

# statistics (section A)

Date : 25.02.2019

$$\textcircled{I} \int_0^\infty e^{-x} x^{n-1} dx = \Gamma_n \quad (\text{smart})$$

$$\textcircled{II} \int_0^\infty e^{-ax} x^{n-1} dx = \frac{\sqrt{n}}{a^n}$$

$$\textcircled{III} \int_0^\infty e^{-ax^2} (x)^{2n-1} dx = \frac{\sqrt{n}}{2 \cdot a^n}$$

— anhänger des Fermionen Produkts

$$\textcircled{IV} \int_{-\infty}^{\infty} e^{-ax^2} dx = \left( \frac{\pi}{a} \right)^{1/2} = \frac{\sqrt{\pi}}{a}$$

$$\textcircled{V} \int_0^1 x^m (1-x)^{n-1} dx = \beta(m, n) \quad [\beta \text{ of 1st kind}]$$

$$\textcircled{VI} \int_0^1 \frac{x^{n-1}}{(1+x)^{m+n}} dx = \beta(m, n) \quad [\beta \text{ of 2nd kind}]$$

$$\textcircled{VII} \beta(m, n) := \frac{\gamma_m \gamma_n}{\gamma_{m+n}}$$

$$\textcircled{VIII} \beta(m, n) = \int_0^\infty \frac{x^{m-1}}{(1+x)^{m+n}} dx$$

$$\begin{aligned} \gamma_n &= (n-1) \gamma_{n-1} = (n-1)(n-2) \gamma_{n-2} \\ \gamma_n &= (n-1)! = (n-1)(n-2)! \end{aligned}$$

## chi-square distribution

( $\chi^2$  distribution)

$$x \sim N(\mu, \sigma^2)$$

mean  $\downarrow$  S.D.

standard normal distribution —

$$z = \frac{x-\mu}{\sigma} \sim N(0, 1)$$

$\chi^2$ -variate:

The square of the standard normal variate

is known as  ~~$\chi^2$ -variate~~  $\chi^2$ -variate with  
1 degrees of freedom (f.d.).

These of  $x \sim N(\mu, \sigma^2)$ . Then

$$z = \frac{x-\mu}{\sigma} \sim N(0, 1)$$

and  $z^2 = \left(\frac{x-\mu}{\sigma}\right)^2$  is a chi-square variate with 1 degrees of freedom (f.d.).

$$\Pr(Z^2 > \chi^2_{\alpha}) = \alpha$$

In general if  $x_i$  ( $i = 1, 2, \dots, n$ ) are  $n$  independent normal variate with mean  $\mu_i$  ( $i = 1, 2, \dots, n$ ) and variance  $\sigma_i^2$  ( $i = 1, 2, \dots, n$ ). Then

$Z^2 = X^2 = \sum_{i=1}^n \left( \frac{x_i - \mu_i}{\sigma_i} \right)^2$  is a chi-square variate with  $n$  degrees of freedom (d.f.).

A random variable  $x$  is said to have a chi-square distribution with  $n$  d.f. if

its p.d.f. is given by

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

where  $n$  is the parameter of the  $\chi^2$ -distribution and also known as degrees of freedom.

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot n^{1/2}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2}$$

Date : 04.08.2017

Properties (of the  $\chi^2$  distribution):

$$\begin{aligned} M_{X^n}(t) &= E(e^{tX^n}) \\ &= \int_0^\infty e^{tx^n} f(x^n) dx^n \\ &= \int_0^\infty e^{tx^n} \frac{1}{2^{n/2} \Gamma(n/2)} e^{-x^{n/2}} (x^n)^{n/2-1} dx^n \\ &= \int_{-\infty}^\infty e^{tx^n} f(x) dx \end{aligned}$$

MGF (Momentary Function)

$$\begin{aligned} M_{X^n}(t) &= \frac{1}{2^{n/2} \Gamma(n/2)} \int_0^\infty e^{-x^{n/2}/2} (1-2t)^{n/2} (x^n)^{n/2-1} dx^n \\ &= \frac{1}{2^{n/2} \Gamma(n/2)} \int_0^\infty x^{n/2} \frac{\sqrt{n/2}}{(1-2t)^{n/2}} e^{-x^{n/2}/2} dx^n \end{aligned}$$

$$M_{X^n}(t) = \frac{1}{2^{n/2} \Gamma(n/2)} \int_0^\infty x^{n/2} \frac{\sqrt{n/2}}{(1-2t)^{n/2}} e^{-x^{n/2}/2} dx^n = \frac{1}{(1-2t)^{n/2}}$$

$$M_{X^n}(t) = (1-2t)^{-n/2}$$

moment function

\* find the modulatory function of the  $x^n$ -distribution.

MGF (momentary Generation Function)

$$M_x(t) = E(e^{tx})$$

CGF : (cumulative generating function)

$$K_{x^n}(t) = \log M_x(t)$$

$$K_{x^n}(t) = \log M_{x^n}(t)$$

$$= \log \left\{ (1-2t)^{-n/2} \right\}$$

$$= -\frac{n/2}{2} \log(1-2t)$$

$$= -\frac{n/2}{2} \left[ 1 + 2t + \frac{(2t)^2}{2} + \frac{(2t)^3}{3} + \dots \right]$$

$$K_{x^n}(t) = nt + 2n \cdot \frac{t^2}{2} + 4n \cdot \frac{t^3}{3} + 8n \cdot \frac{t^4}{4} + \dots$$

$K_1$  = co-efficient of  $\frac{t}{1!}$  = Mean ( $\mu_1$ )

$K_2$  = " "  $\frac{t^2}{2!}$  = Variance ( $\mu_2$ )

$K_3$  = " "  $\frac{t^3}{3!}$  =  $\mu_3$

$K_4$  = " "  $\frac{t^4}{4!}$  . sum of fourth powers

$$K_1 = \text{co-efficient of } \frac{t}{1!} = n = \mu_1 = \text{mean}$$

$$K_2 = \text{co-efficient of } \frac{t^2}{2!} = 2n = \mu_2 = \text{variance}$$

$$K_3 = n \Rightarrow \frac{t^3}{3!} = 8n = \text{skewness}$$

$$K_4 = \text{co-efficient of } \frac{t^4}{4!} = 48n = \text{kurtosis}$$

$$\begin{aligned} \mu_3 &= K_4 + 3K_2 \\ &= 48n + 3 \cdot 4n^2 \\ &= 48n + 12n^2 \end{aligned}$$

$$\text{Skewness } \beta_1 = \frac{(\mu_3)^2}{\mu_2^3} = \frac{(8n)^2}{(2n)^3} = \frac{64n^2}{8n^3} = \frac{8}{n}$$

$$\text{Kurtosis } \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{48n + 12n^2}{(2n)^2} = \frac{12}{n} + 3$$

$$\text{mean } \bar{x} = \frac{n}{12}$$

$$\text{variance } = 2n$$

$$\beta_1 = \frac{8}{n}$$

$$\beta_2 = 3 + \frac{12}{n}$$

$$n = \text{degrees of freedom}$$

$$\beta_1 = 0, \beta_2 = 3$$

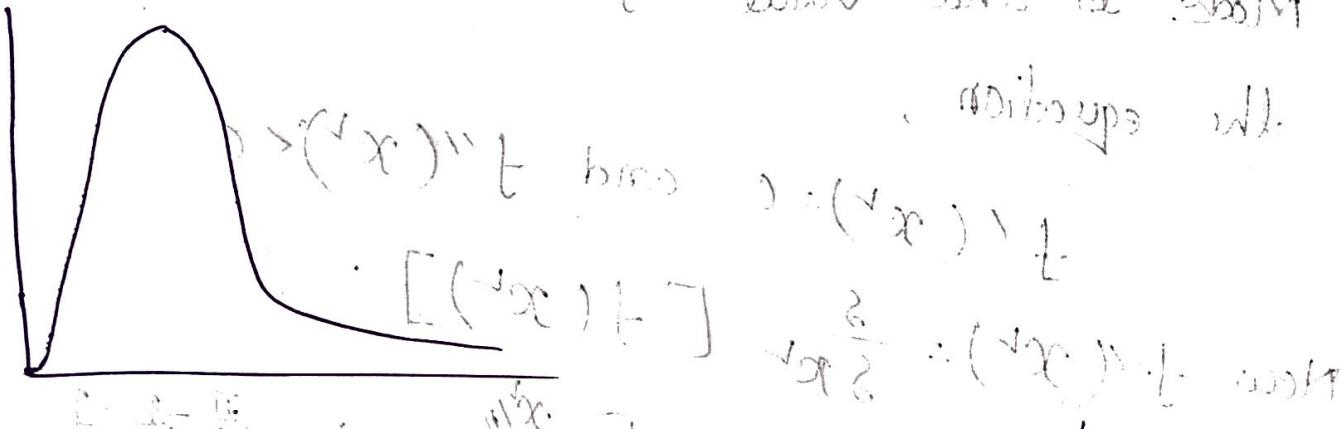
normal distribution

$\beta_1 > 0 \rightarrow$  positive

$\beta_2 > 3 \rightarrow$  Leptokurtic ( $\chi^2$ )

The  $\chi^2$ -distribution is positive, skewed and leptokurtic.

positive at values  $\chi^2$  below 3, 0 at  $\chi^2 = 3$



>Show that  $\chi^2$ -distribution is always normal distribution.

if  $n \rightarrow \infty$  then  $\beta_1 \rightarrow 0$

$\beta_2 \rightarrow 3$

Date : 07.03.2017

~~$x^r$~~

Mode of  $X^r$ -distribution:

$$f'(x) = 0 \quad | \quad f''(x) < 0 \quad (0 < x < n)$$

We know, condition of mode of  $X^r$  is  $f''(x^r) < 0$  and  $(x^r)^{\frac{n}{2}-1}$

$$f(x^r) = \frac{1}{2^{n/2} \sqrt{\frac{n}{2}}} e^{-\frac{x^r}{2}}$$

mode is that value of  $x^r$  which is satisfies

the equation,

$$f'(x^r) = 0 \quad \text{and} \quad f''(x^r) < 0$$

$$\begin{aligned} f'(x^r) &= \frac{s}{s x^r} [f(x^r)] \\ \text{Now } f'(x^r) &= \frac{1}{2^{n/2} \sqrt{\frac{n}{2}}} \left[ e^{-\frac{x^r}{2}} \left( \frac{n}{2} - 1 \right) (x^r)^{\frac{n}{2}-1} \right. \\ &\quad \left. + (x^r)^{\frac{n}{2}-1} e^{-\frac{x^r}{2}} \left( -\frac{1}{2} \right) \right] \\ &= \frac{1}{2^{n/2} \sqrt{\frac{n}{2}}} e^{-\frac{x^r}{2}} (x^r)^{\frac{n}{2}-1} \left[ \left( \frac{n}{2} - 1 \right) (x^r)^{-\frac{1}{2}} \right] \end{aligned}$$

puts  $f'(x^*) = 0$

$$f'(x^*) = 0 \quad \text{and} \quad 0 = \frac{\frac{n}{2} - 1}{2^{n/2}} e^{-x^{*2}/2} (x^*)^{\frac{n}{2}-1} \left[ \left(\frac{n}{2} - 1\right) (x^*)^{-\frac{1}{2}} \right]$$

$$\Rightarrow e^{-x^{\nu/2}} (x^\nu)^{\frac{n}{2}-1} \left[ (\frac{n}{2}-1) (x^\nu)^{-1-\frac{1}{2}} \right] = 0$$

$$\Rightarrow e^{av} \cdot B + B \cdot e^{-av} = -x^{v/2}(x^4)^{m_2-1} \neq 0$$

Either,  $(n-1)(x^r)^{\frac{1}{2}-\frac{1}{2}} = 0$  or  $(x^r)^{\frac{1}{2}} = 0$

$$\therefore \lim_{n \rightarrow \infty} \frac{n-1}{x_n} = \frac{1}{x_0} = \frac{1}{2}$$

$\Rightarrow x^2 = m^2$  shows that all the solutions for  $x$  are  $\pm m$ .

$\Rightarrow$  It shows that  $f'(x^r) < 0$  when the distribution is  $(n-2)$ .

Hence the mode of distribution of  $\chi^2$  is called Chi-square distribution.

for  $n$  degree of freedom. Additive properties of  $\chi^2$ -distribution.

\* Hail - sum of the  $\chi^2$ -variate  
sum of independent  $\chi^2$ -variate  
statement: The sum of independent  $\chi^2$ -variate  
 is also a  $\chi^2$ -variate. Let  $x_1, x_2, \dots, x_k$  be  
 the  $k$ -independent variate with  $n_1, n_2, \dots, n_k$   
 degrees of freedom.

The sum  $\sum_{i=1}^k x_i^r$  is a  $x^r$ -variate with  
 $n_1 + n_2 + \dots + n_k = \sum_{i=1}^k n_i \cdot df$ .

to state & prove  $x^r$ -variate.

Given (known) the m.g.f. of  $x^r$ -variate

$$M_{x^r}(t) = (1-\alpha t)^{n/2} \quad i=1, 2, \dots, k$$

since  $x_1^r, x_2^r, \dots, x_k^r$  be the  $k$ -independent

$x^r$ -variate with  $n_1, n_2, \dots, n_k$  df.

then  $\sum_{i=1}^k x_i^r$  can be written as

The m.g.f. of  $\sum_{i=1}^k x_i^r$  is

$$M_{\sum_{i=1}^k x_i^r}(t) = \prod_{i=1}^k M_{x_i^r}(t) = M_{x_1^r}(t) \cdot M_{x_2^r}(t) \cdots M_{x_k^r}(t)$$

$$= (1-2t)^{n/2} \cdot (1-2t)^{n/2} \cdots (1-2t)^{n/2}$$

$$= (1-2t)^{(n_1/2 + n_2/2 + \dots + n_k/2)}$$

$$= (1-2t)^{-\frac{1}{2} \sum_{i=1}^k n_i}$$

$$= (1-2t)^{-n/2} \quad \text{when } n = n_1 + n_2 + \dots + n_k$$

which is the momentary function of  
 $x^r$ -variate with  $n_1, n_2, \dots, n_k$  df.

so the momentary function of  $\sum_{i=1}^K x_i^r$  is  
a  $\chi^2$ - variate with  $\sum_{i=1}^K n_i$  d.f.

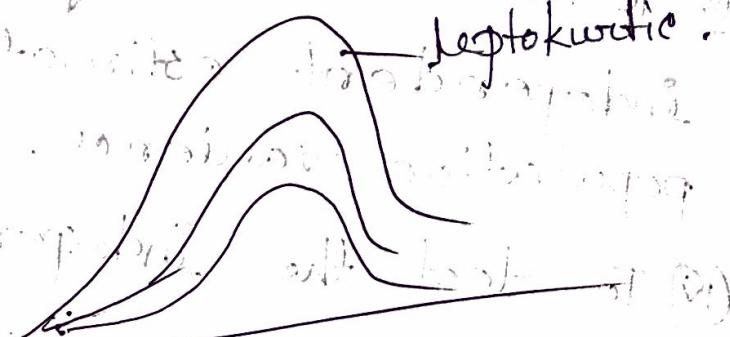
### Application of $\chi^2$ -distribution:

- ① To test if the hypothetical value of the population variance is  $\sigma^2 = \sigma_0^2$  (say)
- ② To test the goodness of fit.  $H_0: \sigma^2 = \sigma_0^2$  (say)
- ③ To test the homogeneity of independent estimate of the population variance.
- ④ To test the independence of attributes (characteristics)
- ⑤ To test the homogeneity of independent estimate of the population correlation coefficient.

### Write down the properties of $\chi^2$ -distribution:

- ①  $\chi^2$  is a continuous type of distribution and its range 0 to  $\infty$ .
- ② The distribution contain only one parameter which is the degrees of freedom of the distribution.

- (10) The mean and variance of the  $\chi^2$ -distribution for  $n$  degrees of freedom are respectively.
- (11) The mode of the  $\chi^2$ -distribution is (a)  $n-2$ , (b)  $n-1$ , (c)  $n$ , (d)  $n+1$ .
- (12)  $\chi^2$ -distribution tends to normal distribution for large degrees of freedom.
- (13)  $\chi^2$ -distribution is positively skewed and leptokurtic.



The distribution of  $\chi^2$  is leptokurtic with much of the probability concentrated near the origin and a long tail extending to the right. It is unimodal and symmetric about its mean.

## T-distribution

Date: 11.03.2017

W.S. Gosset (1908) / student's

student's t-distribution

$$t = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \quad | \quad E(s^2) = \sigma^2$$

Normal distribution,  $D = Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$

$n > 30$  → large normal distribution.

$n > 30 \rightarrow$  normal distribution

$n \leq 30 \rightarrow$  student's t-distribution

t-distribution: Let  $x_1, x_2, \dots, x_n$  be a random sample of size  $n$  from a normal distribution with mean  $\mu$  and variance  $\sigma^2$ . Then student's t-distribution by the

statistic,  $t_m = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$  where  $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$

and  $s^2 = \frac{1}{n} \sum_{i=1}^{n-1} (x_i - \bar{x})^2$  is an unbiased

estimate of population variance  $\sigma^2$  and it follows student's t-distribution with

\*  $\nu = n-1$  df with probability density function,

$$f(t) = \frac{1}{\sqrt{\nu} \rho(\frac{1}{2}, \frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{\frac{\nu+1}{2}}$$

$-\infty < t < \infty$

$$f(t) = \frac{1}{\sqrt{n} \rho(\frac{1}{2}, \frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}} ; n = \nu - 1$$

\* Fisher's  $t$ -distribution:

$\mu = \gamma = N(0, 1)$  standard normal

Fisher's  $t$ -distribution is the ratio of a standard normal distribution variate to the square-root of an independent chi-square variate divided by their degrees of freedom. If

$\mu$  or  $\gamma$  is  $N(0, 1)$  and  $\chi^2$  is

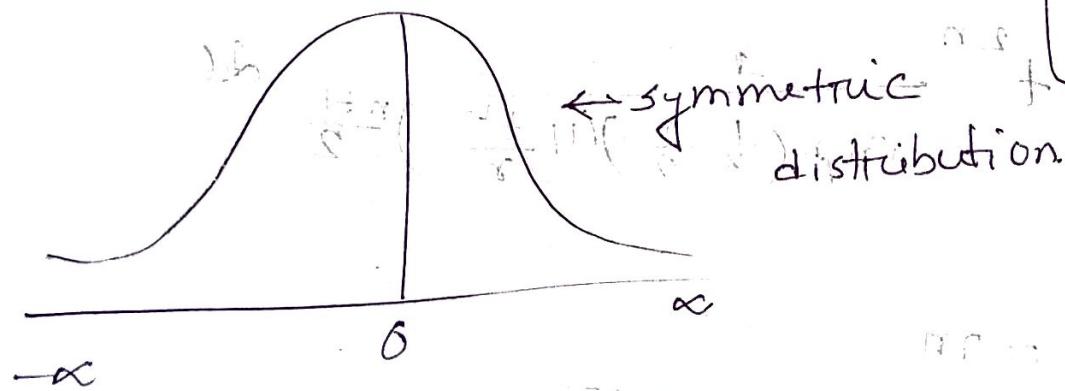
an independent chi-square variate

then, relationship between  $t$  and  $\chi^2$  is

with ~~d~~ n.d.f. then fisher's t-dists  
tribution is defined, as

$$t = \frac{u}{\sqrt{\frac{x^r}{n}}} \text{ or } t = \frac{f}{\sqrt{\frac{x^r}{n}}}$$

$$n = \sqrt{E^2(u)} \quad -\infty < t < \infty$$



$$\begin{aligned} t &= \frac{\bar{x} - \mu}{\sqrt{s^2/n}} \\ &= \frac{\bar{x} - \mu}{\sqrt{s^2/n}} \end{aligned}$$

Odd raw moments

$$\mu'_{2n+1} = 0 \quad n=0, 1, 2, 3, \dots$$

$$\mu'_{2n+1} = 0$$

Since  $f(t)$  is symmetrical about the line  $t=0$  all odd order moment about origin vanish i.e.

$$\mu'_{2n+1} = 0, \quad n=0, 1, 2, \dots$$

Now in particular  $\mu'_1 = 0$  means

Hence the central moments coincide with moment about origin i.e.,  $E(t^n) = \mu_n$ .

$$n=1, 2, 3$$

Even ~~odd~~ moment:

$$\mu_{2n} = E(t^{2n}) = \int_{-\infty}^{\infty} t^{2n} f(t) dt$$

$$= \int_{-\infty}^{\infty} t^{2n} \frac{1}{\sqrt{n} \rho \left( \frac{1}{2} + \frac{n}{2} \right) \left( 1 + \frac{t^2}{n} \right)^{\frac{n+1}{2}}} dt$$

$$\text{Let, } \frac{t^n}{n} = m$$

$$\text{or, } t^n = n \cdot m \Rightarrow t = \sqrt[n]{nm}$$

$$2 \cdot d dt = n dm$$

$$= 2 \int_0^{\infty} \frac{t^{2n}}{\sqrt{n} \rho \left( \frac{1}{2} + \frac{n}{2} \right) \left( 1 + \frac{t^2}{n} \right)^{\frac{n+1}{2}}} dt$$

$$= 2 \int_0^{\infty} \frac{(nm)}{\sqrt{n} \rho \left( \frac{1}{2} + \frac{n}{2} \right) \left( 1 + \frac{nm}{n} \right)^{\frac{nm}{2}}} dt$$

$$\frac{dt}{2t} dt = \frac{ndm}{dm}$$

$$t^{\frac{n}{2}} (t+1) \int \left( \left( \frac{n}{2} \cdot \frac{1}{t} \right) \frac{1}{2t} \right) \frac{dt}{dt} = \frac{ndm}{dm}$$

$$t^{\frac{n}{2}} (t+1) \int \left( \left( \frac{n}{2} \cdot \frac{1}{t} \right) \frac{1}{2t} \right) =$$

Date : 14.03.2017.

or m = t + 1, m > 1

$$m = t + 1$$

$$\text{Let, } t^{\frac{n}{2}} / n = m$$

$$\Rightarrow t^{\frac{n}{2}} = nm$$

$$\Rightarrow t = \sqrt[n]{nm} = (nm)^{\frac{1}{2}}$$

$$\Rightarrow \frac{dt}{dm} = \frac{1}{2} (nm)^{\frac{1}{2}-1}$$

$$\Rightarrow \frac{dt}{dm} = \frac{1}{2} (nm)^{-\frac{1}{2}} \cdot \frac{1}{\sqrt{n}}$$

$$\Rightarrow \frac{dt}{dm} = \frac{1}{2} \cdot \frac{1}{\sqrt{n}} m^{-\frac{1}{2}}$$

$$\text{or } b \quad t^{\frac{n}{2}} (t+1) \int \left( \left( \frac{n}{2} \cdot \frac{1}{t} \right) \cdot \frac{1}{2t} \right) =$$

$$\text{or } b \quad \left( \frac{n}{2} + \frac{1}{2} \right) \int \left( \frac{1}{t} + \frac{1}{t^2} \right) dm = \left( \frac{n}{2} + \frac{1}{2} \right) m$$

$$\text{or } b \quad \left[ \frac{n}{2} \ln t + \frac{1}{2} t^{-1} \right] = \frac{n}{2} m$$

$$\text{or } b \quad \left[ \frac{n}{2} \ln t + \frac{1}{2} t^{-1} \right] = \frac{n}{2} m$$

$$\text{or } b \quad \left[ \frac{n}{2} \ln t + \frac{1}{2} t^{-1} \right] = \frac{n}{2} m$$

$$\text{or } b \quad \left[ \frac{n}{2} \ln t + \frac{1}{2} t^{-1} \right] = \frac{n}{2} m$$

$$\begin{aligned} u_{2n} &= \frac{1}{\sqrt{n} \beta(\frac{1}{2}, \frac{n}{2})} \int_{-\infty}^{\infty} \frac{t^{2n}}{(1 + \frac{t^n}{n})^{\frac{n+1}{2}}} dt \\ &= \frac{1}{\sqrt{n} \beta(\frac{1}{2}, \frac{n}{2})} \cdot 2 \int_0^{\infty} \frac{t^{2n}}{(1 + \frac{t^n}{n})^{\frac{n+1}{2}}} dt \end{aligned}$$

~~from part 2~~

when  $t=0$ ,  $m=0$

$$t=\infty \quad m=\infty$$

$$\begin{aligned} &= \frac{2}{\sqrt{n} \beta(\frac{1}{2}, \frac{n}{2})} \int_0^{\infty} \frac{m^{n+1}}{(1+m)^{\frac{n+1}{2}}} dm \frac{\sqrt{n}}{2\sqrt{m}} \\ &= \frac{n^n}{\beta(\frac{1}{2}, \frac{n}{2})} \int_0^{\infty} \frac{m^{n-\frac{1}{2}}}{(1+m)^{\frac{n+1}{2}+\frac{1}{2}}} dm \end{aligned}$$

$$= \frac{n^n}{\beta(\frac{1}{2}, \frac{n}{2})} \int_0^{\infty} \frac{m^{(\frac{n}{2}-\frac{1}{2})-1}}{(1+m)^{(\frac{n}{2}-\frac{1}{2})+(\frac{n}{2}+\frac{1}{2})}} dm$$

$$= n^n \frac{\Gamma(\frac{n}{2}-\frac{1}{2}) \Gamma(\frac{n}{2}+\frac{1}{2})}{\Gamma(\frac{n}{2}-\frac{1}{2}+\frac{n}{2}+\frac{1}{2})} \quad \frac{\Gamma(\frac{1}{2}) \Gamma(\frac{n}{2})}{\Gamma(\frac{1}{2}+\frac{n}{2})}$$

$$= n^n \frac{\Gamma(\frac{n}{2}-\frac{1}{2}) \Gamma(\frac{n}{2}+\frac{1}{2})}{\Gamma(\frac{n}{2}-\frac{1}{2}+\frac{n}{2}+\frac{1}{2})} \times \frac{\Gamma(\frac{n}{2}+\frac{1}{2})}{\Gamma(\frac{n}{2}+\frac{1}{2})}$$

$$\therefore u'_{2n} = n^n \frac{\Gamma(\frac{n}{2}+\frac{1}{2})}{\Gamma(\frac{n}{2}-\frac{1}{2}) \Gamma(\frac{n}{2}+\frac{1}{2})} \quad ; \pi = 1, 2, 3, \dots$$

if  $n=1$ ,

$$u'_{2,1} = n \cdot \frac{T_{\frac{n}{2}-1} T_{\frac{1}{2}}}{T_{\frac{1}{2}} T_{\frac{n}{2}}} = \frac{n T_{\frac{n}{2}-1} T_{\frac{3}{2}}}{T_{\frac{1}{2}} T_{\frac{n}{2}} (p-r)(q-s)} = \frac{n T_{\frac{n}{2}-1} (\frac{3}{2}-1) T_{\frac{3}{2}-1}}{T_{\frac{1}{2}} (\frac{n}{2}-1) T_{\frac{n}{2}-1} (p-r)(q-s)} = \frac{n \frac{1}{2} T_{\frac{1}{2}}}{T_{\frac{1}{2}} (\frac{n}{2}-1) (p-r)(q-s)} = \frac{\frac{n}{2} \times \frac{2}{n-2}}{T_{\frac{1}{2}} (\frac{n}{2}-1) (p-r)(q-s)} = \frac{n}{n-2}$$

$$\therefore u'_{2,1} = \frac{n}{n-2} \text{ if } n > 2$$

if  $n=2$

$$u'_{2,2} = \frac{n^r T_{\frac{n}{2}-2} T_{\frac{1}{2}+\frac{1}{2}}}{T_{\frac{1}{2}} T_{\frac{n}{2}}} = \frac{n^r T_{\frac{n}{2}-2} T_{\frac{5}{2}}}{T_{\frac{1}{2}} T_{\frac{n}{2}}}$$

$$u'_{2,2} = \frac{n^r T_{\frac{n}{2}-2} T_{\frac{5}{2}}}{T_{\frac{1}{2}} T_{\frac{n}{2}}} = \frac{n^r T_{\frac{n}{2}-2} (\frac{5}{2}+1) (\frac{5}{2}-2) T_{\frac{5}{2}-2}}{T_{\frac{1}{2}} (\frac{n}{2}-1) (\frac{n}{2}-2) T_{\frac{n}{2}-2}}$$

$$\begin{aligned}
 &= \frac{n^2 \cdot \frac{3}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \left(\frac{n}{2}-1\right) \left(\frac{n}{2}-2\right)} \\
 &= \frac{n^2 \cdot \frac{3}{4}}{(n-2)(n-4)} \cdot \frac{\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \frac{1}{2}} \\
 &= n^2 \cdot \frac{3}{4} \times \frac{\frac{1}{4}}{(n-2)(n-4)} \\
 &= \frac{3n^2}{(n-2)(n-4)} \cdot \frac{1}{(1-\frac{2}{n}) \cdot \frac{1}{2}}
 \end{aligned}$$

$$\mu_1 = 0$$

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

$$= \frac{n}{n-2} - (0)^2$$

$$= \frac{n}{n-2} = \text{variance if } n > 2$$

$$\begin{aligned}
 \mu_3 &= \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3 \\
 &= 0 - 3 \cdot \frac{n}{n-2} \cdot 0 + 2(0)^3
 \end{aligned}$$

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$$

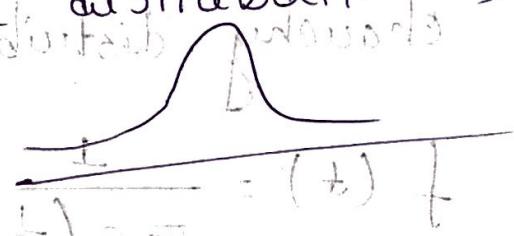
$$\begin{aligned}
 &= \frac{0 - 4 \cdot \frac{n}{n-2} \cdot 0 + 6 \cdot \frac{n}{n-2} \cdot 0^2 - 3 \cdot (0)^4}{(n-2)(n-4)} \\
 &= \frac{3n^2}{(n-2)(n-4)} \quad (\text{if } n > 4)
 \end{aligned}$$

$\therefore$  Skewness's 3' kurtosis  $\beta_1 = \frac{\mu_3}{\mu_2^3}$  residual distribution

residuals of residuals to half width  
 $= \frac{0}{(\frac{n}{n-2})^3} = 0$

Half residuals & remaining body having

If there  $\rho_1 = 0$  then the distribution is symmetric.



$$\text{Kurtosis } \beta_2 = \frac{\mu_4}{\mu_2^3}$$

$$= \frac{3n^2}{(n-2)(n-4)} \quad (L = n \text{ for fitting})$$

$$= \frac{(n+1)(n+2)}{(n-2)(n-4)} \cdot \frac{(n-1)(n-2)}{(n-2)(n-4)}$$

$$= \frac{3n^2}{(n-2)(n-4)} \times \frac{n^2}{n^2}$$

$$= \frac{3(n-2)}{(n-4)}$$

$$(n+1) = \frac{3n-6}{n-4}$$

$$= \frac{3n-12+6}{n-4}$$

$$= \frac{3(n-4)+6}{n-4}$$

$$\therefore \beta_2 = \left( 3 + \frac{6}{n-4} \right); n > 4$$

Here  $\beta_2 > 3$  if  $(n-4) <$

$\therefore t$  distribution is symmetrical & leptokurtic.

④ Show that  $t$  distribution is symmetric & leptokurtic.

⑤ Under what condition  $t$  distribution follows Cauchy distribution with  $\theta = 0$  and  $\beta = 1$ .

$$f(t) = \frac{1}{\sqrt{n} \rho \left(\frac{1}{2} + \frac{n}{2}\right) \left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}}, \text{ for } -\infty < t < \infty$$

putting  $n = 1$ ,

$$f(t) = \frac{1}{\sqrt{1} \rho \left(\frac{1}{2} + \frac{1}{2}\right) \left(1 + t^2\right)^{\frac{1+1}{2}}}$$

$$= \frac{1}{\rho \left(\frac{1}{2} + \frac{1}{2}\right) \left(1 + t^2\right)}$$

$$= \frac{\cancel{\Gamma(1/2 + 1/2)}}{\cancel{\Gamma(1/2)} \cancel{\Gamma(1/2)}} \times \frac{1}{(1 + t^2)}$$

$$= \frac{\cancel{\Gamma(1)}}{\cancel{\sqrt{\pi}} \cancel{\sqrt{\pi}} (1 + t^2)} \times \frac{1}{(1 + t^2)}$$

$$= \frac{1}{\cancel{\pi} \cancel{\sqrt{\pi}} (1 + t^2)} \times \frac{1}{(1 + t^2)}$$

$$= \frac{1}{\pi} \times \frac{1}{(1 + t^2)}$$

which is the form of the standard  
Cauchy distribution.

\* Under what condition the distribution follows  
standard normal distribution ( $t \rightarrow 0$ )

$$\beta_1 = 0$$

$$\beta_2 = 3 + \frac{n(G+1)}{n} = \left(\frac{n}{n+1}\right)^{-1}$$

when  $n \rightarrow \infty$   $\beta_2 \rightarrow 1$

so  $t \rightarrow 0$  as  $n \beta_2 \rightarrow 3$  limit point

When  $n$  is very large  $t$ -distribution  
tends to standard normal distribution.

$$\text{standard normal} \Rightarrow \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$



$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Date : 18.03.2017

We have,

$$f(t) = \frac{1}{\sqrt{n} \rho(\frac{1}{2}, \frac{n}{2})} \left(1 + \frac{t^n}{n}\right)^{\frac{n+1}{2}}$$

or,  $f(t) = \frac{\left(1 + \frac{t^n}{n}\right)^{\frac{n+1}{2}}}{\sqrt{n} \rho(\frac{1}{2}, \frac{n}{2})}$

$$= \frac{\left(1 + \frac{t^n}{n}\right)^{-n/2} \left(1 + \frac{t^n}{n}\right)^{-1/2}}{\sqrt{n} \rho(\frac{1}{2}, \frac{n}{2})}$$

Now taking limit on the both sides we get,

$$\text{L.H.S. } f(t) = \lim_{n \rightarrow \infty} \frac{\left(1 + \frac{t^n}{n}\right)^{-n/2} \left(1 + \frac{t^n}{n}\right)^{-1/2}}{\sqrt{n} \rho(\frac{1}{2}, \frac{n}{2})}$$

$$\text{R.H.S. } \lim_{n \rightarrow \infty} \frac{\sqrt{n} \frac{T^{\frac{1}{2}} T^{\frac{n}{2}}}{T^{\frac{1}{2} + \frac{n}{2}}}}{= \sqrt{n} \frac{T^{\frac{1}{2}}}{T^{\frac{1}{2}}} \lim_{n \rightarrow \infty} \frac{\frac{T^{\frac{n}{2}}}{\sqrt{n}}}{T^{\frac{1}{2} + \frac{n}{2}}}}$$

$$\lim_{n \rightarrow \infty} \frac{\frac{T^{n+1}}{\sqrt{n}}}{T^{\frac{n}{2}}} = n^*$$

$$= \frac{\left(2\sqrt{n} T^{\frac{1}{2}}\right)^{-\frac{1}{2}}}{\lim_{n \rightarrow \infty} \frac{T^{\frac{n}{2}}}{T^{\frac{n}{2}}}}$$

$$\begin{aligned}
 &= \frac{\sqrt{n} T^{\frac{1}{2}}}{\left(\frac{n}{2}\right)^{1/2}} \\
 &= \frac{\sqrt{n} T^{\frac{1}{2}}}{n^{1/2} \left(\frac{1}{2}\right)^{1/2}} \\
 &= \frac{T^{\frac{1}{2}}}{\left(\frac{1}{2}\right)^{1/2}}
 \end{aligned}$$

∴ probability of getting 1 or 2 successes

$$\text{Right most } \Rightarrow T^{\frac{1}{2}} (2)$$

$$\text{Probability } = \sqrt{2} \cdot \frac{T^{\frac{1}{2}}}{2} = \sqrt{2} \cdot \sqrt{T}$$

$$\text{to make } \sqrt{2} \cdot \frac{T^{\frac{1}{2}}}{2} = \sqrt{2} \cdot \sqrt{T}$$

$$\text{most probable value} = \sqrt{2T}$$

$$\text{Let } \left(1 + \frac{t^2}{n}\right)^{\frac{1}{2}} \xrightarrow[n \rightarrow \infty]{} \left(1 + \frac{t^2}{n}\right)^{\frac{1}{2}}$$

Again

$n \rightarrow \infty$  for minimum error with that of  $\sqrt{2}$

$$\text{Let } \left(1 + \frac{t^2}{n}\right)^{\frac{1}{2}} = e^m \quad \text{as part of binomial } t^2/2$$

$$\text{Let } \left(1 + \frac{t^2}{n}\right)^{\frac{1}{2}} \xrightarrow[n \rightarrow \infty]{} \left(1 + \frac{t^2}{n}\right)^{-1/2}$$

$$= \lim_{n \rightarrow \infty} \left\{ \left(1 + \frac{t^2}{n}\right)^{\frac{1}{n}} \right\}^{-1/2}$$

$$= \left( e^{t^2/2} \right)^{-1/2} = e^{-t^2/2}$$

$$= e^{-t^2/2}$$

$$\lim_{n \rightarrow \infty} f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-1/2} \left( \frac{x-\mu}{\sigma} \right)^2$$

$\mu, \sigma$

$$= \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}$$

Application of  $t$ -distribution :-

① To test if the sample mean  $\bar{x}$  differ significantly from the hypothetical value of mean of  $\mu_0$  to observe population mean

(single mean test)

② To test the significance of the difference bet<sup>n</sup> two sample sample means

(double mean test)

③ To test the significance of an observe sample correlation coefficient and sample regression coefficient.

⑩ To test the significance of an observed partial correlation coefficient.

## Properties of the t-distribution:

- ① Student's t-distribution is continuous type and its range is  $-\infty$  to  $+\infty$ .
- ② It has only one parameter (and it is known as degrees of freedom). (Unknown p. = 1 or)
- ③ The mean, median & mode are equal to zero and variance =  $\frac{n}{n-2}$  if  $n > 2$ .
- ④ If  $n=1$  then t-distribution tends to standard cauchy distribution.
- ⑤ t-distribution tends to standard normal distribution when degrees of freedom ( $n$ ) is very large.
- ⑥ All odd moments of t-distribution are zero.
- ⑦ The distribution is symmetric & leptokurtic.

+ - distribution about mean

show that mean deviation of + - distribution =  $\frac{\sqrt{n} \frac{Tn-1}{2}}{\sqrt{\pi} \frac{Tn}{2}}$

$E(t) = E(t) = 0$

$M.D(\bar{t})$  of distribution + - distribution =  $E|t - E(t)|$

m.d. about mean =  $E|t - E(t)|$

$M.D(\bar{x}) = \frac{1}{n} \sum |x_i - \bar{x}|$

MD (mean deviation to compare the measure of dispersion)

$$\text{long PE}(x) = \int_{-\infty}^{\infty} f(x) dx$$

$x < m(t)$

$$= \int_{-\infty}^{m(t)} f(t) dt$$

mean deviation + - result same as LL

$$\text{mean deviation} = \int_{-\infty}^{m(t)} t \cdot \frac{1}{\sqrt{n} \rho(1/2, n/2)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt$$

(constant to vanish makes mean deviation lower)

$$= 2 \int_0^{\infty} t \cdot \frac{1}{\sqrt{n} \rho(1/2, n/2)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt$$

disadvantage of  $\rho$  is it is not symmetric about 0

situation & situation in probability not

$$\text{Ldt}, \frac{dt^n}{n} = 2$$

$$\Rightarrow t^n = n^2$$

$$\Rightarrow 2t dt = n^2 dz$$

$$\Rightarrow t dt = \frac{n^2}{2} dz$$

wihen.  $t=0, z=0$

$t=\infty, z=\infty$

$$= \int_0^\infty \frac{n dz}{\rho(\frac{1+n}{2}) (1+z)^{\frac{n+1}{2}}}$$

zur Zeit  $\rightarrow$  Zeit  $\rightarrow$  Zeit  $\rightarrow$  Zeit

$$\text{Probabilität} = \frac{\sqrt{n}}{\rho(\frac{1+n}{2})} \int_0^\infty (1+z)^{\frac{n+1}{2}}$$

aus  $\rightarrow$  aus  $\rightarrow$  aus  $\rightarrow$  aus

$$\text{Probabilität} = \frac{\sqrt{n}}{\rho(\frac{1+n}{2})} \int_0^\infty \frac{2^{\frac{n+1}{2}}}{(1+z)^{\frac{(n+1)(n+1)}{2}}} dz$$

$$= \frac{\sqrt{n}}{\rho(\frac{1+n}{2})} \times \rho(\frac{(n+1)(n+1)}{2})$$

$$= \frac{\sqrt{n}}{\rho(\frac{1+n}{2})} \times \frac{\frac{n+1}{2}}{\frac{n+1}{2} + 1}$$

$$= \frac{\sqrt{n}}{\frac{\frac{1}{2} + \frac{n}{2}}{\frac{n+1}{2}}} \times \frac{\frac{n+1}{2}}{\frac{n+1}{2} + 1}$$

$$\begin{aligned}
 &= \frac{\sqrt{n} \times T_{\frac{n-1}{2}} \times T_{\frac{n+1}{2}}}{T_{\frac{1}{2}} T_{\frac{n}{2}} T_{\frac{n-1}{2}}} \\
 &= \frac{\sqrt{n} T_{\frac{n-1}{2}}}{\sqrt{n} T_{\frac{n}{2}}}
 \end{aligned}$$

Date: 21.03.2017

F-distribution

Defn: If  $F$  is defined as the ratio of two independent  $\chi^2$  random variable dividing by their respective degrees of freedom.  
 Let  $\chi^2_1$  and  $\chi^2_2$  are the two independent chi-square variate having  $n_1$  and  $n_2$  degree's of freedom respectively then the F. statistics defined as,

$$F = \frac{\chi^2_1/n_1}{\chi^2_2/n_2} \sim F(n_1, n_2)$$

The probability density function of  $F$  is defined as,

$$f(F) = \frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} F^{\frac{n_1}{2}-1}}{\Gamma\left(\frac{n_1}{2}, \frac{n_2}{2}\right) \Gamma\left(\frac{n_1+n_2}{2}\right)} \cdot \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right)}$$

$0 \leq F < \infty$

where  $n_1$  and  $n_2$  are the degrees of freedom of this distribution.

### Properties of the F-distribution

$n$ th raw moment of F-distribution is

$$\begin{aligned} \mu_n &= E(F^n) \\ &= \int_0^\infty F^n f(F) dF \\ &= \int_0^\infty F^n \left( \frac{n_1}{n_2} \right)^{\frac{n_1}{2}} F^{\frac{n_1}{2}-1} \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right)} (1 + \frac{n_1}{n_2} F)^{\frac{n_1+n_2}{2}-1} dF \\ &= \left( \frac{n_1}{n_2} \right)^{\frac{n_1}{2}} \int_0^\infty F^{\frac{n_1}{2}+n-1} \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right)} (1 + \frac{n_1}{n_2} F)^{\frac{n_1+n_2}{2}-1} dF \end{aligned}$$

Let,  $\frac{n_1}{n_2} F = U$       when  $F=0, U=0$

$$\text{or } F = \frac{n_2}{n_1} \cdot U$$

$$\text{or } dF = \frac{n_2}{n_1} du$$

$$\begin{aligned}
 &= \left( \frac{n_1}{n_2} \right)^{\frac{n_1}{2}} \cdot \int_0^{\infty} \frac{\left( \frac{n_2}{n_1} u \right)^{\frac{n_1}{2} + n - 1}}{(1+u)^{\frac{n_1+n_2}{2}}} \Gamma\left(\frac{n_2}{n_1}\right) du \\
 &\quad + P\left(\frac{n_1}{2}, \frac{n_2}{2}\right) \int_0^{\infty} (1+u)^{\frac{n_1+n_2}{2}} \left( \frac{n_2}{n_1} \right)^{\frac{n_1}{2} + n - 1} \Gamma\left(\frac{n_1}{2} + n - 1\right) \cdot \left( \frac{n_2}{n_1} \right)^n du \\
 &= \frac{\left( \frac{n_1}{n_2} \right)^{\frac{n_1}{2}}}{P\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \cdot \int_0^{\infty} \frac{\left( \frac{n_2}{n_1} \right)^{\frac{n_1}{2} + n - 1} u^{\frac{n_1}{2} + n - 1} \left( \frac{n_2}{n_1} \right)^n}{(1+u)^{\frac{n_1+n_2}{2}}} du
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{\left( \frac{n_1}{n_2} \right)^{\frac{n_1}{2}} \left( \frac{n_2}{n_1} \right)}{P\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \int_0^{\infty} \frac{u^{\frac{n_1}{2} + n - 1} \left( \frac{n_2}{n_1} \right)^{n_1 + n_2}}{(1+u)^{\frac{n_1+n_2}{2} + n}} du \\
 &= \left( \frac{n_2}{n_1} \right)^{-\frac{n_1}{2} + \frac{n_1}{2} + n} P\left(\frac{n_1}{2} + n, \frac{n_2}{2} + n\right) \\
 &\quad \times \frac{\beta\left(\frac{n_1}{2} + n, \frac{n_2}{2} + n\right)}{\Gamma\left(\frac{n_1}{2} + n\right) \Gamma\left(\frac{n_2}{2} + n\right)} \\
 &= \frac{\left( \frac{n_2}{n_1} \right)^n}{\frac{\Gamma\left(\frac{n_1}{2} + n\right)}{\Gamma\left(\frac{n_1+n_2}{2}\right)} \times \frac{\Gamma\left(\frac{n_1}{2} + n + \frac{n_2}{2} + n\right)}{\Gamma\left(\frac{n_1}{2} + n + \frac{n_2}{2} + n\right)}} \\
 &= \left( \frac{n_2}{n_1} \right)^n \frac{\Gamma\left(\frac{n_1}{2} + n\right) \Gamma\left(\frac{n_2}{2} + n\right)}{\Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right)}
 \end{aligned}$$

$$u_1' = \left( \frac{n_2}{n_1} \right)^{n_1} \frac{\Gamma(\frac{n_1+1}{2}) \Gamma(\frac{n_2-1}{2})}{\Gamma(\frac{n_1+n_2}{2})}$$

Putting  $n=1$ ,  $\frac{\Gamma(\frac{n_1+1}{2}) \Gamma(\frac{n_2-1}{2})}{\Gamma(\frac{n_1+n_2}{2})}$

$$u_1' = \text{mean} = \left( \frac{n_2}{n_1} \right) \frac{\frac{1}{2} \Gamma(\frac{n_1+1}{2}) \Gamma(\frac{n_2-1}{2})}{\Gamma(\frac{n_1+n_2}{2})}$$

$$= \frac{n_2}{n_1} \left( \frac{n_1+1-1}{2} \right) \frac{\Gamma(\frac{n_1+1-1}{2}) \Gamma(\frac{n_2-1}{2})}{\Gamma(\frac{n_1+n_2}{2})} \boxed{\Gamma(n-\frac{1}{2}) \Gamma(n-1)}$$

$$\cancel{\Gamma(\frac{n_1}{2})}, \cancel{\left( \frac{n_2-1}{2} \right)} \cancel{\Gamma(\frac{n_2}{2}-1)}$$

$$= \frac{\frac{n_2}{2}}{\frac{n_2}{2}-1}$$

$$= \cancel{\frac{n_2}{2}} \times \frac{2}{n_2-2}$$

$$= \frac{n_2}{n_2-2} \quad \text{if } n_2 > 2$$

Q. Find the mean of the F-distribution.

Putting  $n=2$

$$u_2' = \left( \frac{n_2}{n_1} \right)^{n_1} \frac{\Gamma(\frac{n_1+2}{2}) \Gamma(\frac{n_2-2}{2})}{\Gamma(\frac{n_1+n_2}{2})}$$

$$= \left( \frac{n_2}{n_1} \right)^2 \cdot \frac{\left( \frac{n_1+1}{2} \right) \left( \frac{n_1+1}{2} + 1 \right) \cdots \left( \frac{n_1+1}{2} + \frac{n_2-2}{2} \right)}{\left( \frac{n_1+1}{2} - 1 \right) \left( \frac{n_1+1}{2} - 2 \right) \cdots \left( \frac{n_1+1}{2} - \frac{n_2-2}{2} \right)}$$

$$= \frac{n_2^2}{n_1^2} \cdot \frac{\left( \frac{n_1+1}{2} + \frac{n_2-2}{2} + 1 \right) \left( \frac{n_1+1}{2} + \frac{n_2-2}{2} + 2 \right) \cdots \left( \frac{n_1+1}{2} + \frac{n_2-2}{2} + n_1 \right)}{\left( \frac{n_1+1}{2} - 1 \right) \left( \frac{n_1+1}{2} - 2 \right) \cdots \left( \frac{n_1+1}{2} - n_1 \right)} = \text{maxim} = 111$$

$$= \frac{n_2^2}{n_1^2} \cdot \frac{\left( \frac{n_1+1}{2} + \frac{n_2-2}{2} + 1 \right) \left( \frac{n_1+1}{2} + \frac{n_2-2}{2} + 2 \right) \cdots \left( \frac{n_1+1}{2} + \frac{n_2-2}{2} + n_1 \right)}{\left( \frac{n_1+1}{2} - 1 \right) \left( \frac{n_1+1}{2} - 2 \right) \cdots \left( \frac{n_1+1}{2} - n_1 \right)} =$$

$$= n \cdot \left( 1 - \frac{1}{2} \right) \cdot \frac{111}{111}$$

$$\frac{111}{111} \cdot \frac{1}{2} = \frac{111}{222}$$

Class 1:

$$= \frac{n_2^2 (n_1+2)}{n_1 (n_2-2)(n_2-4)} \cdot \frac{\left( \frac{n_1+1}{2} + \frac{n_2-2}{2} + 1 \right) \left( \frac{n_1+1}{2} + \frac{n_2-2}{2} + 2 \right) \cdots \left( \frac{n_1+1}{2} + \frac{n_2-2}{2} + n_1 \right)}{\left( \frac{n_1+1}{2} - 1 \right) \left( \frac{n_1+1}{2} - 2 \right) \cdots \left( \frac{n_1+1}{2} - n_1 \right)} = \text{maxim} = 111$$

$$= \frac{n_2^2 (n_1+2)}{n_1 (n_2-2)(n_2-4)} \cdot \frac{\left( \frac{n_1+1}{2} + \frac{n_2-2}{2} + 1 \right) \left( \frac{n_1+1}{2} + \frac{n_2-2}{2} + 2 \right) \cdots \left( \frac{n_1+1}{2} + \frac{n_2-2}{2} + n_1 \right)}{\left( \frac{n_1+1}{2} - 1 \right) \left( \frac{n_1+1}{2} - 2 \right) \cdots \left( \frac{n_1+1}{2} - n_1 \right)} = \text{maxim} = 111$$

sp. o. s. s. ; o. b. t.

we know,

$$\text{variance} = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2 + \sum_{i=n_1+1}^{n_1+n_2} (x_i - \bar{x}_2)^2}{n_1 + n_2 - 2}$$
$$= \frac{n_1 \bar{x}_1^2 + (n_1+2) \bar{x}_1^2 + n_2 \bar{x}_2^2 + (n_2-2) \bar{x}_2^2 - 2(n_1 \bar{x}_1 \bar{x}_2 + n_2 \bar{x}_1 \bar{x}_2)}{n_1(n_2-2)(n_2-4)}$$
$$= \frac{n_1 \bar{x}_1^2 + n_2 \bar{x}_2^2 - n_1 \bar{x}_1^2 - n_2 \bar{x}_2^2 + 2(n_1 \bar{x}_1 \bar{x}_2 + n_2 \bar{x}_1 \bar{x}_2)}{n_1(n_2-2)(n_2-4)}$$
$$= \frac{n_1 \bar{x}_1^2 + n_2 \bar{x}_2^2 - 2n_1 \bar{x}_1^2 - 2n_2 \bar{x}_2^2 + 2n_1 \bar{x}_1 \bar{x}_2 + 2n_2 \bar{x}_1 \bar{x}_2}{n_1(n_2-2)(n_2-4)}$$
$$= \frac{n_1 \bar{x}_1^2 + n_2 \bar{x}_2^2 - 2n_1 \bar{x}_1^2 - 2n_2 \bar{x}_2^2 + 2n_1 \bar{x}_1 \bar{x}_2 + 2n_2 \bar{x}_1 \bar{x}_2}{n_1(n_2-2)(n_2-4)}$$
$$= \frac{n_1 \bar{x}_1^2 + n_2 \bar{x}_2^2 - 2n_1 \bar{x}_1^2 - 2n_2 \bar{x}_2^2 + 2n_1 \bar{x}_1 \bar{x}_2 + 2n_2 \bar{x}_1 \bar{x}_2}{n_1(n_2-2)(n_2-4)}$$
$$= \frac{2 \cdot n_2 \bar{x}_2^2 + (n_1 + n_2 - 2) \bar{x}_1^2 + 2n_1 \bar{x}_1 \bar{x}_2 + 2n_2 \bar{x}_1 \bar{x}_2}{n_1(n_2-2)(n_2-4)}$$

if  $n_2 > 4$

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

Date: 28.03.2017

## Mode

We have

$$f(F) = \frac{\left(\frac{n_1}{n_2} + 1\right)^{\frac{n_1}{2}} \left(\frac{n_2}{n_1} - 1\right)^{\frac{n_2}{2}}}{P \left(\frac{n_1}{2}\right) \left(\frac{n_2}{2}\right) \left(1 + \frac{n_1 - n_2}{n_1 + n_2} F\right)^{\frac{n_1 + n_2}{2}}}$$

$$f(F) = c \cdot \frac{\left(\frac{n_1}{n_2} + 1\right)^{\frac{n_1}{2}} \left(\frac{n_2}{n_1} - 1\right)^{\frac{n_2}{2}}}{\left(1 + \frac{n_1 - n_2}{n_1 + n_2} F\right)^{\frac{n_1 + n_2}{2}}} \quad 0 < F < \infty$$

where  $c = \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}$

$$P\left(\frac{n_1}{2}, \frac{n_2}{2}\right)$$

(W.P. - W.M.)  $\propto$   $(P - \text{parallel plane wave})^2$

Taking log on both sides we get:

$$\log f(F) = \log c + \left(\frac{n_1 - 1}{2}\right) \log F$$

$$\left(\frac{n_1 + n_2}{2}\right) \log \left(1 + \frac{n_1}{n_2} F\right)$$

$$\log \left(\frac{A}{B}\right) = \log A - \log B$$

Now differentiating w.r.t.  $F$ ,

$$\frac{\delta \log f(F)}{\delta F} = \Theta \left(\frac{n_1}{2} - 1\right) \frac{1}{F} - \left(\frac{n_1 + n_2}{2}\right) \frac{1}{1 + \frac{n_1}{n_2} F} \times \frac{n_1}{n_2}$$

$$\text{or, } \frac{n_1 - 2}{2F} - \left(\frac{n_1 + n_2}{2}\right) \times \frac{1}{\frac{n_1 + n_2}{n_2}} \times \frac{n_1}{n_2} = 0 \quad f'(F) = 0$$

$$\text{or, } \frac{n_1 - 2}{F} - \frac{n_1 + n_2}{2(n_2 + n_1 F)} = 0$$

$$\text{or, } \frac{n_1 - 2}{\cancel{F}} = \frac{n_1^m + n_1 n_2}{(n_2 + n_1 F)} \quad \text{now find initial}$$

$$\text{or, } (n_1 - 2)(n_2 + n_1 F) = (n_1^2 + n_1 n_2)F \quad \text{by}$$

$$\text{or, } n_1 n_2 + \cancel{n_1 F} = 2^n (2 - \cancel{2 n_1 F}) = \cancel{n_1 F} + n_1 m_2 F$$

$$\text{or, } F_b(2n_1 + n_2) = n_1 n_2 - 2n_2$$

$$\text{6th, } F_6 = \frac{\frac{1}{2}b(n_1n_2 - 2n_2)}{2n_1 + n_2} [v] = (\frac{1}{2})b = (\frac{1}{2})+$$

$$t_{k_2} = \frac{\left( \begin{matrix} n_1 \\ n_2 \end{matrix} \right) + \left( \begin{matrix} n_2 \\ n_1 \end{matrix} \right)}{n_1(n_1 + n_2)}$$

$$= \frac{n_1 + 2}{n_1} \cdot \frac{n_2 + 1}{2 + n_2} \quad \text{if } n_1 > 2$$

It can be easily verified that at the point  $F \approx \left(\frac{n_1 - \alpha}{n_1}\right) \left(\frac{n_2}{2 + n_2}\right) \approx \left(1 - \frac{1}{n_1}\right)^2$

$f''(F) < 0$ . Hence, ~~mod~~

$$\text{Mode} = \left( \frac{n_1 - 2}{n_1} \right) \times \left( \frac{n_2}{2 + n_2} \right)$$

Relation between ' $t$ ', and 'F' distribution :

We have,  $f(F) = \frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}}{\beta\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} F^{\frac{n_1}{2}-1} (1-F)^{\frac{n_2}{2}-1}$ ,  $0 < F < 1$

putting,  $n_1 = 1$ ,  $n_2 = n$  (and  $t^2 = F$ )  
or  $2t dt = dF$

$$f(t) = f(F) = |J| \Rightarrow \frac{dF}{dt} = 2t$$

Determining 'J':

$$J = \frac{dF}{dt} = \frac{d}{dt} \left[ \frac{1}{2} t^{\frac{1}{2}} \right] = \frac{1}{2} t^{-\frac{1}{2}}$$

$$= \left( \frac{1}{n} \right)^{\frac{1}{2}} \left( t^{\frac{1}{2}} \right)^{\frac{1}{2}-1}$$

$$\frac{\beta\left(\frac{1}{2}, \frac{n}{2}\right) \left(1 + \frac{t^n}{n}\right)^{\frac{1+n}{2}}}{\left(1 + \frac{t^n}{n}\right)^{\frac{n+1}{2}}} \quad \begin{cases} t^{\frac{1}{2}} = F \\ F = 0, t = 0 \end{cases}$$

$$f(t) = \frac{1}{\sqrt{n} \beta\left(\frac{1}{2}, \frac{n}{2}\right) \left(1 + \frac{t^n}{n}\right)^{\frac{n+1}{2}}} \quad \begin{cases} 0 < t < \infty \\ F = 0, t = \infty \end{cases}$$

$$= \frac{1}{\sqrt{n} \beta\left(\frac{1}{2}, \frac{n}{2}\right) \left(1 + \frac{t^n}{n}\right)^{\frac{n+1}{2}}} \quad \begin{cases} 0 < t < \infty \\ F = 0, t = \infty \end{cases}$$

which is one of the  $t$ -distribution  
with  $n$  degrees of freedom.

Relation between F and  $\chi^2$  distribution.

We have,

$$f(F) = \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}-1} F^{\frac{n_1}{2}-1}$$

$$\rho\left(\frac{n_1}{2}, \frac{n_2}{2}\right) \left(1 + \frac{n_1}{n_2} F\right)^{\frac{n_1+n_2}{2}}$$

$$\frac{T_{\frac{n_1}{2}} + \frac{n_2}{2}}{T_{\frac{n_1}{2}}} \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}-1} F^{\frac{n_1}{2}-1} \cdot \rho\left(\frac{n_1}{2}, \frac{n_2}{2}\right)$$

$$\rho(m, n) = \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)}$$

Now  $\lim_{n_2 \rightarrow \infty}$

$$\frac{\left(\frac{T_{\frac{n_1}{2}} + \frac{n_2}{2}}{T_{\frac{n_1}{2}}}\right)^{\frac{n_1}{2}}}{\left(\frac{n_2}{2} \cdot (n_2)\right)^{\frac{n_1}{2}}} \xrightarrow[n_2 \rightarrow \infty]{} \frac{\left(\frac{T_{\frac{n_1}{2}}}{T_{\frac{n_1}{2}}}\right)^{\frac{n_1}{2}}}{\left(\frac{n_1}{2}\right)^{\frac{n_1}{2}}} = 1$$

$$\text{or, } \lim_{n_2 \rightarrow \infty} \frac{\frac{T_{\frac{n_1}{2}} + \frac{n_2}{2}}{T_{\frac{n_1}{2}}}}{\frac{n_2}{2}} \xrightarrow[n_2 \rightarrow \infty]{} \frac{\frac{1}{1}}{\left(\frac{n_1}{2}\right)^{\frac{n_1}{2}}} = \frac{\left(\frac{n_1}{2}\right)^{\frac{n_1}{2}}}{T_{\frac{n_1}{2}}}$$

$$\lim_{n \rightarrow \infty} \frac{\frac{T_{n+k}}{T_n}}{n} \rightarrow n^k$$

$$= \frac{\left(\frac{n_2}{2}\right)^{\frac{n_1}{2}}}{\left(\frac{n_2}{2}\right)^{\frac{n_1}{2}}} \times \frac{\left(\frac{n_1}{2}\right)^{\frac{n_1}{2}}}{T_{\frac{n_1}{2}}}$$

equilibrium  $\Rightarrow$  form of expected capital  $\underline{F}$

$$\frac{(n_1)^{\frac{n_1}{2}}}{\frac{n_1}{2}! \left( \frac{n_1}{2} + \frac{n_2}{2} \right)_{n_2}^{n_2}} = (1) +$$
$$(1 + \frac{n_1}{2}) \left( \frac{n_1}{2} + \frac{n_2}{2} \right)_{n_2}^{n_2}$$

Again,

$$dt \cdot \left( (n_1)^{\frac{n_1}{2}} \frac{n_1}{n_2} F \right) = \lim_{n_2 \rightarrow \infty} \left( 1 + \frac{n_1}{n_2} F \right)^{\frac{n_1}{2}}$$

$$\underset{n_2 \rightarrow \infty}{\cancel{dt}} \left( 1 + \frac{n_1}{n_2} F \right)^{\frac{n_1}{2}}$$

$$= \lim_{n_2 \rightarrow \infty} \left( 1 + \frac{n_1}{n_2} F \right)^{\frac{n_1}{2}} dt \left\{ \left( 1 + \frac{n_1}{n_2} F \right)^{\frac{n_1}{2}} \right\}$$

---

$$\left\{ \left( 1 + \frac{a}{n} \right)^n \right\}^{\frac{1}{2}} = e^{\frac{a}{2}}$$

1.09.2017

$$\lim_{n \rightarrow \infty} \left\{ \left(1 + \frac{a}{n}\right)^n \right\}^{\frac{1}{n_2}} = e^{ax/2}$$

$$\lim_{n \rightarrow \infty} \left\{ \left(1 + \frac{n_1 F}{n_2}\right)^{\frac{n_2}{2}} \right\} = 1$$

$$\lim_{n_2 \rightarrow \infty} \left[ \left(1 + \frac{n_1 F}{n_2}\right)^{\frac{n_2}{2}} \right]^{\frac{1}{n_2}} \times \lim_{n_2 \rightarrow \infty} \left[ \left(1 + \frac{n_1 F}{n_2}\right)^{\frac{n_2}{2}} \right]$$

is meant and is defined  
to write  $= e^{ax/2}$  as indicated.

Putting all these values in equation (4)

$$f(F) = \frac{\left(\frac{n_1}{2}\right)^{\frac{n_1}{2}} \cdot \left(\frac{n_1}{2}-1\right)}{2^{\frac{n_1}{2}} \cdot \frac{n_1}{2}!} \cdot e^{\frac{n_1 F}{2}}$$

associated probability of finding  $n_1$  atoms to

$$\text{Again, } dt, (nmF) = x^2$$

$$\Rightarrow F = \frac{x^2}{n_1}$$

$$\Rightarrow \frac{dF}{dx^2} = \frac{1}{n_1}$$

$$\therefore f(F) = \frac{\left(\frac{n_1}{2}\right)^{\frac{n_1}{2}} \left(\frac{x^2}{n_1}\right)^{\frac{n_1}{2}-1}}{2^{\frac{n_1}{2}} \cdot \frac{n_1}{2}!} \cdot e^{\frac{n_1 F}{2}}$$

$$= \frac{\left(\frac{n_1}{2}\right)^{\frac{n_1}{2}} \times (n_1^{-1})^{\frac{n_1}{2}-1} (x^2)^{\frac{n_1}{2}-1} (n_1^{-1})^{-\frac{x^2}{2}}}{2^{\frac{n_1}{2}} \cdot \frac{n_1}{2}!}$$

$$f(x^v) = \frac{(n_1)^{\frac{n_1}{2} - \frac{n_1+1}{2}-1} (x^v)^{\frac{n_1}{2}-1} e^{-\frac{x^v}{2}}}{2^{\frac{n_1}{2}} \Gamma(\frac{n_1}{2})}$$

$$= \frac{1}{2^{\frac{n_1}{2}} \Gamma(\frac{n_1}{2})} e^{-\frac{x^v}{2}} (x^v)^{\frac{n_1}{2}-1} \quad (x > 0, 0 < x < \infty)$$

which is the form of the ~~gen formula~~  
with  $x^v$ -distribution with  $n_1$  degree of  
freedom

\* Relation between F-distribution and P  
of second kind

$$\int_0^\infty \frac{x^{\frac{n}{2}-1}}{(1+x)^{\frac{n+m}{2}}} \Rightarrow p(m, n)$$

$$f(x) = \frac{1}{p(m, n)} \cdot \frac{x^{\frac{n}{2}-1}}{(1+x)^{\frac{n+m}{2}}}$$

We have, the pdf of P of second  
kind is

$$(1-m)x^{m-1}(1-x)^{1-n}$$

$$f(x) = \frac{x^{\frac{n_1}{2}-1}}{x^{\frac{n_1}{2}} + (1+\kappa)^{\frac{n_1+n_2}{2}}} \quad x > 0$$

$$f(x) = \frac{1}{\beta(n_1, n_2)} x^{\frac{n_1}{2}-1} (1+\kappa)^{\frac{n_1+n_2}{2}} \quad x > 0$$

Let,  $\frac{n_1}{n_2} F = x$

$$\Rightarrow \frac{dx}{dF} = \frac{n_1}{n_2}$$

$$\therefore f(x) = \frac{1}{\beta(n_1, n_2)} x^{\frac{n_1}{2}-1} (1 + \frac{n_1}{n_2} F)^{\frac{n_1+n_2}{2}} \times \frac{\frac{n_1}{n_2} - 1}{(1 + \frac{n_1}{n_2} F)^{\frac{n_1+n_2}{2}}} \times \frac{n_1}{n_2}$$

$$= \frac{1}{\frac{n_1}{n_2} \Gamma(\frac{n_1}{2}, \frac{n_2}{2})} \times \left( \frac{n_1}{n_2} \right)^{\frac{n_1}{2}-1} F^{\frac{n_1}{2}-1} \left( \frac{n_1}{n_2} \right)' \times \frac{(1 + \frac{n_1}{n_2} F)^{\frac{n_1+n_2}{2}}}{(1 + \frac{n_1}{n_2} F)^{\frac{n_1+n_2}{2}}} \times \frac{n_1}{n_2}$$

$$f(F) = \frac{\left( \frac{n_1}{n_2} \right)^{\frac{n_1}{2}-1} F^{\frac{n_1}{2}-1}}{\beta\left(\frac{n_1}{2}, \frac{n_2}{2}\right) (1 + \frac{n_1}{n_2} F)^{\frac{n_1+n_2}{2}}} \quad \left| \begin{array}{l} \gamma = \frac{dn}{dt} = \frac{n_1}{n_2} \\ f(F) = f(x)|\gamma| \end{array} \right.$$

\* If  $x_1$  and  $x_2$  are two independent random variable having common density function,

$$f(x) = e^{-x} \quad 0 < x < \infty$$

Show that  $u = \frac{x_1}{x_2}$  has an  $F$ -distribution

$$f(x_1) = e^{-x_1}$$

$$f(x_2) = e^{-x_2}$$

The joint distribution of  $x_1$  and  $x_2$  is,

$$f(x_1, x_2) = f(x_1) \cdot f(x_2) = e^{-x_1} \cdot e^{-x_2} = e^{-(x_1 + x_2)}$$

$$\text{Let } U = \frac{x_1}{x_2}$$

$$\text{and } V = x_1 + x_2$$

$$\text{or, } x_1 = ux_2$$

$$(1), x_2 = v - x_1$$

$$\text{or, } x_2 = \frac{v - u - x_1}{1+u}$$

$$\Rightarrow x_2 = \frac{v}{u+1}$$

$$x_1 = u \cdot x_2 = u \cdot \frac{v}{u+1}$$

$$= \frac{uv}{u+1}$$

$$J(f) = \begin{vmatrix} s(x_1, x_2) \\ s(uv) \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial x_1}{\partial u} & \frac{\partial x_1}{\partial v} \\ \frac{\partial x_2}{\partial u} & \frac{\partial x_2}{\partial v} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial x_1}{\partial u} & \frac{\partial x_1}{\partial v} \\ \frac{\partial x_2}{\partial u} & \frac{\partial x_2}{\partial v} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{x(1+u) - uv}{(1+u)} & \frac{u}{1+u} \\ -\frac{v}{(1+u)} & \frac{1}{1+u} \end{vmatrix} = \frac{1}{(1+u)^2} \begin{vmatrix} x(1+u) - uv & u \\ -v & 1 \end{vmatrix}$$

is best. & diff. w. mitbedeutende + verringert die Anzahl der Werte

$$= \frac{v}{(1+u)^3} + \frac{uv}{(1+u)^3} \cdot \text{mehrere } \rightarrow \text{wiederholte Werte}$$

$$= \frac{u+v}{(1+u)^3} = \frac{u(1+u)}{(1+u)^3} = \frac{v}{(1+u)^2}$$

$$f(x_1, x_2) = e^{-(x_1+x_2)}$$

$$g(uv) = f(x_1, x_2) | \uparrow \gamma |$$

$$= e^{-v} \frac{v}{(1+u)^2}$$

$$\left| \begin{array}{l} 0 < u < \infty \\ 0 < v < \infty \end{array} \right.$$

$$f(u) = \int_0^\infty e^{-v} \frac{v}{(1+u)^2} dv$$

$$= \frac{1}{(1+u)^2} \int_0^\infty v \cdot e^{-v} dv$$

$$\begin{aligned}
 &= \frac{1}{(1+u)^n} \left| \int_0^\infty e^{-\frac{1}{2}v^2(2+(2u))} v^{n-1} dv \right| \\
 &= \frac{1}{(1+u)^n} \cdot T_2 \\
 &= \frac{1}{(1+u)^n} \frac{1}{(2-1) \binom{n+1}{2-1}} \frac{v^{\frac{n}{2}-1}}{\sqrt{(n+1)}} \\
 &= \frac{1}{(1+u)^n} \frac{1}{(n+1)} \frac{v^{\frac{n}{2}-1}}{\sqrt{(n+1)}} \\
 &= \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}-1} u^{\frac{n}{2}-1}}{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{n+1}{2}\right) \left(1 + \frac{1}{2}u\right)^{\frac{n+1}{2}}} = 
 \end{aligned}$$

which is follow f-distribution with 2 and 2 degrees of freedom.  $\frac{v_1}{\sigma^2(1+1)} + \frac{v_2}{\sigma^2(1+1)}$

$$\frac{v}{\sqrt{u+1}} \cdot \frac{(n+1)v}{\sigma^2(u+1)} = \frac{v(n+1)}{\sigma^2(u+1)} =$$

$$(X_1 + X_2) \rightarrow (X_1 + X_2) +$$

$$v_1 > v_2 > 0 \quad | \quad |C| (\sin x), \sin(x_1, x_2)$$

$$v_1 \cdot \frac{v}{\sqrt{u+1}} \cdot v_2 \cdot \frac{v}{\sqrt{u+1}} = (u+1)^{-1}$$

$$\sqrt{u+1} \cdot v_1 \cdot v_2 \cdot \frac{v}{\sqrt{u+1}} = (u+1)^{-1}$$

Date: 04.04.2017

$$\beta(m, n) = \int_0^{\infty} x^{m-1} (1+x)^{-n-m} dx$$

$$\beta\left(\frac{m}{2}, \frac{n}{2}\right) = \int_0^{\infty} x^{\frac{m}{2}-1} (1+x)^{\frac{n}{2}} dx$$

We have,

$$f(x) = \frac{(x)}{\beta\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} (1+x)^{\frac{n_1+n_2}{2}}$$

### Properties of F-distribution:

- (i) It has a continuous type of distribution and its range is  $0 \rightarrow \infty$ .
- (ii) The mean and variance of F-distribution are  $\frac{n_2}{n_2-1}$  if  $n_2 > 2$  and  $\frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}$  if  $n_2 > 4$ .
- (iii) The mode of the F-distribution is  $\left(\frac{n_2}{n_2+2}\right)\left(\frac{n_1-2}{n_1}\right)$
- (iv) If  $n_1$  and  $n_2$  are very large which tends to infinite then F-distribution tends to normal distribution.
- (v) The distribution is positive skewed and leptokurtic.
- (vi) If  $F \sim F(n_1, n_2)$  then  $\frac{n_1}{n_2} F \sim \beta_2\left(\frac{n_1}{2}, \frac{n_2}{2}\right)$
- (vii) If  $F \sim F(n_1, n_2)$  then  $\frac{1}{F} \sim F(n_2, n_1)$

## The application of f-distribution

- ① It is used for testing the equality of two population variances [ $H_0: \sigma_1^2 = \sigma_2^2$ ]
- ② It is used for testing the significance of an observed multiple correlation coefficient.
- ③ It is used for testing the significance of sample correlation ratio. [ $r_{yx}$ ]
- ④ It is used for testing linearity of regression.
- ⑤ F-distribution is used to test the equality of several means.

(The F-distribution will be shown with distribution density graph and its behaviour by different methods and results obtained from various distributions known as beta distribution and gamma distribution.)

Date: 8.04.20.

Estimation ~~and~~ ~~in~~ ~~statistical~~ ~~Inference~~ ~~is~~ ~~to~~ ~~estimate~~ ~~population~~ ~~parameters~~ ~~by~~ ~~sample~~ ~~data~~ ~~using~~ ~~statistical~~ ~~method~~ ~~and~~ ~~techniques~~ ~~of~~ ~~estimation~~ ~~and~~ ~~testing~~ ~~hypothesis~~ ~~testing~~ ~~as~~ ~~well~~ ~~as~~ ~~confidence~~ ~~intervals~~ ~~etc.~~

It is the science of drawing valid conclusion from the sample data about the population, from which the sample has been drawn. If it includes the procedures of estimating population parameter and testing whether a tentative statement about a population parameter is supported by evidence from sample elements to both in statistical estimation.

The methods of making judgement about a population parameter are called statistical estimation. There are two types of statistical estimation → ① point estimation ② Interval estimation.

point estimation : point estimation involves the use of sample data to calculate single value which is to be such as serve as best estimate of an unknown population parameter.

If from the sample value at a single value is calculated as an estimate of the population parameter the procedure is termed as point estimation.

### Methods of point estimation:

- ① Method of Maximum Likelihood Estimation (MLE)
- ② Method of moments
- ③ Method of least square
- ④ Method of minimum variance
- ⑤ Method of minimum chi-square
- ⑥ Method of Inverse Probability

### Criteria of good estimator:

- ① Unbiasedness
- ② Consistency
- ③ Efficiency
- ④ Sufficient

Parameter statistic

Estimator

Estimate

Parameters: For drawing valid inference about the population, we in practice deal with sample and obtained the estimate of the population characteristics. The unknown characteristic of the population usually known as parameter.

A statistic is generally a function of a set of sample values. Clearly statistic is a random variable.

Estimator: If any statistic is used to estimate and unknown parameter  $\theta$  /  $\mu$  /  $\sigma^2$  /  $\gamma$  of the distribution, it is called an estimator.

Estimate: Any particular value of the estimator say  $T_n = T(x_1, x_2, \dots, x_n)$  is called an estimate of  $\theta$ .

QUESTION: ~~Find the moment generating function for  $\chi^2$  distribution~~

1. Find the momentary generation function of  $\chi^2$  distribution.

(Ans): M<sub>X^2</sub>(t) =  $E(e^{tX^2})$  set to obtain the waiting time  $t = \int_0^\infty e^{tx^2} f(x^2) dx^2$  now we can write

rearranging as given below

to wait  $t + x^2$  will require  $e^{-\frac{tx^2}{2}} \cdot \frac{n}{2} \Gamma(\frac{n}{2})$  if  $x^2$  follows  $\chi^2$  distribution

so  $t + x^2$  follows  $\chi^2_{n+2}$  distribution by the

$$\text{of above } \frac{1}{2} \int_0^\infty e^{-\frac{x^2}{2}(1-2t)} \cdot \frac{n}{2} \Gamma(\frac{n}{2}) dx^2$$

Now  $E$  of  $\chi^2$  distribution

as  $t = \text{waiting time} \cdot \frac{1}{2}$  from abov

$$2 \frac{n}{2} \Gamma(\frac{n}{2}) \cdot \left( \frac{1-2t}{2} \right)^{\frac{n}{2}}$$

which is  $\frac{1}{2}$  of  $\chi^2$  distribution

so  $\chi^2$  distribution has  $\frac{(1-2t)^{\frac{n}{2}}}{2}$  as abov

$$= (1-2t)^{\frac{n}{2}}$$

2. Find the cumulative generation function: and show that  $\chi^2$ -distribution is always positive skewed and leptokurtic.

$$\text{Ans: } K_{xx}(t) = \log M_{xx}(t)$$

$$\begin{aligned}
 &= \log (1-2t)^{-\frac{n}{2}} \\
 &= -\frac{n}{2} \log 2 + \log(1-2t) \\
 &= -\frac{n}{2} \left[ -2t - \frac{(2t)^2}{2} - \frac{(2t)^3}{3} - \frac{(2t)^4}{4} - \dots \right] \\
 &= n + 2n \cdot \frac{t^2}{2} + 4n \cdot \frac{t^3}{3} + 8n \cdot \frac{t^4}{4} + \dots
 \end{aligned}$$

coefficient of  $K_1 = \frac{t}{1!} = n = \mu_1 = \text{mean}$

$$\therefore \therefore K_2 = \frac{t^2}{2!} = 2n = \mu_2 = \text{variance}$$

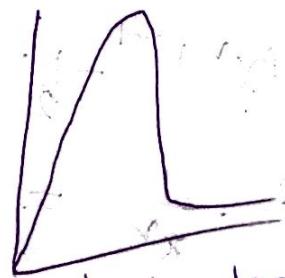
$$\therefore \therefore K_3 = \frac{t^3}{3!} = 8n = \mu_3$$

$$\therefore \therefore K_4 = \frac{t^4}{4!} = 48n = \mu_4$$

$$\begin{aligned}
 \mu_3 &= K_3 + 3K_2 \\
 &= 48n + 3 \cdot 4n^2 \\
 &= 48n + 12n^2
 \end{aligned}$$

$$\beta_1 = \frac{\mu_3}{\mu_2^3} = \frac{\mu_3}{\mu_2^3} = \frac{(8n)^3}{(2n)^3} = \frac{64n^3}{8n^3} = \frac{8}{n}$$

$$\begin{aligned}
 \beta_2 &= \frac{\mu_4}{\mu_2^2} = \frac{48n + 12n^2}{4n^2} \\
 &= 3 + \frac{12}{n}
 \end{aligned}$$



$\beta_1 > 0, \beta_2 > 3$ , positive skewed & leptokurtic

### 3. Mode of $x^n$ -distribution:

We know that,

$$f(x^n) = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} e^{-\frac{x^n}{2}} (x^n)^{\frac{n}{2}-1}$$

Mode is that value of  $x^n$ -distribution which

satisfy the equation,

$$f'(x^n) = 0 \quad \text{and} \quad f''(x^n) < 0$$

Now,

$$\begin{aligned} f'(x^n) &= \frac{d}{dx^n} \left[ \frac{1}{2^{n/2} \Gamma(n/2)} e^{-\frac{x^n}{2}} (x^n)^{\frac{n}{2}-1} \right] \\ &= \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \left[ e^{-\frac{x^n}{2}} (\frac{n}{2}-1)(x^n)^{\frac{n}{2}-2} \right. \\ &\quad \left. + (x^n)^{\frac{n}{2}-1} e^{-\frac{x^n}{2}} (-\frac{1}{2}) \right] \\ &= \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} e^{-\frac{x^n}{2}} (x^n)^{\frac{n}{2}-1} \left[ (\frac{n}{2}-1)(x^n)^{-\frac{1}{2}} \right] \end{aligned}$$

puts  $f'(x^n) = 0$

$$e^{-\frac{x^n}{2}} (x^n)^{\frac{n}{2}-1} \left[ (\frac{n}{2}-1)(x^n)^{-\frac{1}{2}} \right] = 0$$

$$(\frac{n}{2}-1)(x^n)^{-\frac{1}{2}} = 0$$

$$\Rightarrow \frac{n-2}{2} \cdot \frac{1}{x^n} = \frac{1}{2}$$

$$x^n = n-2$$

It is easy to show that at the point

$$x^r = n-2, f''(x^r) < 0$$

Hence the mode of the  $x^r$ -distribution  
is  $(n-2)$  with ~~form~~ degree of freedom.

Q State and prove:  $x^r$ -variate.

Ans: The sum of independent  $x^r$ -variate

is also a  $x^r$ -variate. Let  $x_1^r, x_2^r \dots$   
 $x_k^r$  be the  $k$  independent var variate  
with  $n_1, n_2, \dots, n_k$  degree of freedom

the sum  $\sum_{i=1}^k x_i^r$  is a  $x^r$ -variat with

$\sum_{i=1}^k n_i$  degree of freedom.

We know the momentary generation function

of  $x^r$ -distribution is

$$M_{x^r}(t) = (1-\alpha)^{-\frac{n_i}{2}}$$

since  $x_1^r, x_2^r \dots x_k^r$  be the  $k$  independent  
variate with  $n_1, n_2, \dots, n_k$  degree of freedom.

The m.g.f of  $\sum_{i=1}^k x_i^r$  is written as

$$\begin{aligned}
 M \sum_{i=1}^k x_i^n(t) &= \frac{(1-2t)^{\frac{n_1}{2}}}{k} \cdot M_{x_2}(t) \cdot \dots \cdot M_{x_K}(t) \\
 &\stackrel{\text{Multiplication}}{=} M_{x_1^n}(t) \cdot M_{x_2^n}(t) \cdot \dots \cdot M_{x_K^n}(t) \\
 &= (1-2t)^{\frac{n_1}{2}} \cdot (1-2t)^{\frac{n_2}{2}} \cdot \dots \cdot (1-2t)^{\frac{n_K}{2}} \\
 &= (1-2t)^{\frac{n_1 + n_2 + \dots + n_K}{2}}
 \end{aligned}$$

which is momentary function of  $x_1^n$  variate

with  $n_1, n_2 \dots$  as degree of freedom

so momentary function of  $\sum_{i=1}^k x_i^n$  is also a  $x^n$  variate with ~~and its~~  $\sum_{i=1}^k n_i$  degrees of freedom.

x

$$f(t) = \frac{1}{\sqrt{n} \cdot p_0 \left(\frac{1}{2}, \frac{n}{2}\right) \left(1 + \frac{t^n}{n}\right)^{\frac{n+1}{2}}}$$

$$u'_2 = n^n \frac{\overline{T_{\frac{n}{2}-1}} \overline{T_{1+\frac{1}{2}}}}{\overline{T_{\frac{1}{2}}} \overline{T_{\frac{n}{2}}}}$$

$n=1$

$$\begin{aligned} u'_2 &= n^n \frac{\overline{T_{\frac{n}{2}-1}} \overline{T_{1+\frac{1}{2}}}}{\overline{T_{\frac{1}{2}}} \overline{T_{\frac{n}{2}}}} \\ &= \frac{n \cancel{\overline{T_{\frac{n}{2}-1}}} \left(\frac{3}{2}-1\right) \cancel{\overline{T_{\frac{3}{2}-1}}}}{\overline{T_{\frac{1}{2}}} \left(\frac{1}{2}-1\right) \cancel{\overline{T_{\frac{n}{2}-1}}}} \\ &= \frac{n}{\cancel{n}} \cdot \frac{\cancel{n}}{\cancel{n-2}} \\ &= \frac{n}{n-2} \end{aligned}$$

$n=2$

$$\begin{aligned} u'_4 &= n^n \frac{\overline{T_{\frac{n}{2}-2}} \overline{T_{2+\frac{1}{2}}}}{\overline{T_{\frac{1}{2}}} \overline{T_{\frac{n}{2}}}} \\ &= \frac{n^n \cancel{\overline{T_{\frac{n}{2}-2}}} \left(2+\frac{1}{2}-1-1\right) \cancel{\overline{T_{\frac{5}{2}-2}}}}{\cancel{n^n} \cancel{\overline{T_{\frac{n}{2}}}} \left(2+\frac{1}{2}-1\right) \left(2+\frac{1}{2}-1-1\right)} \\ &= \frac{n^n \cdot \cancel{\frac{3}{2}} \cdot \frac{1}{2} \left(\frac{n}{2}-2\right) \cancel{\overline{T_{\frac{n}{2}-2}}}}{\cancel{(n-2)} \cancel{\frac{3}{4}} \cancel{(n-4)} - \cancel{(n-2)} \cancel{(n-4)}} = \frac{3n^n}{(n-2)(n-4)} \end{aligned}$$

$$u_1 = 0$$

$$u_2 = u_2' - (u_1')^2$$

$$u_3 = u_3' - 2u_2'u_1' + (u_1')^3$$

$$u_4 = u_4' - 4u_3'u_2' + 6u_2'u_1' - 3u_1'^4$$

for  
giv

Date: 29.04.2017

### Section A

\* Show that in estimating for the parameter "u" is the poision distribution, &

$$f(x, u) = \frac{e^{-u} u^x}{x!} ; x = 0, 1, 2, \dots$$

$\bar{x}$  is a sufficient statistic for  $u$ .

<b>likelihood function</b> : on density (joint probability function) (pdf)	$\theta$ , on $u$ $p(\theta), p(u)$
--	--

Let  $x_1, x_2, \dots, x_n$  be a random sample of size  $n$  from a population with density function  $f(x, \theta)$ . Then the likelihood function of the sample values  $x_1, x_2, \dots, x_n$  usually denoted by  $L(\theta)$ , on  $L(x_1, x_2, \dots, x_n; \theta)$  on  $(x, \theta)$  is their joint density function given by

$$L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdots f(x_n, \theta)$$

Now the likelihood function is defined as,

$$\begin{aligned} L(u) &= f(x_1, u) \cdot f(x_2, u) \cdots f(x_n, u) \\ &= \frac{e^{-u} u^{x_1}}{x_1!} \cdot \frac{e^{-u} u^{x_2}}{x_2!} \cdots \frac{e^{-u} u^{x_n}}{x_n!} \end{aligned}$$

EDUCACIONES 1967

## A synthesis

Furthermore ~~we~~ still need  $x_1 + x_2 + \dots + x_n$  and  $\beta$ .

E. ~~the~~ <sup>e!</sup> ~~in~~ <sup>Dee</sup> ~~n~~ ~~is~~ ~~missing~~ ~~int~~ ~~si~~

$$\pi(x) = \sum_{n=1}^{\infty} \frac{x_n}{n+1}$$

$\sum_{i=1}^n u_i$  ist die Summe der positiven Werte von  $u_i$ .

$$u_{10} \cdot x_1 = \frac{\prod_{i=1}^n x_i!}{\text{jumlah faktorial binomial}} \cdot \frac{\text{(jumlah faktorial triangel)}}{\text{(jumlah faktorial binomial)}}$$

Wir wollen die Wahrscheinlichkeit  $\lambda_{\text{max}}^{\text{optimal}}$  für das Maximum der Log-Likelihood-Funktion bestimmen. Dazu betrachten wir die Log-Likelihood-Funktion  $\ell(\theta, x)$  mit  $x = (x_1, x_2, \dots, x_n)$  und  $\theta = (\theta_1, \theta_2, \dots, \theta_m)$ . Die Log-Likelihood-Funktion ist definiert als

$\bar{x}$  is a sufficient statistic for  $\mu$ .

$$h(x_1, x_2, \dots, x_n) = \frac{(nx)!}{n!}$$

$$f(x) = \sum_{i=1}^n x_i$$

Let  $x_1, x_2, \dots, x_n$  be a random sample of size 'n' drawn from normal population with mean ' $\theta$ ' and variance  $\sigma^2$ . Show that,

$$t = \frac{(\sum x_i) - n\bar{x}}{\sqrt{n}} \text{ is a minimal sufficient statistic for } \theta.$$

The pdf is,

$$f(x, \theta, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\theta}{\sigma})^2} \quad -\infty < x < \infty$$

Now the likelihood function is as follows:

$$\begin{aligned} L = L(\theta, \sigma^2) &= f(x_1, \theta, \sigma^2) f(x_2, \theta, \sigma^2) \dots f(x_n, \theta, \sigma^2) \\ &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x_1-\theta}{\sigma})^2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x_2-\theta}{\sigma})^2} \dots \end{aligned}$$

$$\text{base is } (\text{m.e.})^{(n)} \text{ i.e. } e^{-\frac{1}{2}(\frac{x_n-\theta}{\sigma})^2}$$

$$\begin{aligned} \text{L.H.T.F. is } &\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x_1-\theta}{\sigma})^2} \dots \text{ division and } \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2}\sum_{i=1}^n \left(\frac{x_i-\theta}{\sigma}\right)^2} \quad \text{as } \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \text{ is constant} \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\bar{x})^2} \quad \text{as } \sigma^2 \text{ is constant} \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n \{(x_i-\bar{x})^2 + n(\bar{x}-\theta)^2\}} \end{aligned}$$

i.e. ~~(n)~~ ~~(n)~~

$$\therefore \sum (x_i - \bar{x}) = 0$$

$$\begin{aligned} &+ 2 \sum (x_i - \bar{x})(\bar{x} - \theta) \\ &= (\bar{x} - \theta) \sum (x_i - \bar{x}) \\ &+ \sum \{(x_i - \bar{x})^2 + n(\bar{x} - \theta)^2\} \\ &= \sum (x_i - \bar{x})^2 + n(\bar{x} - \theta)^2 \end{aligned}$$

to determine maximum is said max  $\{e^{\sum_{i=1}^n (x_i - \bar{x})^2 / 2\sigma^2}\}$

After multiplying  $\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2}$  with side

we get  $\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2}$  which is a minimum when  $\sum_{i=1}^n (x_i - \bar{x})^2 = 0$

but it holds  $g(\bar{x}, \theta) \ln(x_1, x_2, \dots, x_n) = 0$

where  $g(\bar{x}, \theta) = e^{-\frac{n}{2\sigma^2} (\bar{x} - \theta)^2}$

$h(x_1, x_2, \dots, x_n) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2}$

$\therefore \bar{x}$  is a sufficient statistic for  $\theta$ .

and  $T(\bar{x}, \theta) + T(\bar{x}, \theta_1) + \dots + T(\bar{x}, \theta_k) = 1$

(Unbiasedness)

An estimator  $T_n = T(x_1, x_2, \dots, x_n)$  is said to be unbiased estimator of  $T(\theta)$  if

$E(T_n) = T(\theta)$  for all  $\theta \in \mathcal{S}$

$T_n$  is unbiased estimator of  $T(\theta)$

Ex-1

Let  $x_1, x_2, \dots, x_n$  is a random sample from  $N(\mu, \sigma^2)$  show that  $t = \frac{\sum_{i=1}^n x_i^2}{n}$  is

an unbiased estimator of  $\mu^2 + \sigma^2$

an unbiased estimator of  $\mu^2$

$$\{(x-\mu)^2 + (x-\mu)^2\} = 2 +$$

$$\{(x-\mu)^2 + (x-\mu)^2\}$$

estimator of  $\mu^2 + \sigma^2$

$$0 = (x-\mu)^2$$

$$E(x) = u$$

$\therefore$  bounded  $v_n(x)$  is bounded minimum to obtain  $v$

We know  $v_n(x) = E(x^n) - \{E(x)\}$

$$E(x^{n+1}) = v(x) + \frac{\{E(x)\}}{1+u^n} = u^{n+1} \quad (1)$$

$$\begin{aligned} \text{or } E(t_n) &= \left[ \frac{\sum_{i=1}^n x_i^n}{n} \right]_0 = (2) \text{ pol } \\ &= \frac{1}{n} \sum_{i=1}^n E(x_i^n) \text{ pol } \\ &= \frac{1}{n} \sum_{i=1}^n (u^n + 1) \text{ pol } \end{aligned}$$

$\Rightarrow$  bounded sequence to obtaining  $u^n$   
 $\Rightarrow$  to statement  $\frac{1}{n}(u^n + 1)$  is also bounded  
 $= (u^n + 1) \rightarrow$  as  $n \rightarrow \infty$  is divided

$\Rightarrow$  for which will be  $(nx + ex + 1)x/b = 0$

$\therefore$   $x = 0$  with (2) is similar divided

$\Rightarrow$  to obtain bounded minimum  
 comparison with (1) is similar how

$\Rightarrow$  result with the obtained bounded

$$(2) \text{ pol}$$

$$(1) \text{ pol}$$

Date : 30.04.2017

## Principle of maximum likelihood method :

$$L(\theta) = (x_1, x_2, \dots, x_n, \theta) \quad \frac{\partial L(\theta)}{\partial \theta} = 0$$

$$\frac{\partial \log L(\theta)}{\partial \theta} = 0 \quad \frac{\partial^2 L(\theta)}{\partial \theta^2} < 0$$

$$\frac{\partial^2 \log L(\theta)}{\partial \theta^2} < 0$$

The principle of maximum likelihood is taking that value as the estimator of  $\theta$  which is maximize  $L(\theta)$ .

$\hat{\theta} = d(x_1, x_2, \dots, x_n)$  is the value of  $\theta$  which maximize  $L(\theta)$  then  $\hat{\theta}$  is the maximum likelihood estimator of  $\theta$ . We shall maximize  $L(\theta)$  The maximum likelihood estimator is the solution of the equation,

$$\frac{\partial^2 L(\theta)}{\partial \theta^2} = 0 \quad \frac{\partial^2 L(\theta)}{\partial \theta^2} < 0$$

Also  $L(\theta)$  and  $\log L(\theta)$  have maximum at the same value of  $\theta$ . Sometimes it is similar to take  $\log L(\theta)$  instead of  $L(\theta)$ .

$$\cancel{\frac{d \log L(\theta)}{d\theta} = 0} \text{ and } \frac{d^2 \log L(\theta)}{d\theta^2} < 0.$$

(Fix  $\theta$ )  $\frac{d \log L(\theta)}{d\theta} = \frac{s^2 \log L(\theta)}{s^2 \theta} > 0$

Q. What is the principle of maximum likelihood method?

Example: Let  $x_1, x_2, \dots, x_n$  be a random sample from a poision distribution with p.d.f is given by,

$$f(x, \lambda) = \frac{e^{-\lambda} \lambda^n}{x!} \quad ; \quad \lambda = 0, 1, 2, \dots, n$$

① Find the maximum likelihood estimator(MLE) of  $\lambda$

② Show that the estimator is likelihood.

$$\text{We have } f(x, \lambda) = \frac{e^{-\lambda} \lambda^n}{x!} \quad ; \quad f(x_i, \lambda) = \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$$

$$x_i = 0, 1, \dots, n$$

$$f(x, \lambda) = \frac{e^{-\lambda} \lambda^n}{x!} \quad ; \quad f(x_i, \lambda) = \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$$

Then the likelihood function  $L(\lambda) = f(x_1, \lambda) f(x_2, \lambda) \dots f(x_n, \lambda)$

$$= \frac{e^{-\lambda} \lambda^{x_1}}{x_1!} \cdot \frac{e^{-\lambda} \lambda^{x_2}}{x_2!} \dots \frac{e^{-\lambda} \lambda^{x_n}}{x_n!}$$

$$= \frac{e^{-n\lambda} \lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}$$

By simplifying on the both sides, we get

∴ Taking log on both sides, we get

$$(1) \log L(\theta) = -n\bar{x} + \sum_{i=1}^n x_i \log \theta$$

$$\therefore D> \frac{d}{d\theta} \log L(\theta) = \sum_{i=1}^n \frac{x_i}{\theta}$$

$$= -n\bar{x} + \sum_{i=1}^n x_i \log \theta - \log (\sum_{i=1}^n x_i)$$

Find maximum for  $\log L(\theta)$  w.r.t.  $\theta$  and

Now differentiating  $\log L(\theta)$  w.r.t.  $\theta$  and set equal to zero we get, boundary point

obtains  $\log L(\theta)$  and  $n\bar{x}$  ...  $n\bar{x}$  for  $\log L(\theta)$

~~Manipulating~~  $\frac{d}{d\theta} \log L(\theta) = 0$  condition for maximum

$$\Rightarrow -n\bar{x} + \sum_{i=1}^n x_i = 0 \quad (F.B.W)$$

(LM) provides boundary maximum w.r.t.  $\theta$  (I)

$$\Rightarrow -n\bar{x} + \sum_{i=1}^n x_i = 0$$

boundary in  $n$  points with boundary  $\theta$

$$\Rightarrow n\bar{x} = \sum_{i=1}^n x_i$$

$$\Rightarrow \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$$\Rightarrow \bar{x} = \bar{x}$$

Hence, the MLE of  $\theta$  is  $\bar{x}$  and it exist

$$\textcircled{1} \quad \hat{\theta} = \bar{x} \quad \therefore E(\hat{\theta}) = \gamma(\theta)$$

Taking expectation on both sides (from P)  $\Rightarrow E(\hat{\theta}) =$

$$E(\hat{\theta}) = E(\bar{x}) = E\left(\frac{\sum x_i}{n}\right)$$

$$= \frac{1}{n} \sum_{i=1}^n E(x_i) = \frac{1}{n} \sum_{i=1}^n A \quad \text{B. I. L. B. I. L. B. I. L. B. I. L. B. I. L.}$$

$$= \frac{1}{n} \cdot nA = A$$

$\therefore \hat{\theta}$  is an unbiased estimator of  $A$ .

Example: Let  $x_1, x_2, \dots, x_n$  be a random sample

drawn from the distribution having

$$\text{p.d.f is, } f(x, p) = {}^n c_x p^x q^{n-x} ; x = 0, 1, \dots, n$$

① Find the MLE of  $P$ .  $L(p) = p^n$  is likelihood.

② Show that the estimator  $\hat{p}$  is

We have  $f(x, p) = {}^n c_x p^x q^{n-x} ; x = 0, 1, \dots, n$

Now the likelihood function is,

$$\begin{aligned} L(p) &= f(x_1, p) f(x_2, p) \dots f(x_n, p) \\ &= {}^n c_{x_1} p^{x_1} q^{n-x_1} \cdot {}^n c_{x_2} p^{x_2} q^{n-x_2} \dots {}^n c_{x_n} p^{x_n} q^{n-x_n} \\ &= \prod_{i=1}^n {}^n c_{x_i} p^{\sum_{i=1}^n x_i} q^{n - \sum_{i=1}^n x_i} \end{aligned}$$

Taking log on both sides,

$$\begin{aligned} \log L(p) &= \log \left( \prod_{i=1}^n {}^n c_{x_i} \right) + \sum_{i=1}^n x_i \log p \\ &\quad + (n - \sum_{i=1}^n x_i) \log q \end{aligned}$$

$$(1) P = (q, p)$$

$$= \log(\prod_{i=1}^n x_i) + \sum_{i=1}^n x_i \log p \quad \left| \begin{array}{l} p+q=1 \\ q=p-1-p \end{array} \right. \\ + (n - \sum_{i=1}^n x_i) \cdot \log(1-p) \quad = (2)$$

Now differentiating both sides w.r.t.  $p$

$$\frac{\partial \log L(p)}{\partial p} = 0$$

$$\Rightarrow \frac{\sum_{i=1}^n x_i}{p} + \frac{n - \sum_{i=1}^n x_i}{1-p} \times (-1) = 0$$

$$\Rightarrow (1-p) \sum_{i=1}^n x_i + p(n - \sum_{i=1}^n x_i) = 0 \quad \text{[Eq. 3]}$$

$$\Rightarrow \sum_{i=1}^n x_i - p \sum_{i=1}^n x_i + p n - p \sum_{i=1}^n x_i = 0 \quad \text{[Eq. 4]}$$

$$\Rightarrow \sum_{i=1}^n x_i - n p = 0 \quad \text{[Eq. 5]} \quad \text{to sum out terms}$$

$$\Rightarrow \hat{p} = \frac{\sum_{i=1}^n x_i}{n} \quad \text{[Eq. 6] and Eq. 5]}$$

$$\text{Hence } \hat{p} = \frac{1}{n} \cdot \sum_{i=1}^n x_i \quad \text{[Eq. 6] and Eq. 5]}$$

$$\text{Hence } \hat{p} = \frac{1}{n} \cdot \sum_{i=1}^n x_i \quad \text{[Eq. 6]}$$

Hence  $\frac{\sum_{i=1}^n x_i}{n}$  is the MLE of  $\hat{p}$ .

$$\text{Prob. of } \frac{\sum_{i=1}^n x_i}{n} + \left( \frac{1}{n} \right) \text{ Prob. of } (q, 1-q)$$

$$\text{Prob. of } \left( \frac{\sum_{i=1}^n x_i}{n} + \frac{1}{n} \right)$$

⑪ Taking expectation on both sides,

$$E(\hat{P}) = E\left(\frac{\bar{x}}{n}\right) = \begin{cases} \mu = E(x) = np \\ \text{variance } v(x) = npq \end{cases}$$

$\hat{P}$  is to name  $= E\left(\frac{\sum x_i}{n}\right)$

it is the sum of individual random variable

$$\begin{aligned} &= E\left(\frac{1}{n} \sum_{i=1}^n x_i\right) \\ &= \frac{1}{n} \sum_{i=1}^n E(x_i) \\ &\stackrel{\text{sum of n r.v.}}{=} \frac{1}{n} \cdot n \cdot np \end{aligned}$$

which is the unbiased estimator of  $P$

$$(x_{\text{mean}})^{\hat{P}} = (x_{\text{mean}})$$

$$E(\hat{P}) = \frac{1}{n} \sum_{i=1}^n E(x_i)$$

$$E(x_1) + E(x_2) + \dots + E(x_n) = np$$

$$np + np + \dots + np = np$$

This is the result by first

(second)  $\hat{P}$  is the best est. of  $P$  because it got

$$(x_{\text{mean}})^{\hat{P}} = (x_{\text{mean}}) \text{ for } \hat{P} =$$

Example:  
 Let  $x_1, x_2, \dots, x_n$  be a random sample of size  $n$  from normal distribution with p.d.f is,

$$f(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2}$$

- $\infty < x < \infty$   
 $\sigma^2 > 0$   
 $-\infty < \mu < \infty$

Find the MLE of  $\mu$  and  $\sigma^2$ .

Solution: We have  $f(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2}$

Then the likelihood function is,

$$\begin{aligned} L(\mu, \sigma^2) &= \prod_{i=1}^n f(x_i, \mu, \sigma^2) \\ &= f(x_1, \mu, \sigma^2) f(x_2, \mu, \sigma^2) \dots f(x_n, \mu, \sigma^2) \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} (x_i - \mu)^2} \\ &= \left( \frac{1}{2\pi\sigma^2} \right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2} \end{aligned}$$

Taking  $\log$  on both sides.

$$\begin{aligned} \log L(\mu, \sigma^2) &= \frac{n}{2} \log \left( \frac{1}{2\pi\sigma^2} \right) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \\ &= -\frac{n}{2} \log (2\pi\sigma^2) + \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \end{aligned}$$

$$= -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) + \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

Differentiating  $\log L(\mu, \sigma^2)$  w.r.t.  $\mu$  &  $\sigma^2$   
and set to zero. equal zero. to zero.

$$\frac{\partial \log L(\mu, \sigma^2)}{\partial \mu} = 0$$

$$\Rightarrow -0 - 0 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \times (-1) = 0$$

$$\Rightarrow \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0$$

$$\Rightarrow \sum_{i=1}^n (x_i - \mu) = 0$$

$$\Rightarrow \sum_{i=1}^n x_i - n\mu = 0$$

$$\Rightarrow n\mu = \sum_{i=1}^n x_i$$

Hence the MLE of  $\mu$  is equal to the sample mean  $\bar{x}$ .

Again differentiating w.r.t.  $\sigma^2$

$$\frac{\partial \log L(\mu, \sigma^2)}{\partial \sigma^2} = 0$$

$$\Rightarrow 0 - \frac{n}{2} \frac{1}{\sigma^4} - (-1) \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2$$

$$\Rightarrow \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2 = \frac{n}{2\sigma^2}$$

$$\Rightarrow \sum_{i=1}^n (x_i - \hat{\mu})^2 = \frac{n}{\hat{\sigma}^2} \sum_{i=1}^n (x_i - \hat{\mu})^2$$

$\Rightarrow$  For  $\hat{\mu}$  to be minimum  $\frac{\sum (x_i - \hat{\mu})^2}{n}$  must be minimum which is sum of squared errors.

Hence the MLE of  $\hat{\mu}$  is  $\frac{\sum (x_i - \hat{\mu})^2}{n}$

④ Binomial, poisson, normal.

যেখন গোটো কোর্স

$$\hat{\sigma}^2 = \frac{(n-1)s^2}{n}$$

Properties of the MLE:

- ① MLE are generally unbiased
- ② MLE are consistent
- ③ MLE are most efficient
- ④ MLE are sufficient
- ⑤ MLE is unique under regularity condition
- ⑥ MLE have the invariance property.

Efficiency | most efficient  
with the estimator.

$$E(\text{efficiency}) = \frac{V(T_1)}{V(T_2)} = \frac{1}{1} (\max)$$

most of time

$$E < 1$$

### Consistency estimator.

$$T_n \xrightarrow[n \rightarrow \infty]{P} \theta = \infty$$

$$T_n = T(x_1, x_2, \dots, x_n)$$

### Consistency:

An estimator  $T_n = T(x_1, x_2, \dots, x_n)$  based on a random sample of size  $n$  is said to be consistent estimator of  $\gamma(\theta)$ ;  $\theta \in \Omega$  the parameter space if  $T_n$  converges to  $\gamma(\theta)$  in probability.

$$\text{i.e } T_n \xrightarrow{P} \gamma(\theta) \text{ as } n \rightarrow \infty.$$

### Efficiency:

If  $T_1$  is the most efficient estimator with variance  $v(T_1)$  and  $T_2$  is any other estimator with variance  $= v(T_2)$  then the efficiency 'E' of  $T_2$  is defined as

$$E = \frac{v(T_1)}{v(T_2)}$$

obviously 'E' can not exceed unity.

What is the difference between  $t$  and  $t'$  in terms of  $\theta$ ?

### Point Estimator

Let  $x_1, x_2, \dots, x_n$  is a sample from a density function  $f(x, \theta)$  or  $f(x/\theta)$  where ' $\theta$ ' is unknown fixed value for which we can assume any value in one-dimensional real parameter space  $\Omega$ .

Let ' $t$ ' be a function of  $x_1, x_2, \dots, x_n$ . So that ' $t$ ' is a statistic and hence a random variable. If it is used to estimate ' $\theta$ ', then ' $t$ ' is called a point estimate of ' $\theta$ '.

$$f(x, \theta) = \theta^x (1-\theta)^{x-1}$$

$$t = f(x_1, x_2, \dots, x_n)$$

' $t$ ' is a sufficient statistic.

Let  $f(x, \theta)$  be a density of a random variable  $(x)$  where ' $\theta$ ' is unknown fixed parameter and  $\theta \in \Omega$ . Let  $x_1, x_2, \dots, x_n$  be a random sample from this distribution/density function. Let ' $t$ ' and ' $t'$  are two statistics such that  $t'$  is not a function of ' $t$ '. If the conditional distribution of ' $t'$  for given ' $t$ ' be independent

of ' $\theta$ '. Then ' $t$ ' is called a sufficient statistic for ' $\theta$ '.

Let the joint distribution of  $t$  and  $t'$  be  $h(t, t'/\theta)$  as follows

$$h(t, t'/\theta) = g(t, \theta) \cdot h(t'/t)$$

$$\text{or } h(t'/t) = \frac{h(t, t'/\theta)}{g(t, \theta)}$$

Where  $g(t, \theta)$  is the marginal p.d.f. of  $t$  for fixed ' $\theta$ ', and  $h(t'/t)$  is the conditional distribution of  $t'$  given  $t$ .

A statistic  $t_n = t(x_1, x_2, \dots, x_n)$  is a sufficient estimator of parameter ' $\theta$ ' if and only if the likelihood function (joint p.d.f. of the sample) can be expressed as,

$$L(\theta) = \prod_{i=1}^n f(x_i, \theta)$$

$$= g(t_n, \theta) h(x_1, x_2, \dots, x_n)$$

where  $g(t_n, \theta)$  is the p.d.f. of the statistic  $t_n$  and  $h(x_1, x_2, \dots, x_n)$  is the function of sample observation only.

Example: Let  $x_1$  and  $x_2$  be independent distributed with  $f(x_i; \theta) = \theta x_i^{\theta-1}$ ,  $x_i > 0$ ,  $0 < \theta < 1$ . Show that  $\prod_{i=1}^2 x_i$  is a sufficient statistic for  $\theta$ .

Now the likelihood function

$$L = \prod_{i=1}^2 f(x_i; \theta) = \theta^2 x_1^{\theta-1} x_2^{\theta-2}$$

$$= \theta^2 \left( \prod_{i=1}^2 x_i \right)^{\theta-1} (\theta-1)^{x-2}$$

$$= \theta^2 \left( \prod_{i=1}^2 x_i \right)^{\theta-1} (\theta-1)^{x-2} \cdot \frac{(x-1)(x-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)}$$

$$= \theta^2 \left( \prod_{i=1}^2 x_i \right)^{\theta-1} \frac{x(x-1)}{2!} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)}$$

$$= \theta^2 \left( \prod_{i=1}^2 x_i \right)^{\theta-1} \frac{x(x-1)}{2!} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)}$$

$$= g\left(\prod_{i=1}^2 x_i, \theta\right) \cdot h(x_1, x_2, \dots, x_n)$$

$$\text{where } g\left(\prod_{i=1}^2 x_i, \theta\right) = \theta^2 \left( \prod_{i=1}^2 x_i \right)^{\theta-1} \cdot \frac{(\theta-1)(\theta-2)}{2!} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)} \cdot \frac{(\theta-1)(\theta-2)}{(\theta-1)(\theta-2)}$$

$\therefore \prod_{i=1}^2 x_i$  is a sufficient statistic for  $\theta$ .

Example: Let  $x_1, x_2, \dots, x_n$  be a random sample from the density  $f(x_i, \theta) = \theta x_i (1-\theta)^{1-x_i}$  where  $0 < \theta < 1$ .

Let  $\sum_{i=1}^n x_i$  be a statistic.

Then show that  $\sum_{i=1}^n x_i$  is a sufficient statistic for  $\theta$ . In other words,  $\sum_{i=1}^n x_i$  is a function of the sufficient statistic.

Now the likelihood function or joint density function is,

$$L(\theta) = \prod_{i=1}^n f(x_i, \theta) = \prod_{i=1}^n \theta^{x_i} (1-\theta)^{1-x_i} = \prod_{i=1}^n \theta^{\sum_{j=1}^n x_j - i} (1-\theta)^{n-i}$$

$$= \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n - \sum_{i=1}^n x_i}$$

$$= g\left(\sum_{i=1}^n x_i, \theta\right) \cdot h(x_1, x_2, \dots, x_n)$$

$$\text{where } g\left(\sum_{i=1}^n x_i, \theta\right) = \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n - \sum_{i=1}^n x_i}$$

$$\text{and } h(x_1, x_2, \dots, x_n) = 1$$

$\sum_{i=1}^n x_i$  is a sufficient statistic for  $\theta$

the  $\sum_{i=1}^n x_i$  is fixed and  $n - \sum_{i=1}^n x_i$  is random

$$L = \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n - \sum_{i=1}^n x_i}$$

Hence the marginal distribution of  $\sum_{i=1}^n x_i$  is,

$$g\left(\sum_{i=1}^n x_i, \theta\right) = \binom{n}{\sum_{i=1}^n x_i} \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n - \sum_{i=1}^n x_i}$$

$$f(x, p) = n c x^x q^{n-x}$$

$$\begin{aligned} h(t'/t) &= \frac{g(t'/\theta)}{g(t/\theta)} \\ &= \frac{\theta^{\sum x_i} (1-\theta)^{n - \sum x_i}}{\binom{n}{\sum x_i} \theta^{\sum x_i} (1-\theta)^{n - \sum x_i}} \\ &= \frac{1}{n c \sum x_i} \end{aligned}$$

mitobrings to sufficient if

different if

3rd tutorial Section B is on April 24,

2nd tutorial will be held on April 24, 2017 at 8:15 am.

Test about correlation coefficient,

# Correlation Analysis:

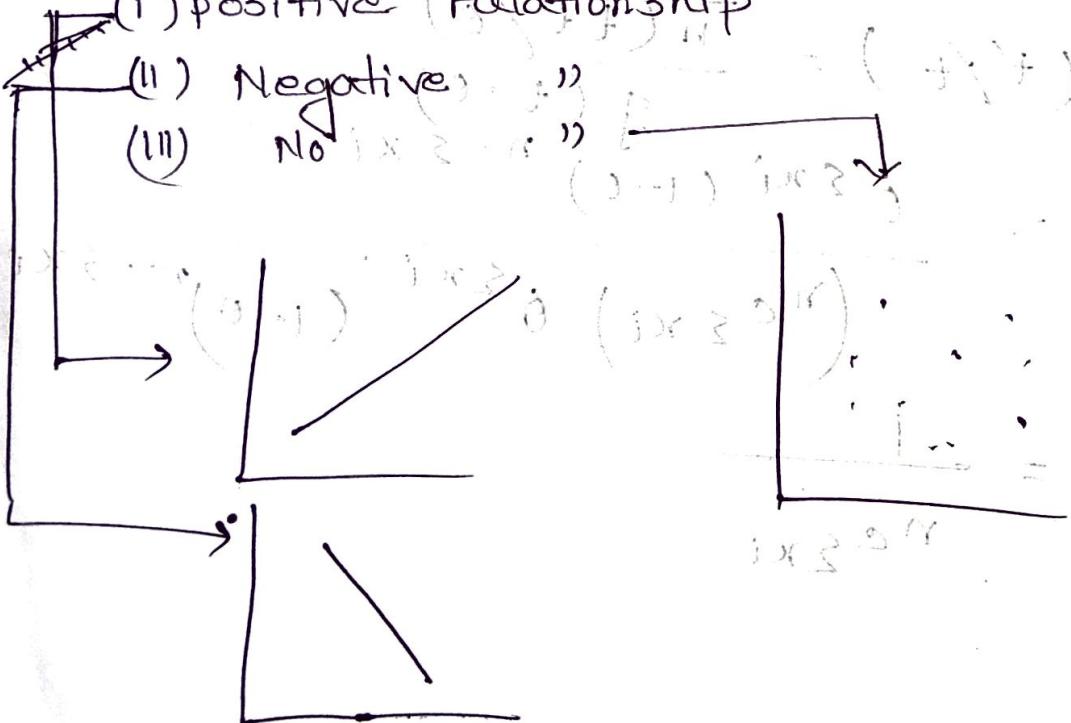
$$y = \alpha + \beta x$$

# scatter plots:

(I) positive relationship

(II) Negative

(III) No relationship



# properties of correlation:

# Limitation.

- $\Rightarrow r=0$   $\rightarrow$  no relationship, i.e. no dependency  
 $\Rightarrow 0 < r \leq 0.25 \rightarrow$  weak positive  
 $\Rightarrow 0.25 < r \leq 0.7 \rightarrow$  intermediate  
 $\Rightarrow 0.75 < r \leq 1 \rightarrow$  strong positive  
 $\Rightarrow r=1 = \text{perfect}$

Opposite to test if

# Example:

### 9.7 Test about correlation coefficients.

$$t = \frac{n \sqrt{n-2}}{\sqrt{1-r^2}}$$

$$df = 4$$

$$\alpha = 0.05$$

- ⊗ test procedure
- ⊗ type-I and Type-II error (difference)
- ⊗ def" ( a hypothetical test, ... )
- ⊗ Best critical region
- ⊗ math correlation coefficient on test

Problem: The coefficient of correlation obtained from a random sample of 28 pairs of values is 0.45. Is this significant at 5% level?

# Test of significance?

\* rank & run test

$$P = 46$$