Frequently Asked Questions

1. Define a chi-square variate. Write any one application of it.

"The square of a standard normal variate is called a chi-square variate with 1 degree of freedom".

That is, if $X \sim N(\mu, \sigma^2)$, we define $Z = (X - \mu)/\sigma \sim N(0, 1)$, then $Z^2 \sim \chi^2$ with 1 d.f.

Thus in general, if $X_1, X_2, \ldots, X_n$ are n independent random variables, distributed as $N(\mu_i, \sigma_i^2)$, for i = 1, 2,…, n, then the random variable χ^2 defined by

$$\chi^2 = \sum_{i=1}^{n}\left(\frac{X_i - \mu_i}{\sigma_i}\right)^2$$

is said to follow chi-square distribution with n d.f., with pdf is given by

$$f(\chi^2) = \frac{1}{2^{n/2}\Gamma n/2} e^{-\frac{\chi^2}{2}} \chi^{2(n/2 \ -1)}, \ 0 \le \chi^2 < \infty$$

2. Write the properties of chi-square distribution.

If , the random variable $X \sim \chi^2_{(n)}$ d.f., then

i. Mean = n and variance =2n.

ii. MGF = $(1-2t)^{-n/2}$, provided $|2t| < 1$.

iii. Mode = $n - 2$, for $n > 2$.

iv. Sum of k independent chi-square variates with n_i d.f. is a chi-square variate with

$n = \sum_{i=1}^{k} n_i$ d.f.

v. If $X \sim \chi^2$ with n d.f. and
$Y \sim \chi^2$ with m d.f., then $(X/Y) \sim \beta_2(n/2, m/2)$ and $X/(X+Y) \sim \beta_1(n/2, m/2)$.

vi. Let $X \sim U(0, 1)$, then $Y = -2\log X$ has χ^2 with 2 d.f.

3. Derive the MGF of a chi-square variate. Hence find mean and variance of chi-square variate.

By definition of MGF of a random variable X

$$M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

Since X~ χ^2 distribution with n d.f. we have

$$M_X(t) = \int_0^{\infty} e^{tx} \frac{1}{2^{n/2}\Gamma n/2} e^{-\frac{x}{2}} x^{n/2 - 1} dx, \quad 0 \leq x < \infty;$$

$$= \frac{1}{2^{n/2}\Gamma n/2} \int_0^{\infty} e^{-\frac{1}{2}(1-2t)x} x^{n/2 - 1} dx, \quad 0 \leq x < \infty;$$

$$= \frac{1}{2^{n/2}\Gamma n/2} \frac{\Gamma n/2}{[(1-2t)/2]^{n/2}}$$

$$\therefore M_X(t) = (1-2t)^{-n/2}, \text{ iff } |2t| < 1. \qquad (1)$$

Which is the required mgf of a chi-square distribution

Using binomial expansion for negative index, we get from (1)

$$M_X(t) = 1 + \frac{n}{2}(2t) + \frac{\frac{n}{2}(\frac{n}{2}+1)}{2!}(2t)^2 + \ldots + \frac{\frac{n}{2}(\frac{n}{2}+1)(\frac{n}{2}+2)\ldots(\frac{n}{2}+r-1)}{r!}(2t)^r.$$

Therefore,

$\mu_r' =$ coefficient of $\frac{t^r}{r!}$ in the expansion of $M_X(t)$

$= 2^r \frac{n}{2}(\frac{n}{2}+1)(\frac{n}{2}+2)\ldots(\frac{n}{2}+r-1)$

=n(n+2)(n+4)….(n+2r-2)

When r = 1 then

$\mu_1' =$ coefficient of ‘t’ in the expansion of $M_X(t)$ = n = Mean

When r = 2,

$\mu_2' =$ coefficient of ‘$t^2/2$’ in the expansion of $M_X(t)$ = n (n + 2)

Variance = $\mu_2' - (\mu_1')^2$

= n(n+2) - n^2 = 2n.

4. **State and prove additive property of chi-square distribution.**

The sum of independent χ^2 variates is also a chi-square variate. That is, if X_i

(i=1,2,…,k)are independent χ^2 variates with n_i d.f. respectively then the sum $\sum_{i=1}^{k} X_i$ is also a chi-square variate with n = $\sum_{i=1}^{k} n_i$ d.f.

Proof:

Given X~ χ^2 Distribution with n d.f. then by definition of mgf,

$$M_{X_i}(t) = (1 - 2t)]^{-n_i/2}, \text{ iff } |2t|<1 \text{ and } \forall i=1,2,..k$$

Then by uniqueness theorem of mgf when X_i's are independent, we have

$$M_{\sum_{i=1}^{k} X_i^2}(t) = \prod_{i=1}^{k} M_{X_i^2}(t) = [M_{X_i^2}(t)]^k$$

$$=(1-2t)]^{-n_1/2} \times (1-2t)]^{-n_2/2} \times ... \times (1-2t)]^{-n_k/2}$$

$$=(1-2t)]^{-(n_1+n_2+..+n_k)/2} = (1-2t)]^{-\frac{1}{2}\sum_{i=1}^{k} n_i}$$

Which is the mgf of a chi-square variate with n= $\sum_{i=1}^{k} n_i$ d.f.

5. **Write applications of chi-square distribution**.

Chi-square distribution has large number of applications some of which are as follows:

i. To test the significance of population variance $\sigma^2 = \sigma_0^2$.
ii. To test the goodness of fit.
iii. To test the independence of attributes.
iv. To test the homogeneity of independent estimates of the population variance.
v. To test the homogeneity of independent estimates of the population correlation coefficient.

6. **Let $x_1, x_2, \ldots, x_n$ be a random sample from normal population with mean μ and variance σ^2, then show that**

   i. $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$ **and**

   ii. $\frac{ns^2}{\sigma^2} = \sum_{i=1}^{n}(x_i \quad \bar{x})^2$ **is a chi-square variate with (n-1)d.f. and i) and ii) are independently distributed.**

Ans: The joint probability differential of $x_1, x_2, \ldots, x_n$ is given by

$$dP(x_1.x_2...x_n) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_i)^2} dx_1 dx_2 ... dx_n \quad \infty < x_i < \infty; \ \forall i = 1,2,...,n.$$

Consider the transformation to the variables Y_i(I = 1,2,…,n)by menas of a linear orthogonal transformation Y=AX, where

$$A = \begin{bmatrix} a_{11} & a_{12} & . & . & a_{1n} \\ a_{21} & a_{22} & . & . & a_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{n1} & a_{n2} & & & a_{nn} \end{bmatrix}$$

In particular, $a_{11}=a_{12}=\ldots=a_{1n}=1/\sqrt{n}$

=> $y_1= (1/\sqrt{n})(x_1+ x_2+ \ldots, +x_n)=\sqrt{n}\,\bar{x}$ (1)

Then $dy_1=\sqrt{n}\, d\bar{x}$

It can be easily shown that the above choice of $a_{11}, a_{12}, \ldots, a_{1n}$ satisfies the condition of orthogonality, so that, $\sum_{j=1}^{n} a_{ij}^2 = 1$. Since the transformation is orthogonal, we have

$$\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 + n\bar{x}^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 + y_1^2,$$ **from (1)**

=> $$\sum_{i=2}^{n} y_i^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2.$$ **(2)**

$$\sum_{i=1}^{n}(x_i - \mu)^2 = \sum_{i=1}^{n}(x_i - \bar{x} + \bar{x} - \mu)^2 = .\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - \mu)^2 = \sum_{i=2}^{n} y_i^2 + n(\bar{x} - \mu)^2,$$
from(2)

Now $|A'A|=|I_n|=1$, and therefore the jacobian transformation J = ± 1. Thus the joint density function of $(x_1, x_2, \ldots, x_n)$is given by

$$dG(y_1.y_2...y_n) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}\left(\sum_{i=2}^{n} y_i^{\ 2} + n(\bar{x} - \mu)^2\right)} |J| dy_1.dy_2....dy_n \quad .$$

$$= \left(\frac{1}{(\sigma/\sqrt{n})\sqrt{2\pi}}\right) e^{-\frac{n(\bar{x} - \mu)^2}{2\sigma^2}} d\bar{x} \times \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{n-1} e^{-\frac{1}{2\sigma^2}\sum_{i=2}^{n} y_i^2} dy_2.dy_2...dy_n$$

Therefore, we have on simplification, we get

$g(y_1.y_2.y_3....y_n) = g(y_1).g(y_2.y_3....y_n)$ => y_1 and $(y_2,y_3,....y_n)$ are independent

where $g(y_1)$ is the pdf of $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$ and

$$\sum_{i=2}^{n} y_i^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 = ns^2 = (n-1)S^2$$

are independently distributed. Moreover, $\sum_{i=2}^{n} y_i^2 / \sigma^2 = ns^2 / \sigma^2 \sim \chi^2_{(n-1)}\ d.f.$

7. **If X ~ χ^2 with n d.f. and Y ~ χ^2 with m d.f., and X and Y are independent then show that (X/Y)~β₂(n/ 2 , m/2).**

Proof : Since X and Y are independent chi-square variates we have

$$f(x,y) = \frac{1}{2^{n/2}\Gamma n/2} e^{-\frac{x}{2}} x^{n/2-1} \frac{1}{2^{m/2}\Gamma m/2} e^{-\frac{y}{2}} y^{m/2-1}, \quad 0 \le x, y < \infty$$

$$= \frac{1}{2^{(n+m)/2}\Gamma n/2 \Gamma m/2} e^{-\frac{(x+y)}{2}} x^{n/2-1} y^{m/2-1}, \quad 0 \le x, y < \infty$$

Consider the transformation, u = x/y and v = y then we have x = uv and y = v

The jacobian transformation J is given by

$$J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} = v$$

Thus the joint pdf of random variables U and V is

$$g(u,v)=\frac{1}{2^{(n+m)/2}\Gamma n/2\Gamma m/2}e^{-\frac{(1+u)v}{2}}(uv)^{n/2\ -1}v^{m/2\ -1}|\mathrm{J}|,\ 0\leq u,v<\infty$$

$$=\frac{1}{2^{(n+m)/2}\Gamma n/2\Gamma m/2}e^{-\frac{(1+u)v}{2}}u^{n/2\ -1}v^{m/2+n/2\ 1}\ ,\ 0\leq u,v<\infty$$

Now to get the maginal distribution of U we have to integrate g(u,v) w.r.t. v and thus,

$$g(u)=\frac{1}{2^{(n+m)/2}\Gamma n/2\Gamma m/2}u^{n/2\ -1}\int_0^\infty e^{-\frac{(1+u)v}{2}}v^{m/2+n/2\ 1}\ \mathrm{dv},$$

$$=\frac{1}{2^{(n+m)/2}\Gamma n/2\Gamma m/2}u^{n/2\ -1}\int_0^\infty e^{-\frac{(1+u)v}{2}}v^{m/2+n/2\ 1}\ \mathrm{dv},$$

$$=\frac{1}{2^{(n+m)/2}\Gamma n/2\Gamma m/2}u^{n/2\ -1}\frac{\Gamma(m/2+n/2)}{\left(\frac{1+u}{2}\right)^{m/2+n/2}}$$

$$=\frac{1}{B(n/2,\ m/2)}\frac{u^{n/2\ -1}}{(1+u)^{m/2+n/2}};0\leq u<\propto$$

Which is the pdf of a $\beta_2(n/2,\ m/2)$ variate, where

$B(\frac{n}{2},\frac{m}{2})=\frac{\Gamma\frac{n}{2}\times\Gamma\frac{m}{2}}{\Gamma(\frac{n+m}{2})}$.

Thus, if X ~ χ^2 with n d.f. and Y ~ χ^2 with m d.f., and X and Y are independent then (X/Y)~β_2(n/ 2 , m/2)

8. **If X ~ χ^2 with n d.f. and Y ~ χ^2 with m d.f., and X and Y are independent then show that X / (X+ Y) ~ β_1(n/ 2 , m/2).**

   Proof : Since X and Y are independent chi-square variates we have

$$f(x,y)=\frac{1}{2^{n/2}\Gamma n/2}e^{-\frac{x}{2}}x^{n/2\ -1}\frac{1}{2^{m/2}\Gamma m/2}e^{-\frac{y}{2}}y^{m/2\ -1},\ \ 0\le x,y<\infty$$

$$=\frac{1}{2^{(n+m)/2}\Gamma n/2\Gamma m/2}e^{-\frac{(x+y)}{2}}x^{n/2\ -1}y^{m/2\ -1},\ \ 0\le x,y<\infty$$

Here, we consider the transformation in variables as
u=x/(x + y) and v = x + y then we have x=uv and y= (1-u)v

The jacobian transformation J is given by

$$J=\frac{\partial(x,y)}{\partial(u,v)}=\begin{vmatrix} v & u \\ v & 1\ \ u \end{vmatrix}=v$$

Thus the joint pdf of random variables U and V is

$$g(u,v)=\frac{1}{2^{(n+m)/2}\Gamma n/2\Gamma m/2}e^{-\frac{v}{2}}(uv)^{n/2\ -1}[(1-u)v]^{m/2\ -1}\,|\mathrm{J}|,\ \ 0\le u\le 1,;\ 0\le \mathrm{v}<\infty.$$

$$=\frac{1}{2^{(n+m)/2}\Gamma n/2\Gamma m/2}e^{-\frac{v}{2}}u^{n/2\ -1}(1\text{-}u)^{m/2\ -1}\mathrm{v}^{\frac{n+m}{2}-1}$$

Now to get the maginal distribution of U we have to integrate g(u,v) w.r.t. v and thus,

$$g(u)=\frac{1}{2^{(n+m)/2}\Gamma n/2\Gamma m/2}u^{n/2\ -1}(1\text{-}u)^{m/2\ 1}\int_0^\infty e^{-\frac{v}{2}}v^{\frac{n+m}{2}\ 1}\ \mathrm{dv}$$

$$=\frac{u^{n/2\ -1}(1\text{-}u)^{m/2-1}}{2^{(n+m)/2}\Gamma n/2\ \Gamma m/2}\frac{\Gamma(\frac{n+m}{2})}{(1/2)^{\frac{n+m}{2}}}$$

$$g(u)=\frac{1}{B(\frac{n}{2},\frac{m}{2})}u^{n/2\ -1}(1\text{-}u)^{m/2-1}\ ;0\le \mathrm{u}<1$$

Which is the pdf of a $\beta_1(n/2, \ m/2)$ variate, where

$$B(\frac{n}{2},\frac{m}{2})=\frac{\Gamma\frac{n}{2}\times\Gamma\frac{m}{2}}{\Gamma(\frac{n+m}{2})}.$$

Thus, if $X \sim \chi^2$ with n d.f. and $Y \sim \chi^2$ with m d.f., and X and Y are independent then $X / (X + Y) \sim \beta_1(n/2, m/2)$.

9. If X ~ U(0, 1), then show that Y= -2logX has χ^2 with 2 d.f.

Proof: Given X ~ U(0, 1), then f(x) = 1, 0 < x < 1.

Consider, y= - 2log(x) then $x = e^{-y/2} \Rightarrow dx = e^{-y/2}(-1/2)\,dy$.

Here, when x = 0 y = ∞ and when x = 1, y = 0.

Therefore,

$g(y) = f(x).|dx/dy| = (½)\, e^{-y/2}$

$$\Rightarrow g(y) = \frac{1}{2^{2/2}\Gamma(2/2)} e^{\ y/2} y^{2/2 \ \ 1};\ 0 \leq y < \infty$$

which is a chi-square variate with 2 d.f. Thus, if X ~ U(0, 1), then Y= -2logX has χ^2 with 2 d.f.

10. State and prove limiting form of χ^2 Distribution.

Statement : Let X~ χ^2 distribution with n d.f. then for large n (i.e., as n-- > ∞) χ^2 distribution is asymptotically distributed as standard normal distribution

Proof : Let X~ χ^2 distribution with n d.f. then

$M_X(t) = (1 - 2t)]^{-n/2}$, iff $|2t| < 1$.

The MGF of standard χ^2 variate $Z=(x-n)/\sqrt{2n}$ is

$$M_Z(t) = e^{-nt/\sqrt{2n}}(1 \ - \ t\sqrt{\tfrac{2}{n}})]^{n/2}, \text{ iff } |2t|<1.$$

Taking logarithm on bothsides of $M_Z(t)$,

$$Log\, M_Z(t) = -t\sqrt{\tfrac{n}{2}} - \frac{n}{2}\log(1 - t\sqrt{\tfrac{2}{n}}),$$

$$Log\, M_Z(t) = \ -t\sqrt{\tfrac{n}{2}} + \frac{n}{2}\left[t\sqrt{\tfrac{2}{n}} + \frac{t^2}{2}\frac{2}{n} + \frac{t^3}{3}\left(\frac{2}{n}\right)^{3/2} + \ldots \right]$$

$$= -t\sqrt{\tfrac{n}{2}} + t\sqrt{\tfrac{n}{2}} + \frac{t^2}{2} + O(n)^{-1/2}$$

$$\log M_Z(t) = \frac{t^2}{2} + O(n)^{-1/2}$$

Where $O(n^{-1/2})$ be the order of n containing $n^{1/2}$ and higher powers of n in the denominator.

$$\therefore \lim_{n\to\infty} \log M_Z(t) = \frac{t^2}{2}$$

$$\Rightarrow \lim_{n\to\infty} M_Z(t) = e^{\frac{t^2}{2}},$$

which is the MGF of a standard normal variate. Thus by uniqueness theorem of mgf, χ^2 distribution tends to normal distribution asymptotically. That is as n-- >∞, χ^2 distribution tends to normal distribution.