2), h_2(y_1,y_2)) \cdot |J|$$

now,

$$J = \det \begin{pmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{pmatrix}$$

$$= \det \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$= \frac{1}{2} \cdot \frac{1}{2} - \left(\frac{-1}{2}\right)\left(\frac{1}{2}\right) = \boxed{\frac{1}{2}}$$

Now,

$$f_{y_1,y_2}(y_1,y_2) = f_{x_1,x_2}(h_1(y_1,y_2), h_2(y_1,y_2)) \cdot |J| =$$

$$e^{-\frac{y_1+y_2}{2}} \cdot e^{\frac{-y_2-y_1}{2}} \cdot \left|\frac{1}{2}\right| =$$

$$\boxed{\frac{1}{2} e^{-y_2}}$$

Conditions:

$$\begin{cases} x_1 > 0 \\ x_2 > 0 \end{cases} \Rightarrow \begin{cases} \frac{y_1+y_2}{2} > 0 \\ \frac{y_2-y_1}{2} > 0 \end{cases} \Rightarrow \begin{cases} y_2 > -y_1 \\ y_2 > y_1 \end{cases}$$