02/08/2017

1. Classical Definition

2. Axioms

3. Proof of classical definition

4. Frequency definition

Classical definition of probability

Let E be an experiment such that its event space and contains n no. of event points which are equally likely or mutually symmetrical. If A be an event connected with E containing m no. of event points, then probability of A is denoted by P(A) and defined by $P(A) = \frac{m}{n}$

Axioms

(i) $0 \le P(A) \le 1$

Probability of an impossible event is 0 and probability of certain event is 1

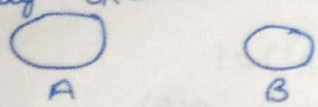
(ii) If $A_1, A_2, \ldots A_n$ are mutually exclusive events

$(A_i \cap A_j = \phi,\ i \ne j,\ i, j = 1, 2, \ldots n)$

then $P(A_1 \cup A_2 \cup \ldots A_n) = P(A_1) + P(A_2) \ldots + P(A_n)$

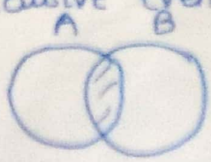
or $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$

If A and B are mutually exclusive events



$P(A \cup B) = P(A) + P(B)$

If A and B are not mutually exclusive events



$A - B = A \cap B^c$

$A = \underbrace{(A - A\cap B)}_{C} \cup (A \cap B)$ — (i)

$P(A) = P(A - A\cap B) + P(A\cap B)$ (mutually exclusive)

$B = (A\cap B) \cup \underbrace{(B - A\cap B)}_{D}$

$P(B) = P(A\cap B) + P(B - A\cap B)$ — (ii)

$A \cup B = (A - A\cap B) \cup (A\cap B) \cup (B - A\cap B)$

$P(A\cup B) = P(A - A\cap B) + P(A\cap B) + P(B - A\cap B)$

$= P(A) - \cancel{P(A\cap B)} + \cancel{P(A\cap B)} + P(B) - P(A\cap B)$

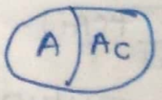
$$\boxed{P(A\cup B) = P(A) + P(B) - P(A\cap B)}$$

NOTE: If A and B are mutually exclusive, then

$A \cap B = \phi$ $\quad P(A\cap B) = 0$

$\therefore P(A\cup B) = P(A) + P(B)$

- $A \cup A^c = S$
- $A \cap A^c = \phi$



$P(A\cup A^c) = P(S) = 1$

$P(A) + P(A^c) = 1$

$P(A^c) = 1 - P(A)$

Scanned by CamScanner

AND — ×

OR — +

$P(A)P(\bar{B})P(\bar{C}) + P(\bar{A})P(B)P(\bar{C}) + P(\bar{A})P(\bar{B})P(C)$

$+ P(A).P(B).P(\bar{C}) \ldots$

$+ P(A).P(B).P(C) = 1 - P(\bar{A})P(\bar{B})P(\bar{C})$

$P(\text{getting at least one six}) = 1 - \left(\frac{5}{6}\right)^{10}$

Proof: of classical Definition

Let $A_1, A_2, \ldots A_n$ be mutually exclusive events

$S = A_1 \cup A_2 \cup \ldots \cup A_n$

$P(A_1) = P(A_2) \ldots \ldots P(A_n) = \frac{1}{n}$

$A = A_1 \cup A_2 \ldots \cup A_m$

$P(A) = P(A_1) + P(A_2) \ldots + P(A_m)$

$= \frac{m}{n}$

Frequency ~~distribution~~ definition

Def: Let an experiment E be repeated N no. of times under uniform condition. If A be an event connected with E is found to occur N(A) no. of times, then N(A) → absolute frequency

$\frac{N(A)}{N} \rightarrow$ frequency ratio or relative frequency

$\lim_{N \to \infty} \frac{N(A)}{N}$ approaches to a constant

This phenomenon of stability for long sequence of repetition is called statistical regularity or probability

$f(A) = \frac{N(A)}{N}$

$\lim_{N \to \infty} f(A) = P(A)$

03/08/2017

1	5	
5	1	
2	4	$\frac{5}{36}$
4	2	
3	3	

1	6	
6	1	
2	5	$\frac{6}{36}$
5	2	
3	4	
4	3	

A and B play with two dice

A	B
6	7

wins

If A starts then probability of A's winning

$= P(A) + P(\bar{A})P(\bar{B})P(A) + \cdots$

B's $= P(\bar{A})P(B) + P(\bar{A})P(\bar{B})P(\bar{A})P(B)$

markovian/Non-markovian

Independent events

Two events A and B are said to be independent or stochastically independent if $P(A \cap B) = P(A).P(B)$

completely dependent on time

e.g. If A and B are independent events then show that (i) A and B^c are independent (ii) A^c and B^c are indepen.

solution: $\because$ A and B are independent events

$P(A \cap B) = P(A).P(B)$

(i) $B \cup B^c = S$, $B \cap B^c = \phi$

$A \cap (B \cup B^c) = A \cap S$

$(A \cap B) \cup (A \cap B^c) = A$

$[(A \cap B) \cap (A \cap B^c)$

$P(A \cap B) + P(A \cap B^c) = P(A)$

$= A \cap (B \cap B^c)$

$= A \cap \phi$

$= \phi]$

$P(A \cap B^c) = P(A) - P(A \cap B)$

$= P(A) - P(A).P(B)$

$P(A \cap B^c) = P(A)(1 - P(B))$

$P(A \cap B^c) = P(A) P(B^c)$

De-morgan's law $(A \cup B)^c = A^c \cap B^c$

$P(A^c \cap B^c) = P((A \cup B)^c)$

$= 1 - P(A \cup B)$

$= 1 - [P(A)] - P(B) + P(A).P(B)$

$= (1 - P(A))(1 - P(B))$

$= P(A^c) P(B^c)$

Conditional Probability

The conditional probability of an event B on the hypothesis that another event 'A' has occurred is denoted by $P(B/A)$ and denoted by

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \quad (\because P(A) \neq 0)$$

Scanned by CamScanner

Baye's Theorem

- If $A_1, A_2, \ldots, A_n$ are mutually exclusive events $(A_i \cap A_j = \phi)$, $i \neq j$, $i, j = 1, 2, 3, \ldots, n)$ and $S = A_1 \cup A_2 \cup \ldots A_n$

then for any arbitary event X,

$$P(X) = P(A_1)P\left(\frac{X}{A_1}\right) + P(A_2)P\left(\frac{X}{A_2}\right) + P(A_3)P\left(\frac{X}{A_3}\right) + \ldots + P(A_n)P\left(\frac{X}{A_n}\right)$$

and for $P(X) \neq 0$

$$P\left(\frac{A_i}{X}\right) = \frac{P(A_i)P\left(\frac{X}{A_i}\right)}{P(X)}$$

Proof :

$$S = A_1 \cup A_2 \ldots \cup A_n$$

$$S \cap X = (A_1 \cup A_2 \ldots \cup A_n) \cap X$$

$$X = (A_1 \cap X) \cup (A_2 \cap X) \cup (A_3 \cap X) \ldots (A_n \cap X)$$

$$= (A_i \cap X) \cup (A_j \cap X) \qquad i \neq j \qquad (A_i \cap X) \cap (A_j \cap X)$$

$$= (A_i \cap A_j) \cup X$$

$$X = \phi \cup X = \phi X$$

$$\left[\cancel{(A_i \cap X)} \cap (A_j \cap X) \right.$$
$$[\cancel{(A_i \cap X)} \cap A_j]$$
$$\left. \cap [(A_i \cap X) \cap \cancel{X}] \right]$$
$$\underbrace{(A_i \cap A_j)}_{\phi} \cap (X \cap A_j) \cap (A_i \cap X) = \phi$$

$$(A_i \cap A_j) \cap \ldots$$
$$\phi \cap X$$
$$= \phi$$

$$P(X) = P(A_1 \cap X) + P(A_2 \cap X) \ldots + P(A_n \cap X)$$

$$[P(A_i \cap X) = P(A_i) \cdot P(X/A_i)]$$

$$= P(A_1)P\left(\frac{X}{A_1}\right) + \ldots + P(A_n)P\left(\frac{X}{A_n}\right)$$

a.

A_1	A_2	A_3
2W	5W	3W
3B	2B	7B

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}$$

Let X be the event of drawing white ball.

$$P(X) = P(A_1)P\left(\frac{X}{A_1}\right) + P(A_2)P\left(\frac{X}{A_2}\right) + P(A_3)P\left(\frac{X}{A_3}\right)$$

$$= \frac{1}{3} \times \frac{2}{5} + \frac{1}{3} \times \frac{5}{7} + \frac{1}{3} \times \frac{3}{10}$$

$$= \frac{1}{3}\left(\frac{2}{5} + \frac{5}{7} + \frac{3}{10}\right]$$

$$= \frac{1}{3} \times$$

$$P\left(\frac{A_1}{X}\right) = \frac{P(A_1)\,P(X/A_1)}{P(X)}$$

$\vdots$

$$P\left(\frac{A_1}{X}\right) + P\left(\frac{A_2}{X}\right) + P\left(\frac{A_3}{X}\right) = 1$$

Random Variable

08/08/2017

For each value of a set U of an event space S if we get unique real value $X = X(U)$, then X is called random variable or variate.

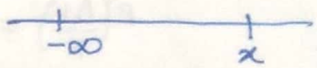
Set of X – spectrum (Range)

If spectrum is continuous, random variable is continuous.

Distribution Function

The distribution function of a random variable X, is a function of a real variable is denoted by $F(x)$ and defined by

$$F(x) = P(-\infty < X \leq x)$$



$$P(a < X \leq b) = P(-\infty < X \leq b) - P(-\infty < X \leq a)$$

$$= F(b) - F(a)$$

$$= \int_a^b F'(x)\,dx$$

$$= \int_a^b f(x)\,dx,$$

where $f(x) = F'(x)$ is called Probability Density Function.

$$\int_{-1}^{1} \frac{dx}{x}$$

$$= \lim_{\varepsilon \to 0} \int_{-1}^{0-\varepsilon} \frac{dx}{x} + \lim_{\delta \to 0} \int_{0+\delta}^{1} \frac{dx}{x}$$

$$= \lim \ln\left|\frac{\varepsilon}{\delta}\right|$$

does not exist

If $\varepsilon = \delta$

$$\int_{-1}^{1} \frac{dx}{x} = \ln 1 = 0$$

Principal value.

$$P(a < X < b) = \int_a^b f(x).dx$$

Making $a \to -\infty$, $b = x$

$$P(-\infty < X < x) = \int_{-\infty}^{x} f(x).dx$$

$$F(x) = \int_{-\infty}^{x} f(x).dx$$

Making $a \to -\infty$, $b \to +\infty$

$$P(-\infty < X < \infty) = \int_{-\infty}^{\infty} f(x).dx = 1$$

$$\boxed{\int_{-\infty}^{\infty} f(x).dx = 1}$$

If discrete

$$\sum_{i=-\infty}^{+\infty} f_i = 1$$

Ex1 Find the value of k so that $f(x)$ ~~$f(x) = kx(1-x)$~~ is a probability density function :

$$f(x) = \begin{cases} kx(1-x) & 0 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Hence find the corresponding distribution function.

Sol^n: $\because f(x)$ is a probability density function $\int_{-\infty}^{\infty} f(x).dx = 1$

$$\int_0^1 kx(1-x)dx = 1$$

$$k\int_0^1 (x - x^2)\,dx = 1$$

$$k\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = 1$$

$$k \times \frac{1}{6} = 1$$

$$\boxed{k = 6}$$

$$f(x) = \begin{cases} 6x(1-x), & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

If $\quad -\infty < x < 0,$

$$F(x) = \int_{-\infty}^{x} f(x).dx$$

$$= \int_{-\infty}^{x} 0.dx = 0$$

If $\quad 0 < x < 1,$

$$F(x) = \int_{-\infty}^{x} f(x).dx = \int_{-\infty}^{0} f(x).dx + \int_{0}^{x} f(x).dx$$

$$F(x) = 0 + \int_{0}^{x} f(x).dx$$

$$= \int_{0}^{x} 6x(1-x).dx$$

$$= \left(\frac{6x^2}{2} - \frac{6x^3}{3}\right)_0^x$$

$$F(x) = (3x^2 - 2x^3)$$

If $\quad 1 < x < \infty$

$$F(x) = \int_{-\infty}^{x} f(x).dx = \cancel{\int_{-\infty}^{0} f(x).dx} + \int_{0}^{1} f(x).dx + \cancel{\int_{1}^{x} f(x).dx}$$

$$= \int_{0}^{1} (3x^2 - 2x^3)\,dx$$

$$= \left(\frac{3x^3}{3} - \frac{2x^4}{4}\right)_0^1$$

$$= \int_{0}^{1} 6x(1-x).dx$$

$$= (3x^2 - 2x^3)_0^1$$

$$= 1$$

$$F(x) = 0 \qquad -\infty < x < 0$$
$$= 3x^2 - 2x^3 \qquad 0 < x < 1$$
$$= 1 \qquad x > 1$$

NOTE:- $F(-\infty) = 0$

$F(\infty) = 1$

Q. Find the value of K so that

$$f(x) = x \qquad 0 < x < 1$$
$$= K - x \qquad 1 < x < 2$$
$$= 0 \quad , \quad \text{elsewhere}$$

is a p.d.f. Hence find the corresponding distribution function

Also find $P\left(\frac{1}{2} < x < \frac{3}{2}\right)$

Sol^n

$$\int_{-\infty}^{\infty} f(x).dx = 1$$

$$\int_{-\infty}^{0} 0.dx + \int_{0}^{1} x.dx + \int_{1}^{2} (K-x)dx + \int_{2}^{\infty} 0.dx = 1$$

$$\frac{1}{2} + \left(Kx - \frac{x^2}{2}\right)_1^2 = 1$$

$$\frac{1}{2} + (2K - 2) - \left(K - \frac{1}{2}\right) = 1$$

$$1 + K - 2 = 1$$

$$\boxed{K = 2}$$

If $-\infty < x < 0$
$F(x) = 0$

If $0 < x < 1$

$$F(x) = \int_{-\infty}^{x} f(x).dx = \int_{-\infty}^{0} f(x) + \int_{0}^{x} f(x).dx$$

$$= 0 + \int_{0}^{x} x.dx$$

$$F(x) = \frac{x^2}{2}$$

Scanned by CamScanner

If $1 < x < 2$

$$F(x) = \int_{-\infty}^{x} f(x).dx = \int_{-\infty}^{0} f(x).dx + \int_{0}^{1} f(x).dx + \int_{1}^{x} f(x).dx$$

$$= 0 + \int_{0}^{1} x.dx + \int_{1}^{x} (2-x).dx$$

$$= \frac{1}{2} + \left(2x - \frac{x^2}{2}\right)_{1}^{x}$$

$$= \frac{1}{2} + \left(2x - \frac{x^2}{2}\right) - \left(2 - \frac{1}{2}\right)$$

$$= -1 + 2x - \frac{x^2}{2}$$

If $x > 2$

$$F(x) = \int_{-\infty}^{x} f(x).dx = \int_{-\infty}^{0} f(x).dx + \int_{0}^{1} f(x).dx + \int_{1}^{2} f(x).dx + \int_{2}^{\infty} f(x).dx$$

$$= 1$$

$$P\left(\frac{1}{2} < x < \frac{3}{2}\right) = F\left(\frac{3}{2}\right) - F\left(\frac{1}{2}\right)$$

$$= -1 + 2 \times \frac{3}{2} - \frac{9}{4 \times 2} - \left[\frac{1}{4 \times 2}\right]$$

$$= 2 - \frac{9}{8} - \frac{1}{8}$$

$$= 2 - \frac{\cancel{10}\,5}{\cancel{8}\,4}$$

$$= \frac{3}{4}$$

MT Thursday

10/08/2017

(n)

sfff ssf . . . .

$\underbrace{s.s.s.s}_{r}\,\underbrace{f.f.\dots f}_{n-r}$

Probability of success does not change with trial

$P(S) = p$

$P(f) = 1-p = q$

$\underbrace{p.p.\dots p}_{r}\,\underbrace{.q.\dots q}_{n-r}$

$P(A_r) = {}^nC_r\, p^r q^{n-r}$ Bernoulli trials $r = 0, 1, 2, \dots n$

↳ Binomial law

Let $p = \frac{\mu}{n}$ where μ is moderate value.

As $n \to \infty$ $p \to 0$

$$P(A_r) = \frac{n!}{r!(n-r)!} p^r (1-p)^{n-r}$$

$$= \frac{n.(n-1)\dots(n-r+1)}{r!} \left(\frac{\mu}{n}\right)^r \left(1-\frac{\mu}{n}\right)^{n-r}$$

$$= \frac{\cancel{n^r}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\left(1-\left(\frac{r-1}{n}\right)\right)}{r!} \frac{\mu^r}{\cancel{n^r}} \frac{\left(1-\frac{\mu}{n}\right)^n}{\left(1-\frac{\mu}{n}\right)^r}$$

as $n \to \infty$

$$= \frac{1}{r!} \mu^r e^{-\mu}$$

$r = 0, 1, 2, \dots n$

$$\boxed{P(A_r) = \frac{e^{-\mu}\mu^r}{r!}}$$ → Poisson process

$$\sum_{r=-\infty}^{\infty} f_r = \sum_{r=0}^{n} {}^nC_r p^r q^{n-r} = (p+q)^n = 1$$

$$\sum_{r=-\infty}^{\infty} f_r = \sum_{r=0}^{\infty} \frac{e^{-\mu}\mu^r}{r!} = e^{-\mu}\sum_{r=0}^{\infty}\frac{\mu^r}{r!} = e^{-\mu}.e^{\mu} = 1$$

NOTE : Poisson process is a limiting case of Bernoulli

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

Normal distribution function

$$\int_{-\infty}^{\infty} f(x).dx = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

$$\frac{x-m}{\sqrt{2}\sigma} = z$$

$$dx = \sqrt{2}\sigma dz$$

$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-z^2} dz$$

$$= \frac{1}{\sqrt{\pi}} \times \sqrt{\pi}$$

$$= 1$$

$$\int_{-\infty}^{\infty} e^{-x^2}.dx = \sqrt{\pi}$$

$$\int_{0}^{\infty} e^{-x^2}.dx = \frac{\sqrt{\pi}}{2}$$

Symm.

Symm.

with

P(Ai

P(Ai

symm. about line is even

symm. w.r.t point is odd

~~with~~ with replacement

N_1 white

N_2 black

$N = N_1 + N_2$

$$P(A_i) = {}^nC_i \cdot p^i q^{n-i}$$

$$P(A_i) = {}^nC_i \left(\frac{N_1}{N}\right)^i \left(\frac{N_2}{N}\right)^{n-i}$$

without replacement.

$$\frac{{}^{N_1}C_i \; {}^{N_2}C_{n-i}}{{}^NC_n} \xrightarrow{} \text{finite}$$

$N \to \infty \Rightarrow N_1, N_2 \to \infty$ [illegible] p is fixed

$N \longrightarrow \infty \longrightarrow$ p is finite

$$\frac{\dfrac{N_1(N_1-1)\ldots(N_i-i+1)}{i!} \; \dfrac{N_2(N_2-1)\ldots(N_2-n+i+1)}{(n-i)!}}{} \times \frac{n!}{N(N-1)(N-2)\ldots(N-n+1)}$$

$$\frac{N_1^i \left[1\cdot\left(1-\frac{1}{N_1}\right)\ldots\left(1-\frac{(i-1)}{N_1}\right)\right]}{i!} \; \frac{N_2^{n-i}\left(1\cdot\left(1-\frac{1}{N_2}\right)\ldots\left(1-\frac{(n-i-1)}{N_2}\right)\right)}{(n-i)!} \times \frac{n!}{N^n\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\ldots\left(1-\left(\frac{n-1}{N}\right)\right)}$$

as $N \longrightarrow \infty$

$$\frac{{}^nC_i \, N_1^i N_2^{n-i}}{N^n}$$

$$= {}^nC_i \left(\frac{N_1}{N}\right)^i \left(\frac{N_2}{N}\right)^{n-i}$$

Limiting case of with~~out~~ replacement is without replacement.

17/08/2017

GAMMA FUNCTION

(Given by Euler)

$\Gamma \rightarrow$ Mathematical convergence.

$$\Gamma(n) = \int_0^\infty e^{-x} x^{n-1}\, dx \qquad (n>0)$$

$$\Gamma(n+1) = \int_0^\infty e^{-x} x^{n} . dx$$

$$= \left[-x^n e^{-x}\right]_0^\infty + n\int_0^\infty e^{-x} x^{n-1} . dx$$

$$\lim_{n\to\infty} \frac{x^n}{e^x} = 0 \qquad = \qquad 0 + n\Gamma(n)$$

$$\Gamma(n+1) = n\,\Gamma(n-1+1)$$

$$= n(n-1) \ldots \Gamma(1)$$

$$\boxed{\Gamma(n+1) = n! . 1}$$

$$\beta(m,n) = \int_0^1 x^{m-1} (1-x)^{n-1}\, dx \qquad m,n>0$$

$$x = \sin^2\theta$$

$$= 2\int_0^{\pi/2} \sin^{2m-1}\theta \cos^{2n-1}\theta . d\theta$$

$m = n = 1 \quad \beta(1,1) = 1$

$m = n = \frac{1}{2} \quad \beta(\frac{1}{2},\frac{1}{2}) = \pi$

Relation between β and Γ function

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

Put

$m=n=\frac{1}{2}$

$$\beta\left(\tfrac{1}{2},\tfrac{1}{2}\right) = \pi$$

$$\pi = \frac{\left[\Gamma\left(\tfrac{1}{2}\right)\right]^2}{\Gamma(1)}$$

$$\Gamma\left(\tfrac{1}{2}\right) = \sqrt{\pi}$$

$$\int_0^\infty e^{-\alpha x} x^n \cdot dx$$

$\alpha x = z$

$\alpha dx = dz$

$$\int_0^\infty e^{-z}\left(\frac{z}{\alpha}\right)^n \frac{dz}{\alpha}$$

$$\frac{1}{\alpha^{n+1}} \int_0^\infty e^{-z} z^{(n+1)-1} dz$$

$$= \frac{\Gamma(n+1)}{\alpha^{n+1}}$$

$$\int_0^\infty e^{-x^2} \cdot dx$$

$x^2 = z$

$dz = 2x \cdot dx$

$$\int_0^\infty e^{-z} \frac{dz}{2\sqrt{z}}$$

$$\frac{1}{2}\int_0^\infty e^{-z} z^{-1/2} dz \qquad \frac{1}{2}\int_0^\infty e^{-z} z^{\frac{1}{2}-1} dz$$

$$\frac{1}{2}\Gamma\left(\tfrac{1}{2}\right) = \frac{\sqrt{\pi}}{2}$$

Gamma Distribution

$$f(x) = \begin{cases} \dfrac{e^{-x} x^{n-1}}{\Gamma(n)} , & 0 < x < \infty \\ = 0 , & \text{elsewhere} \end{cases}$$

Beta Distribution $\beta(m,n) = \beta(n,m)$

Two Beta distributions.

Beta distribution of first kind

$$f(x) = \begin{cases} \dfrac{x^{m-1}(1-x)^{n-1}}{\beta(m,n)} , & 0 < x < 1 \\ 0 , & \text{elsewhere} \end{cases}$$

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

$$1+x = \frac{1}{z}$$

$$dx = \frac{-1}{z^2} dz$$

$$\int \left(\frac{1}{z}-1\right)^{m-1} \left(-\frac{dz}{z^2}\right) \cdot$$

$$-\int_1^0 \frac{dz}{z^2} \frac{(1-z)^{m-1}}{z^{m-1}}$$

$$\int_0^1 \frac{dz\,(1-z)^{m-1}}{z^{m+1}}$$

$$\int_0^1 z^{-m-1} (1-z)^{m-1} dz$$

$$\int_0^1 z^{-m+1} (1-z)^{m-1} dz$$

$$-\int_1^0 z^{m+n} \left(\frac{1}{z}-1\right)^{m-1} \frac{dz}{z^2}$$

$$\int_0^1 \frac{z^{m+n} (1-z)^{m-1}}{z^{m+1}} dz$$

$$\int_0^1 z^{n-1} (1-z)^{m-1} dz$$

$$= \beta(n,m) = \beta(m,n)$$

$\Gamma(-m)\,\Gamma(m)$

Beta distribution of second kind

NOTE :

another one

Beta distribution of second kind

$$f(x) = \frac{x^{m-1}}{(1+x)^{m+n}\beta(m,n)}, \quad 0<x<\infty$$

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

$$\int_0^\infty x^2 e^{-x^2} dx$$

$$= \frac{1}{2}\int_0^\infty z e^{-z} z^{-1/2} dz$$

$$= \frac{1}{2}\int_0^\infty e^{-z} z^{3/2-1} dz$$

$$= \frac{1}{2}\Gamma\left(\frac{3}{2}\right)$$

$$= \frac{1}{2}\times\frac{\sqrt{\pi}}{2}$$

NOTE: $\Gamma(n+1) = n\,\Gamma(n)$

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

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

$$-2\sqrt{\pi} = \Gamma\left(\frac{-1}{2}\right)$$

zero and negative integer not defi[ned] -e

$\frac{1}{\Gamma(0)}, \frac{1}{\Gamma(-1)}, \frac{1}{\Gamma(-2)}$ is defined.

mean:

The k-th ~~abo~~ moment about the origin is denoted by

~~k^{th} central moment~~

$$\alpha_k = E(X^k) = \int_{-\infty}^{\infty} x^k f(x).dx$$

$$= \sum_{i=-\infty}^{\infty} i^k f_i$$

↳ probability density function

$m = E(X) \longrightarrow$ Mean

i.e. The first moment about the origin

Variance:

The k^{th} central moment about the mean is denoted by μ_k

and defined by

$$\mu_k = E(X-m)^k$$

μ_2 – The second moment about the mean is called the variance.

$$\text{Variance} = \mu_2 = E(X-m)^2$$

$$= E(X^2 + m^2 - 2mX)$$

$$= E(X^2) + m^2 - 2m\underbrace{E(X)}_{m}$$

$$= E(X^2) - m^2$$

$$= E(X^2) - (E(X))^2$$

$$= \alpha_2 - \alpha_1^2$$

$$E(X-m) - $$
$$= E(X) - m$$
$$= m - m$$
$$= 0$$

Mean:

Physical significance:

The physical significance of mean is, it represents the centre of mass of Probability mass distribution.

This gives the rough position of the bulk of distribution.

Thus it is called measure of location.

Thus mean is not the only measure of location. There are other physical quantities for memory location.

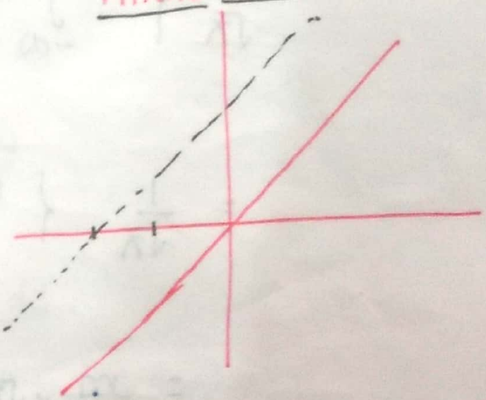
Variance

The variance represents how widely the probability masses are distributed about the mean. Thus it is called the measure of dispersion. The physical interpretation of the variance is, it represents the moment of inertia of the line of distribution about a line through the mean $\perp^r$ to the line

This is not only the measure of dispersion

Standard deviation $= +\sqrt{\text{Variance}}$

Affine function



$\alpha_1 = m$

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

$$\alpha_1 = \text{mean} = \int_{-\infty}^{\infty} x f(x).\,dx = E(X)$$

$$= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}.\,dx$$

put $\frac{x-m}{\sqrt{2}\sigma} = z$

$$= \int_{-\infty}^{\infty} \frac{(\sqrt{2}\sigma z + m) e^{-z^2}}{\sqrt{\pi}} dz \qquad dz = \frac{dx}{\sqrt{2}\sigma}$$

$$= \frac{1}{\sqrt{\pi}} \left\{ \int_{-\infty}^{\infty} \sqrt{2}\sigma z e^{-z^2} dz + m\int_{-\infty}^{\infty} e^{-z^2}.\,dz \right\}$$

$$= \frac{1}{\sqrt{\pi}} \left\{ 0 + m\int_{-\infty}^{\infty} e^{-z^2} dz \right\}$$

$$= \frac{m}{\sqrt{\pi}} \times \sqrt{\pi}$$

$$\boxed{\alpha_1 = m}$$

Varia

$$\alpha_2 = \int_{-\infty}^{\infty} x^2 f(x).\,dx$$

$$= \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

$$\frac{x-m}{\sqrt{2}\sigma} = z$$

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (m^2 + 2\sqrt{2}\sigma m z + 2\sigma^2 z^2)\, e^{-z^2} dz$$

$$= \frac{1}{\sqrt{\pi}} \left[m^2\sqrt{\pi} + 0 + 2\sigma^2 \int_{-\infty}^{\infty} z^2 e^{-z^2} dz \right.$$

$$= \frac{1}{\sqrt{\pi}} \left[m^2\sqrt{\pi} + 4\sigma^2 \int_{0}^{\infty} z^2 e^{-z^2} dz \right]$$

put $z^2 = y$

$2z\,dz = dy$

$$\int_0^{\infty} \sqrt{y}\, e^{-y} \frac{dy}{2}$$

$$\frac{1}{2}\int_0^{\infty} y^{3/2-1} e^{-y} dy$$

$$\frac{1}{2}\Gamma\left(\frac{3}{2}\right) = \frac{1}{2} \times \frac{\sqrt{\pi}}{2} = \frac{\sqrt{\pi}}{4}$$

$$= \frac{1}{\sqrt{\pi}} \left[m^2\sqrt{\pi} + 4\sigma^2 \times \frac{\sqrt{\pi}}{4} \right]$$

$$\alpha_2 = m^2 + \sigma^2$$

Variance $= \alpha_2 - \alpha_1^2$

$= m^2 + \sigma^2 - m^2$

$= \sigma^2$

standard deviation $= \sigma$

Significance of σ:

NOTE: Put $m=0, \sigma=1$,

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

↳ Standard normal distribution

$\frac{x-m}{\sigma}$ is a <u>scaling function</u>

22/08/2017

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

$$\mu_K = E(X-m)^K$$

$$= \int_{-\infty}^{\infty} (x-m)^K f(x).dx$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-m)^{K-1} (x-m) e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

$$\int (x-m) e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

put $\frac{x-m}{\sqrt{2}\sigma} = z$ $\quad dz = \frac{dx}{(\sqrt{2}\sigma)}$

$$\int (z\sqrt{2}\sigma) e^{-z^2} (\sqrt{2}\sigma) dz$$

$$2\sigma^2 \int z e^{-z^2} dz$$

$$\sigma^2 (-e^{-z^2})$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\left\{\left(-\underbrace{(x-m)^{k-1}\sigma^2 e^{-\frac{(x-m)^2}{2\sigma^2}}}_{0}\right)_{-\infty}^{\infty} + \sigma^2(k-1)\int_{-\infty}^{\infty}(x-m)^{k-2} e^{-\frac{(x-m)^2}{2\sigma^2}}dx\right\}$$

$$= \sigma^2(k-1)\mu_{k-2}$$

$$\boxed{\mu_k = \sigma^2(k-1)\mu_{k-2}}$$

$\mu_0 = 1$

$k=2$

$\mu_2 = \sigma^2(1)\mu_0 = \sigma^2$

$k=4$

$\mu_4 = 3\sigma^2\mu_2 = 3\sigma^4$

$k=6$

$\mu_6 = 5.3.1\,(\sigma^2)^3$

$\mu_{2k} = 1.3.5\ \ldots\ (2k-1)\,(\sigma^2)^k$

$= 1.3.5\ldots(2k-1)\,\sigma^{2k}$

Put $k=3$

$\mu_3 = \sigma^2.2.\mu_1 = 0$

$\mu_{2k+1} = 0$

$$I_7 = \int_0^{\pi/2} \sin^7 x.dx$$

$$= \frac{6.4.2}{7.5.3.1}.1$$

$$I_8 = \int_0^{\pi/2} \sin^8 x\,dx$$

$$= \frac{7.5.3.1}{8.6.4.2}.\frac{\pi}{2}$$

Scanned by CamScanner

Binomial Theorem

$$P(A_r) = {}^nC_r\, p^r q^{n-r} \quad , \; r = 0,1,2,\dots n$$

f_r is the probability of A_r

$$\alpha_1 = E(X) = \sum_{r=-\infty}^{\infty} r f_r$$

$$= \sum_{r=0}^{n} r f_r$$

$$= \sum_{r=1}^{n} r f_r$$

$$= \sum_{r=1}^{n} \frac{r\,n!}{(n-r)!\,r!} p^r q^{n-r}$$

$$= \sum_{r=1}^{n} \frac{n!}{(n-r)!\,(r-1)!} p^r q^{n-r}$$

$$= np \sum_{r=1}^{n} \frac{(n-1)!}{(n-r)!\,(r-1)!} p^{r-1} q^{n-r}$$

$$\sum_{r=0}^{n} \frac{n!}{r!\,(n-r)!} p^r q^{n-r} = (p+q)^n$$

$$= np\,(p+q)^{n-1}$$

$$= np$$

$$\alpha_2 = E(X^2)$$

$$= \sum_{r=0}^{n} r^2 f_r$$

$$= \sum_{r=0}^{n} [r(r-1)+r] f_r$$

$$= \sum_{r=2}^{n} r(r-1) f_r + \sum_{r=0}^{n} r f_r$$

$$= \sum_{r=2} \frac{n!\, p^r q^{n-r}}{(n-r)!\,(r-2)!} + \alpha_1$$

$$= n(n-1)p^2 \sum_{r=2}^{n} \frac{(n-2)!\, p^{r-2} q^{n-r}}{(n-r)!\,(r-2)!}$$

$$= n(n-1)\, p^2\, (p+q)^{n-2} + np$$

$$\alpha_2 = n(n-1)p^2 + np$$

Variance

$$\mu_2 = \alpha_2 - \alpha_1^2$$

$$= (n^2 - n)p^2 + np - n^2p^2$$

$$= -np^2 + np$$

$$= np(1-p)$$

$$\boxed{\mu_2 = npq}$$

standard deviation $= \sqrt{npq}$

Q. Find mean and variance of Poisson process

$$P(A_r) = \frac{e^{-\mu}\mu^r}{r!}$$

$$\alpha_1 = E(X) = \sum_{r=0}^{\infty} \left(\frac{e^{-\mu}\mu^r}{r!}\right) \cdot r$$

$$= \sum_{r=0}^{\infty} \frac{e^{-\mu}\mu^r}{(r-1)!}$$

$$= e^{-\mu}\mu \sum_{r=1}^{\infty} \frac{\mu^{r-1}}{(r-1)!}$$

$$= \mu e^{-\mu} e^{\mu}$$

$$\alpha_1 = \mu$$

$\alpha_2 = E(X^2)$

$$= \sum_{r=0}^{\infty} r^2 \frac{e^{-\mu}\mu^r}{r!}$$

$$= \sum_{r=1}^{\infty} \frac{r^2 e^{-\mu}\mu^r}{r!}$$

$$= \sum_{r=1}^{\infty} \frac{(r(r-1)+r)\, e^{-\mu}\mu^r}{r!}$$

$$= \sum_{r=1}^{\infty} \frac{r(r-1)e^{-\mu}\mu^r}{r!} + \mu$$

$$= e^{-\mu}\mu^2 \sum_{r=2}^{\infty} \frac{\mu^{r-2}}{(r-2)!}$$

$$\alpha_2 = \mu^2 + \mu .$$

$$\text{variance} = \alpha_2 - \alpha_1^2$$

$$= \mu^2 + \mu - \mu^2$$

$$= \mu .$$

NOTE: This is the distribution where mean and variance is same

Uniform distribution:

$$f(x) = \frac{1}{b-a} \quad , \quad a < x < b$$

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

Cauchy distribution:

$\lambda, \mu \rightarrow$ parameters

$$f(x) = \frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x-\mu)^2} \quad , \quad -\infty < x < \infty$$

$$\int_{-\infty}^{\infty} \frac{\lambda}{\lambda^2 + (x-\mu)^2} \cdot dx$$

$$\frac{\lambda}{\lambda} \int \frac{dx}{\lambda^2 + (x-\mu)^2}$$

$$\frac{\lambda}{\lambda} \left[\tan^{-1}\left(\frac{x-\mu}{\lambda}\right) \right]_{-\infty}^{+\infty} = 1$$

$$E(X) = \int_{-\infty}^{\infty} x f(x) \cdot dx$$

$$= \int_{-\infty}^{\infty} \frac{x \lambda \cdot dx}{\pi (\lambda^2 + (x-\mu)^2}$$

$$= \frac{\lambda}{\pi} \int_{-\infty}^{\infty} \frac{x \cdot dx}{\lambda^2 + (x-\mu)^2}$$

$$\lambda^2 + (x-\mu)^2 = t$$

$$dt = 2\cancel{(x-\mu)}\, dx$$

$$x - \mu = z$$

$$dz = dx$$

$$= 2 \frac{\lambda}{\pi} \int_{0}^{\infty} \frac{dt}{2(x-\mu) t}$$

$$= \frac{\lambda}{\pi} \int_{-\infty}^{\infty} \frac{(z+\mu)}{\lambda^2 + z^2} dz$$

$$= \frac{\lambda}{\pi} \left\{ \int_{-\infty}^{\infty} \frac{z}{\lambda^2 + z^2} \cdot dz + \mu \int_{-\infty}^{\infty} \frac{dz}{\lambda^2 + z^2} \right.$$

Does not exist

$$\frac{1}{2}\int_0^{\infty} \cancel{dt} \qquad \frac{\lambda}{\pi}\left[\cancel{0} \quad + \frac{\mu}{\lambda} \times \frac{\pi}{2}\right]$$

$\cancel{=\mu}$

The mean does not exist

24/08/2017

Transformation of Variables

Let, $y = g(x)$ is continuously differentiable function.

$$\left[\begin{array}{l} f \in C^n[a,b] \\ f^n(x) \text{ is continuous in } [a,b] \\ \text{If } C^{\infty}\text{, then smooth function} \end{array}\right]$$

Let y is a monotonic

Case 1: If $\frac{dy}{dx} > 0$

$$X \le x \Rightarrow g(X) \le g(x)$$
$$Y \le y$$

~~P(X ≤ x~~

$$P(-\infty < X \le x) = P(-\infty < Y \le y)$$
$$F(x) = F(y)$$
$$F'(x)\,dx = F'(y)\,dy$$
$$f(x)\,dx = f(y)\,dy$$
$$f(y) = f(x)\,\frac{dx}{dy}$$

ca

Q1. If X

of a

Sol^n:

Case 2 : If $\frac{dy}{dx} < 0$

$X \leq x \Rightarrow g(X) \geq g(x)$

$g(x) \leq g(X)$

$y \leq Y$

$P_r(-\infty < X \leq x) = P(Y \geq y)$

$= 1 - P(-\infty < Y \leq y)$

$F(x) = 1 - F(y)$

$F'(x)dx = -F'(y)dy$

$f(x)dx = -f(y)dy$

$f(y) = -f(x)\frac{dx}{dy}$

$$f(y) = f(x)\left|\frac{dx}{dy}\right|$$

Q1. If $X \sim N(m, \sigma)$, then find the distribution

↓ Normal distribution function

of $aX+b$, when a, b are constants

Sol^n: $\because X \sim N(m, \sigma)$

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

Let $Y = aX+b$

$y = ax+b$ $\qquad \frac{y-b}{a} = x$

$\frac{dy}{dx} = a$ $\qquad$ as $x \to -\infty$ to ∞, $y \to -\infty$ to ∞

$$\therefore f(y) = f(x)\left|\frac{dx}{dy}\right|$$

$$= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \frac{1}{|a|}$$

$$= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\left(\frac{y-b}{a}-m\right)^2}{2\sigma^2}} \frac{1}{|a|}$$

$$= \frac{1}{\sqrt{2\pi}\sigma|a|} e^{-\frac{(y-b-am)^2}{2\sigma^2 a^2}}$$

$Y \sim N(am+b, \sigma|a|)$

Ex2 If X represents $\beta_2(m,n)$ then show that $\frac{1}{X}$ represents $\beta_2(m,n)$

Solution:

$$f(x) = \frac{x^{m-1}}{(1+x)^{m+n}\,\cancel{\beta(m,n)}\;\beta_2(m,n)}, \quad 0 < x < \infty$$

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

Let $Y = \frac{1}{X}$

$y = \frac{1}{x}$

$\frac{dy}{dx} = -\frac{1}{x^2} < 0$ (Monotonic)

as $x \to 0$ to ∞

$y \to 0$ to ∞

$$f(y) = f(x)\left|\frac{dx}{dy}\right|$$

$$= \frac{x^{m-1}}{(1+x)^{m+n}\beta_2(m,n)}(x^2)$$

$$= \frac{x^{m+1}}{(1+x)^{m+n}\beta_2(m,n)}$$

$$= \frac{\left(\frac{1}{y}\right)^{m+1}}{\left(1+\frac{1}{y}\right)^{m+n}\beta_2(m,n)}$$

$$= \frac{y^{m+n}}{y^{m+1}(1+y)^{m+n}\beta_2(m,n)}$$

$$= \frac{y^{n-1}}{(1+y)^{m+n}\beta_2(m,n)}$$

Ex3

Solution

F

Ex3 If $X \sim N(0,1)$ then s.t. $\frac{1}{2}X^2 \sim \Gamma(\frac{1}{2})$

Solution If $X \sim N(0,1)$

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

Let $Y = \frac{1}{2}X^2$

$\Rightarrow y = \frac{1}{2}x^2$

$\frac{dy}{dx} = x > 0$ in $(0, \infty)$

< 0 in $(-\infty, 0)$

Moreover if x traverses to $-\infty$ to ∞ y traverses in $(0,\infty)$ in opposite direction

$P(y < Y \leq y+dy)$

$= P\left(\frac{1}{2}x^2 < \frac{1}{2}X^2 \leq \frac{1}{2}(x+dx)^2\right)$

$= P(x^2 < X^2 \leq (x+dx)^2)$

$= P(x < X \leq x+dx) + P(-x-dx \leq X \leq -x)$

$f(y)\,dy = 2P(x < X \leq x+dx)$

$f(y)\,dy = 2f(x).dx$

$\frac{F(y+dy) - F(y)}{dy}\,dy$

$F'(y)dy$

$f(y)dy$

$f(y) = 2f(x)\frac{dx}{dy}$

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

$= \frac{\sqrt{2}}{\sqrt{\pi}} \frac{e^{-y}}{\sqrt{2}\sqrt{y}}$

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

$= \frac{1}{\sqrt{\pi}} e^{-y} y^{\frac{1}{2}-1} \quad 0 < y < \infty$

$= \frac{e^{-y} y^{\frac{1}{2}-1}}{\Gamma(\frac{1}{2})} \quad 0 < y < \infty$

Moment generating function (m.g.f)

The m.g.f of a random variable X is defined by

$$\psi(t) = E(e^{tX})$$

$$= \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$= \sum_{r=-\infty}^{\infty} e^{tr} f(r) \quad (1)$$

$$\psi(t) = E\left(1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \frac{t^r x^r}{r!}\right)$$

$$= 1 + tE(x) + \frac{t^2}{2!}E(x^2) + \frac{t^3}{3!}E(x^3) \cdots + \frac{t^r}{r!}E(x^r)$$

$$= 1 + t\alpha_1 + \frac{t^2}{2!}\alpha^2 + \frac{t^3}{3!}\alpha^3 - \cdots + \frac{t^r}{r!}\alpha^r$$

$$= \sum_{r=0}^{\infty} \frac{t^r}{r!}\alpha^r \quad (2)$$

- α_r can be found by comparing the coeffs. of $\frac{t^r}{r!}$

Characteristic function

Characteristic function of a random variable X is defined by

$$\chi(t)$$

$$\chi(t) = E(e^{itX}) \quad (3)$$

$$|e^{itX}| = 1$$

$$\chi(t) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!}\alpha^r \quad (4)$$

α_r can be found by comparing the coeffs. of $\frac{(it)^r}{r!}$

. characteristic function uniquely determines the moment of the given distribution f's.

NOTE:
Entire function which is differentiable everywhere

Q1. Find the characteristic function of the $N(m,\sigma)$ distribution and hence find mean and variance

Solution:

$$\chi(t) = E(e^{itX})$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{itx} e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

put $\frac{x-m}{\sqrt{2}\sigma} = z$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{it(\sqrt{2}\sigma z+m)} e^{-z^2} dz\, \sqrt{2}\sigma$$

$$= \frac{e^{mit}}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{it\sqrt{2}\sigma z - z^2} dz$$

$$= \frac{e^{imt}}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-(z^2 - \sqrt{2}\sigma z it)} dz$$

$$= \frac{e^{imt}}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-\left\{\left(z^2 - 2z\left(\frac{it\sigma}{\sqrt{2}}\right) + \left(\frac{it\sigma}{\sqrt{2}}\right)^2 - \left(\frac{it\sigma}{\sqrt{2}}\right)^2\right\}} dz$$

$$= \frac{e^{imt}}{\sqrt{\pi}} e^{-\frac{t^2\sigma^2}{2}} \int_{-\infty}^{\infty} e^{-\left(z - \frac{it\sigma}{\sqrt{2}}\right)^2} dz$$

$$= \frac{e^{imt - \frac{t^2\sigma^2}{2}}}{\sqrt{\pi}} \cdot \sqrt{\pi}$$

$$\chi(t) = e^{imt - \frac{\sigma^2 t^2}{2}}$$

$$= 1 + \left(imt - \frac{\sigma^2 t^2}{2}\right) + \frac{\left(imt - \frac{\sigma^2 t^2}{2}\right)^2}{2!}$$

$\alpha_1 = \text{coeff of } (it) = m$

$\alpha_2 = \text{coeff of } \frac{(it)^2}{2!} = \sigma^2 + m^2$

$\text{Variance} = \sigma^2 + m^2 - m^2$

$= \sigma^2$

Q. Find μ_3 for Gamma distribution function using characteristic function.

$\mu_3 = E(X-m)^3$

$= E(X^3 - m^3 - 3mX^2 + 3m^2X)$

$= E(X^3) - m^3 - 3mE(X^2) + 3m^2E(X)$

$= \alpha_3 - m^3 - 3m\alpha_2 + 3m^3$

$= \alpha_3 - 3m\alpha_2 + 2m^3$

$\mu_3 = \alpha_3 - 3\alpha_1\alpha_2 + 2\alpha_1^3$

$$f(x) = \begin{cases} \dfrac{e^{-x} x^{n-1}}{\Gamma(n)}, & 0 < x < \infty \\ 0, & \text{elsewhere} \end{cases}$$

$\chi(t)$

$(1+x)^{-n} = 1-$

$$\chi(t) = E(e^{itx})$$

$$= \int_0^\infty \frac{e^{itx} e^{-x} x^{n-1}}{\Gamma(n)} dx$$

$$= \frac{1}{\Gamma(n)} \int_0^\infty e^{itx-x} x^{n-1} dx$$

$$= \cancel{\frac{1}{\Gamma(n)} \int_0^\infty e^{i(tx+ix)} x^{n-1} dx}$$

$$= \frac{1}{\Gamma(n)} \int_0^\infty e^{-x(1-it)} x^{n-1} dx$$

$$x(1-it) = z$$

$$dz = dx(1+it)$$

$$= \frac{1}{\Gamma(n)} \int_0^\infty e^{-z} \left(\frac{z}{1-it}\right)^{n-1} \frac{dz}{(1-it)}$$

$$= \frac{1}{\Gamma(n)} \frac{1}{(1-it)^n} \int_0^\infty e^{-z} z^{n-1} dz$$

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

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

$$\boxed{(1+x)^{-n} = 1 - nx + \frac{n(n+1)x^2}{2!} - \frac{1}{3!} n(n+1)(n+2)x^3}$$

$$= (1-it)^{-n}$$

$$= 1 + nit + \frac{n(n+1)}{2!}(+it)^2 + \frac{n(n+1)(n+2)}{3!}(it)^3 + \frac{n(n+1)(n+2)(n+3)}{4!}(it)^4$$

$\alpha_1 = n$

$\alpha_2 = n(n+1)$ $\qquad \mu_2 = \alpha_2 - \alpha_1^2 = n^2+n-n^2$, $\mu_2 = n$

$\alpha_3 = n(n+1)(n+2)$

$\alpha_4 = n(n+1)(n+2)(n+3)$

$\mu_3 = \alpha_3 - 3\alpha_1\alpha_2 + 2\alpha_1^3$

$= n(n+1)(n+2) - 3n \times n(n+1) + 2n^3$

$= n(n^2+3n+2) - 3(n^3+n^2) + 2n^3$

$= \cancel{n^3} + \cancel{3n^2} + 2n - \cancel{3n^3} - \cancel{3n^2} + \cancel{2n^3}$

$= 2n$

$\mu_4 = \alpha_4 - 4\alpha_1\alpha_3 + 6\alpha_1^2\alpha_2 - 3\alpha_1^4 = 3n^2+6n$

Q1. Find ~~for~~ the chara. function of ~~poisson~~ Binomial ~~process~~ distribution and hence find mean & variance

$X(t) = E(e^{itX})$

$= \sum_{r=0}^{n} e^{itr}\, {}^nC_r p^r q^{n-r}$

$= \sum_{r=0}^{n} {}^nC_r (pe^{it})^r q^{n-r}$

$X(t) = (pe^{it}+q)^n$

$= (pe^{it}+1-p)^n$

$= (1+p(-1+e^{it}))^n$

~~$= 1 + p(1-e^{it})$~~ ...

$= 1 + p(1-(1+it+\frac{(it)^2}{2!}+\frac{(it)^3}{3!}--))$

$= \left[1+p\left(+it+\frac{i^2t^2}{2!}+\frac{i^3t^3}{3!}+-\right)\right]^n$

$= 1+np\left(+it+\frac{i^2t^2}{2!}+\frac{(it)^3}{3!}--\right)$

$+ \frac{n(n-1)p^2}{2!}\left(it+\frac{i^2t^2}{2!}+\frac{(it)^3}{3!}-\right)^2$

Q2. Find the

hence fi

Ans

$= 1+\mu$

Scanned by CamScanner

$$\alpha_1 = np$$

$$\alpha_2 = np + n(n-1)p^2$$

$$\text{variance} = \alpha_2 - \alpha_1^2$$
$$= np + n(n-1)p^2 - n^2p^2$$
$$= np + \cancel{n^2p^2} - np^2 - \cancel{n^2p^2}$$
$$= np(1-p)$$
$$= npq$$

Q2. Find the characteristic function of poisson process and hence find mean and variance.

Ans
$$\chi(t) = E(e^{itx})$$

$$f(x) = \frac{e^{-\mu}\mu^x}{x!} \qquad 0 < x < \infty$$

$$= \sum_{x=0}^{\infty} e^{itx} \frac{e^{-\mu}\mu^x}{x!}$$

$$= e^{-\mu} \sum_{x=0}^{\infty} e^{itx} \cdot \frac{\mu^x}{x!}$$

$$= e^{-\mu} \sum_{x=0}^{\infty} \frac{(e^{it} \cdot \mu)^x}{x!}$$

$$= e^{-\mu} e^{\mu e^{it}}$$

$$\chi(t) = e^{\mu(e^{it}-1)}$$

$$= 1 + \mu(e^{it}-1) + \frac{\mu^2(e^{it}-1)^2}{2!} + \frac{\mu^3}{3!}(e^{it}-1)^3 \dots$$

$$= 1 + \mu\left(it + \frac{(it)^2}{2!} + \frac{(it)^3}{3!} \dots\right) + \frac{\mu^2}{2!}\left(it + \frac{(it)^2}{2!} + \frac{(it)^3}{3!}\right)^2 + \frac{\mu^3}{3!}\left(\right.$$

$\alpha_1 = \mu$

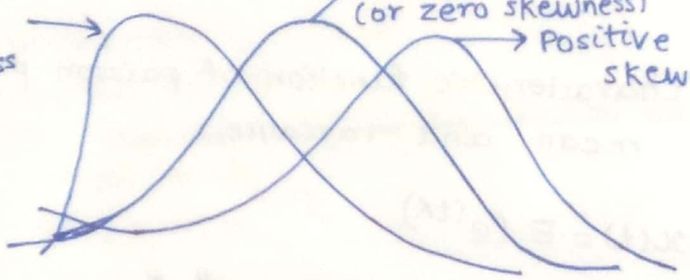
$\alpha_2 = \mu + \mu^2$

$\mu_2 = \alpha_2 - \alpha_1^2$

$= \mu + \mu^2 - \mu^2$

$\boxed{\mu_2 = \mu}$

Skewness 29/08/2017



coeff. of skewness

$\gamma_1 = \frac{\mu_3}{\sigma^3}$

For Gamma distribution

$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{2n}{n^{3/2}} = \frac{2}{\sqrt{n}}$

$\mu_{2k} = 1.3.5 \ldots (2k-1)\sigma^{2k}$

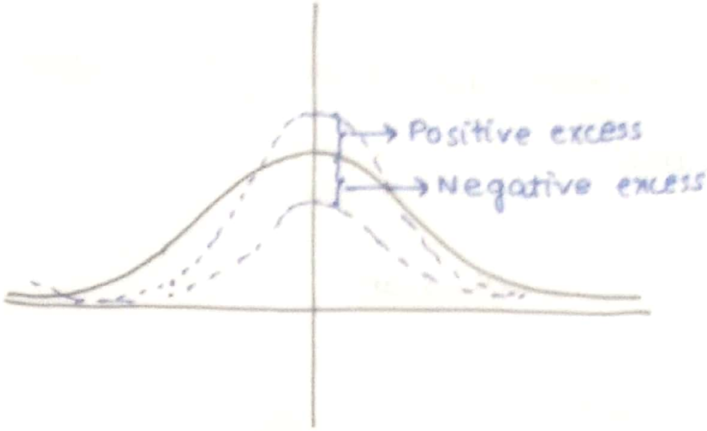
$\mu_{2k+1} = 0$

Kurtosis : The coeff. of kurtosis

$\beta_2 = \frac{\mu_4}{\sigma^4}$

For normal distribution $\beta_2 = \frac{3\sigma^4}{\sigma^4} = 3$

Excess of Kurtosis : $\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$



Another measure of skewness $= \frac{m - M}{\sigma}$ (M → Mode)

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

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

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

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

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

$$x = n-1$$

For Gamma distribution

$$= \frac{n-(n-1)}{\sqrt{n}}$$

$$= \frac{1}{\sqrt{n}}$$

[NOTE : For same distribution $\frac{1}{2}\gamma_1$, becomes coeff. of skewness]
(if not same)

Median

distribution function $\longrightarrow F(\mu) = 1/2 \rightarrow$ median

NOTE : median is also a measure of location.

Quantiles

$X(t)$ $\xi_p \longrightarrow$ quantile $(0<p<1)$ (Hyponym of quartiles)

$F(\xi_p) = p$

$F(\xi_{1/2}) = \frac{1}{2} \Rightarrow \xi_{1/2}$ is the median

lower quartile $\leftarrow$ 1/4   1/2   3/4 $\rightarrow$ upper quartile

0 —— $\xi_{1/4}$ —— $\xi_{1/2}$ —— $\xi_{3/4}$ —— 1

Quartile deviation $= \frac{1}{2}(\xi_{3/4} - \xi_{1/4})$

measure of dispersion

Cauchy distribution

$$f(x) = \frac{1}{\pi} \frac{\lambda}{\lambda^2+(x-\mu)^2} \quad , -\infty<x<\infty$$

$$F(x) = \int_{-\infty}^{x} f(x)\,dx$$

$$= \frac{1}{\pi}\int_{-\infty}^{x} \frac{\lambda}{\lambda^2+(x-\mu)^2}\,dx$$

$$= \frac{1}{\pi}\left[\tan^{-1}\left(\frac{x-\mu}{\lambda}\right)\right]_{-\infty}^{x}$$

$$= \frac{1}{\pi}\left[\tan^{-1}\left(\frac{x-\mu}{\lambda}\right)+\frac{\pi}{2}\right]$$

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

$F(\mu-\lambda) = \frac{1}{4}$

$\xi_{1/4} = \mu-\lambda.$

$F(\lambda+\mu) = \frac{3}{4}$

$\xi_{1/4} = \frac{1}{4}$

$\xi_{3/4} = \frac{3}{4}$

quartile deviation $= \frac{1}{2}(\xi_{3/4} - \xi_{1/4})$

$= \lambda$

31/08/2017

Two dimensional distribution

Let X and Y are two random variables in the space S.

The distribution function of X and Y is defined by

$$F(x, y) = P(-\infty < X \leq x, -\infty < Y \leq y)$$

The joint occurrence

$$(-\infty < X \leq x, -\infty < Y \leq y) = (-\infty < X \leq x)(-\infty < Y \leq y)$$

NOTE: If x and y are independent variables

$$P(-\infty < X \leq x, -\infty < Y \leq y) = P(-\infty < X \leq x) . P(-\infty < Y \leq y)$$

$$\boxed{F_{x,y}(x, y) = F_x(x) . F_y(y)}$$

marginal distribution function

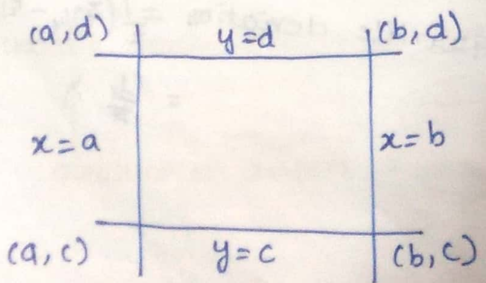
$$(-\infty < X \leq x, -\infty < Y < \infty) = (-\infty < X \leq x) . (-\infty < Y < \infty)$$

$$= (-\infty < X \leq x) . S$$

$$= (-\infty < X \leq x)$$

$$F(x, \infty) = F_x(x)$$

Similarly, $F_y(y) = F(\infty, y)$



$P(a < X \leq b,\ c < Y \leq d)$

$$= F(b,d) + F(a,c) - F(b,c) - F(a,d)$$

$$\int_{y=c}^{d} \int_{x=a}^{b} \frac{\partial^2 F}{\partial x \partial y}\, dx\, dy$$

$$= \int_{y=c}^{d} dy \left(\cancel{\partial^2} \int_{x=a}^{b} \cancel{dy}\, \frac{\partial^2 F}{\partial x . \partial y}\, dx \right)$$

$$= \int_{y=c}^{d} \left(\frac{\partial F}{\partial y} \right)_{x=a}^{b} dy$$

$$= \int_{y=c}^{d} \left(\frac{\partial F(b,y)}{\partial y} - \frac{\partial F(a,y)}{\partial y} \right) . dy$$

$$= \int_{y=c}^{d} \frac{\partial F(b,y)}{\partial y} . dy - \int_{y=c}^{d} \frac{\partial F(a,y)}{\partial y} . dy$$

$$= F(b,d) - F(b,c) - F(a,d) + F(a,c)$$

$$= P(a < X \leq b,\ c < Y \leq d)$$

$P(a < X \le b, c < Y \le d)$

$$= \int_{y=c}^{d}\int_{x=a}^{b} \frac{\partial^2 F}{\partial x \cdot \partial y}\, dx\, dy = \int_{y=c}^{d}\int_{x=a}^{b} f(x,y)\, dx\, dy$$

$$\frac{\partial^2 F}{\partial x \cdot \partial y} = f(x,y) \downarrow$$

Two dimensional probability density function

$P(-\infty < X \le x, -\infty < Y \le y)$

$$F(x,y) = \int_{y=-\infty}^{y}\int_{x=-\infty}^{x} f(x,y)\, dx\, dy$$

$P(-\infty < X < \infty, -\infty < Y < \infty)$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, dx\, dy = 1$$

$$F(x,y) = \int_{x=-\infty}^{x}\int_{y=-\infty}^{y} f(x,y)\, dx\, dy$$

$$y = \int_{-\infty}^{x} f(x)\, dx$$

$$F(x,\infty) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(x,y)\, dy\, dx$$

$$\frac{dy}{dx} = f(x)$$

$$F'(x,\infty) = \int_{-\infty}^{\infty} f(x,y)\, dy$$

$$F_x'(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$$

$$f_x(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$$

marginal density function

$$f_x(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$$

$$f_y(y) = \int_{-\infty}^{\infty} f(x,y)\, dx$$

Q.

$$f(x,y) = k.(1-x-y), \quad x>0, y>0, x+y \le 1$$
$$= 0 \quad , \text{ elsewhere}$$

Hence find probability of

$$P(X<\tfrac{1}{2}, Y>\tfrac{1}{4})$$

Also show whether f(x,y) is independent.

✓ (K=6)

$$P(\cancel{-\infty} < X < \tfrac{1}{2}, \tfrac{1}{4} < Y \cancel{\le \infty})$$

$$= \int_{1/4}^{1/2} \int_{0}^{1/2} K(1-x-y)\, dx\, dy$$

$$= \int_{1/4}^{1/2} \left[K\left(x - \frac{x^2}{2} - xy\right)\right]_0^{1/2} dy$$

$$= \int_{1/4}^{1/2} K\left(\frac{1}{2} - \frac{1}{8} - \frac{y}{2}\right) dy$$

$$= \int_{1/4}^{1/2} K\left(\frac{3}{8} - \frac{y}{2}\right) dy$$

$$= K\left(\frac{3}{8}y - \frac{y^2}{4}\right)_{1/4}^{1/2}$$

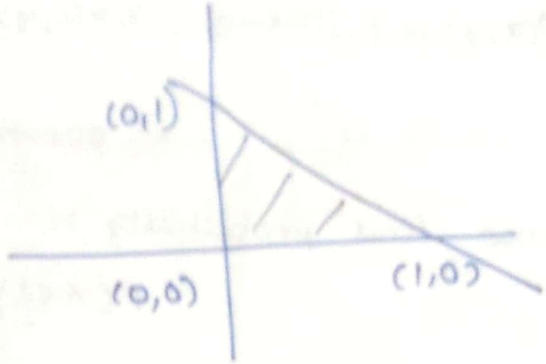
$$= K\left[\left(\frac{3}{8}\times\frac{1}{2} - \frac{1}{16}\right) - \left(\frac{3}{8}\times\frac{1}{4} - \frac{1}{64}\right)\right]$$

$$= K\left[\frac{2}{16} - \left(\frac{5}{64}\right)\right]$$

$$= \frac{3K}{64}$$

Scanned by CamScanner

18



$$\int_0^1 \int_{y=0}^{1-x} \ldots \, dx\,dy$$

$x \geq 0, y \geq 0, z \geq 0,$
$x+y+z \leq 1$

$P(X < \frac{1}{2}, Y > \frac{1}{4})$

$$= \int_0^{1/2} \int_{1/4}^{1-x} f(x,y) \quad dy\,dx$$

$$= \frac{13K}{48 \times 4} = \frac{13}{32}$$

04/09/2017

$$P(x < X \leq x+dx)$$

$$= F(x+dx) - F(x)$$

$$= \frac{F(x+dx) - F(x)}{dx} \cdot dx$$

$$= F'(x) \cdot dx$$

$$= f(x) \cdot dx$$

$$P(x < X \leq x+dx,\ y < Y \leq y+dy)$$

$$= F(x+dx, y+dy) + F(x,y) - F(x, y+dy) - F(x+dx, y)$$

$$\boxed{f(x+h, y+k) = f(x,y) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f(x,y) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 \frac{f(x,y)}{2!} + \cdots}$$

$$= \cancel{F(x,y)} + \left(\cancel{dx\frac{\partial}{\partial x}} + \cancel{dy\frac{\partial}{\partial y}}\right) F(x,y) + \frac{1}{2!}\left(dx\frac{\partial}{\partial x} + dy\frac{\partial}{\partial y}\right)^2 F(x,y)$$

$$+ \cancel{F(x,y)} - \left[\cancel{F(x,y)} + \left(\cancel{dx\frac{\partial}{\partial x}}\right) F(x,y) + \frac{1}{2!}\left(dx\frac{\partial}{\partial x}\right)^2 F(x,y)\right]$$

$$- \left[\cancel{F(x,y)} + \cancel{dy\frac{\partial}{\partial y}} F(x,y) + \frac{1}{2!}\left(dy\frac{\partial}{\partial y}\right)^2 F(x,y)\right]$$

$$= \frac{1}{2!} \times 2\, dx\, dy \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

$$= f(x,y)\, dx\, dy$$

Transformation of variables for two dimension:

if negative then we have to [illegible] [illegible] [illegible]

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

$$f(u,v) = f(x,y) \left| \frac{\partial(x,y)}{\partial(u,v)} \right|$$

[e.g. $x = r\cos\theta \qquad y = r\sin\theta$

$$\frac{\partial x}{\partial \theta} = -r\sin\theta \qquad \frac{\partial y}{\partial \theta} = r\cos\theta$$

$$\frac{\partial x}{\partial r} = \cos\theta \qquad \frac{\partial y}{\partial r} = \sin\theta$$

$$\begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r$$

$$\iint (x^2+y^2)^{n/2}\, dx\, dy$$

$$\iint r^n |J|\, dr\, d\theta$$

Q1. The joint distribution of two random variables X and Y are

$$f(x,y) = x+y, \quad 0<x<1,\ 0<y<1$$

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

Find the distribution XY.

Solution: Let $u = xy$, $v = x$

$\cancel{y = u - v}$

$y = \frac{u}{v}$

$\cancel{x = v}$

$$f(u,v) = f(x,y) \left| \frac{\partial(x,y)}{\partial(u,v)} \right|$$

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

$\frac{\partial x}{\partial u} = 0$

$\frac{\partial x}{\partial v} = 1$

$\frac{\partial y}{\partial u} = \frac{1}{v} \qquad \frac{\partial y}{\partial v} = -\frac{u}{v^2}$

$$= \begin{vmatrix} 0 & 1 \\ \frac{1}{v} & -\frac{u}{v^2} \end{vmatrix} = -\frac{1}{v}$$

Tomorrow off

$= f(x,y)\left(\frac{1}{v}\right)$

$= (x+y)\left(\frac{1}{v}\right)$

$= \left(v+\frac{u}{v}\right)\frac{1}{v}$

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

$0<\frac{u}{v}<1$

$0<u<v$

$0<v<1$

$0<u<v<1$

$f(u) = \int_{-\infty}^{\infty} f(u,v)\,dv$

$= \int_{u}^{1} \left(1+\frac{u}{v^2}\right)dv$

Q2. If $X \sim \Gamma(l)$, $Y \sim \Gamma(m)$ are independent X and Y

Show that
(i) $X+Y \sim \Gamma(m+l)$ and (ii) $X/Y \sim \beta_2(l,m)$

Solution

Given

$$f(x,y) = f(x).f(y)$$
$$= \frac{e^{-x}x^{l-1}}{\Gamma(l)} \cdot \frac{e^{-y}y^{m-1}}{\Gamma(m)}$$
$$0<x<\infty,\ 0<y<\infty$$

(i) ~~$u = x+y$~~

$$(x,y) \to (u,v)$$
$$u = x+y \qquad v = \frac{x}{y}$$

We have to find $f(u)$ and $f(v)$

$$u = vy+y$$
$$y = \frac{u}{v+1}$$
$$v = \frac{x(v+1)}{u}$$
$$x = \frac{uv}{v+1}$$

$$\frac{\partial x}{\partial u} = \frac{v}{v+1} \qquad \frac{\partial x}{\partial v} = u\left[\frac{(v+1)-v}{(v+1)^2}\right] = \frac{u}{(v+1)^2}$$

$$\frac{\partial y}{\partial u} = \frac{1}{v+1} \qquad \frac{\partial y}{\partial v} = \frac{-u}{(v+1)^2}$$

$$\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

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

$$f(u,v) = f(x,y)\left|\frac{\partial(x,y)}{\partial(u,v)}\right|$$
$$= \frac{e^{-x}x^{l-1}.e^{-y}y^{m-1}}{\Gamma(l).\Gamma(m)} \cdot \frac{u}{(v+1)^2}$$
$$= \frac{e^{-u}.u^{l+m-1}v^{m-1}}{\Gamma(l)(v+1)^{l+m}\Gamma(m)}$$
$$= \frac{e^{-u}(u)^{l+m-1}v^{m-1}}{\beta_1(l,m)\ \Gamma(l+m)(v+1)^{l+m}} \quad 0<u<\infty,\ 0<v<\infty$$

$$f(u) = \int_0^\infty \frac{e^{-u} u^{l+m-1}}{\Gamma(l+m)} \frac{v^{m-1} \cdot dv}{\beta(l,m)(v+1)^{l+m}}$$

$$= \frac{e^{-u}}{} \Gamma(l+m) \cdot 1$$

$$f(v) = \beta_2(l,m) \cdot 1$$

07/09/2017

Q. $X \sim B(n_1, p)$, $Y \sim B(n_2, p)$

and X and Y are independent then

S.T. $X+Y \sim B(n_1+n_2, p)$

solution:

$$f_{x_i} = {}^{n_1}C_i\, p^i q^{n_1-i}, \quad i = 0,1,2,\dots n_1$$

$$f_{y_j} = {}^{n_2}C_j\, p^j q^{n_2-j}, \quad j = 0,1,2,\dots,n_2$$

$$f_{u_k} = P(X=i, Y=j)$$

$$k = 0,1,2,\dots,n_1+n_2$$

$$f_{u_k} = \sum_{i+j=k} f_{x_i} f_{y_j}$$

$$= \sum_{i+j=k} {}^{n_1}C_i\, p^i q^{n_1-i}\, {}^{n_2}C_j\, p^j q^{n_2-j}$$

$$= p^k q^{n_1+n_2-k} \sum_{i+j=k} {}^{n_1}C_i\, {}^{n_2}C_j$$

$$(1+x)^{n_1+n_2} = (1+x)^{n_1} \cdot (1+x)^{n_2}$$

${}^{n_2}C_j (x)^j$

kth coeff

$$\sum_{k=0}^{n_1+n_2} {}^{n_1+n_2}C_k (x)^k = \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} {}^{n_1}C_i (x^i)$$

$$\sum_{k=0}^{n_1+n_2} \sum_{i+j=k}^{n_2} {}^{n_1}C_i\, {}^{n_2}C_j (x)^k$$

$${}^{n_1+n_2}C_k = \sum_{i+j=k} {}^{n_1}C_i\, {}^{n_2}C_j$$

$= {}^{n_1+n_2}C_k\, p^k\, q^{n_1+n_2-k}$

Bivariant normal distribution

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-p^2}}\left[e^{-\frac{(x-m_x)^2}{\sigma_x^2}} - 2p\left(\frac{x-m_x}{\sigma_x}\right)\left(\frac{y-m_y}{\sigma_y}\right) + e^{-\left(\frac{x-m_y}{\sigma_y}\right)^2}\right]$$

$$e^{-\frac{1}{2}\left(\frac{x-m_x}{\sigma_x}\right)^2 - 2p\left(\frac{x-m_x}{\sigma_x}\right)\left(\frac{y-m_y}{\sigma_y}\right) + \frac{1}{2}\left(\frac{x-m_y}{\sigma_y}\right)^2}$$

$-\infty < x < y < \infty$

$f_x(x,y) = f_x(x) \cdot f_y(y)$

$f(x,y) = \frac{\partial^2 F}{\partial x \cdot \partial y}$

$F(x,y) = F(x) \cdot F(y)$

$\frac{\partial F}{\partial y} = F(x) F'(y)$

$\frac{\partial F}{\partial x} = F'(x) F(y)$

$\frac{\partial^2 F}{\partial x \partial y} = F'(x) F'(y)$

$f(x,y) = f(x) \cdot f(y)$

$$f_x(x) = \int_{y=-\infty}^{\infty} f(x,y)\, dy$$

$$= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{x'^2}{2} - 2\rho x' y' - \frac{y'^2}{2}} \sigma_y\, dy'$$

$\frac{x-m_x}{\sigma_x} = x'$

$$= \frac{e^{-x'^2/2}}{2\pi\sigma_x\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-1/2(y'^2 + 2\rho x' y')}\, dy'$$

$\frac{y-m_y}{\sigma_y} = y'$

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-p^2}}\left[e^{-\frac{1}{2(\sqrt{1-\rho^2})^2}\left[\left(\frac{x-m_x}{\sigma}\right)^2 - 2\rho\left(\frac{x-m_x}{\sigma}\right)\left(\frac{y-m_y}{\sigma}\right) + \left(\frac{y-m_y}{\sigma}\right)^2\right]}\right]$$

$$f_x(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$$

$\frac{x-m_x}{\sigma_x} = x'$, $\frac{y-m_y}{\sigma_y} = y'$

$$= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(\sqrt{1-\rho^2})^2}\left[x'^2 - 2\rho x' y' + y'^2\right]} \sigma_y\, dy'$$

$$= \frac{1}{2\pi\sigma_x\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(\sqrt{1-\rho^2})^2}[(y'-\rho x')^2+(1-\rho^2)x'^2]} dy'$$

$$= \frac{1}{2\pi\sigma_x\sqrt{1-\rho^2}} e^{-\frac{x'^2}{2}} \int_{-\infty}^{\infty} e^{-\left(\frac{y'-\rho x'}{\sqrt{2}\sqrt{1-\rho^2}}\right)^2} dy'$$

put $\frac{y'-\rho x'}{\sqrt{2}\sqrt{1-\rho^2}} = z$

$$dz = \frac{dy'}{\sqrt{2}\sqrt{1-\rho^2}}$$

$$= \frac{1}{2\pi\sigma_x\cancel{\sqrt{1-\rho^2}}} e^{-x'^2/2} \int_{-\infty}^{\infty} e^{-z^2} \sqrt{2}\, dz\, \cancel{\sqrt{1-\rho^2}}$$

$$= \frac{1}{\sqrt{2}\cdot\pi\sigma_x} e^{-x'^2/2} \times \sqrt{\pi}$$

$\rho \to$ density
$(-1$ to $1)$
(as $1-\rho^2>0$
$1>\rho^2$
$\rho^2<1$

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

$\left(\frac{y-m_y}{\sigma}\right)^2]$

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2}\left(\frac{x-m_x}{\sigma_x}\right)^2}$$

$$\int_{-\infty}^{\infty} f_x(x)\,dx = 1$$

$$\boxed{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1}$$

conditional marginal density function

- $f_x\left(\frac{x}{y}\right) = \frac{f(x;y)}{f_y(y)}$

- $f_y\left(\frac{x}{y}\right) \neq \quad f_y\left(\frac{y}{x}\right) = \frac{f(x,y)}{f_x(x)}$

- $P(a < x \leq b / Y = y) = \frac{\int_a^b f(x,y)dx}{f_y(y)}$

- $F_x\left(\frac{x}{y}\right) = P(-\infty < X \leq x / Y = y)$
$= \int_{-\infty}^{x} f_x\left(\frac{x}{y}\right) dx$

Theorem :

18/09/2017

χ^2- distribution

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

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

where n is the no. of degrees of freedom.

Theorem: If $X \sim \Gamma\left(\frac{n}{2}\right)$, then $2X$ represents $\chi^2(n)$-distribution.

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

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

Let $y = 2x$ as $0<x<\infty$
$\Rightarrow 0<y<\infty$

$$\frac{dy}{dx} = 2 > 0 \text{ (monotonic)}$$

$$f(y) = f(x)\left|\frac{dx}{dy}\right|$$

$$= \frac{1}{2}f(x)$$

$$= \frac{e^{-x}x^{n/2-1}}{2\,\Gamma\left(\frac{n}{2}\right)}$$

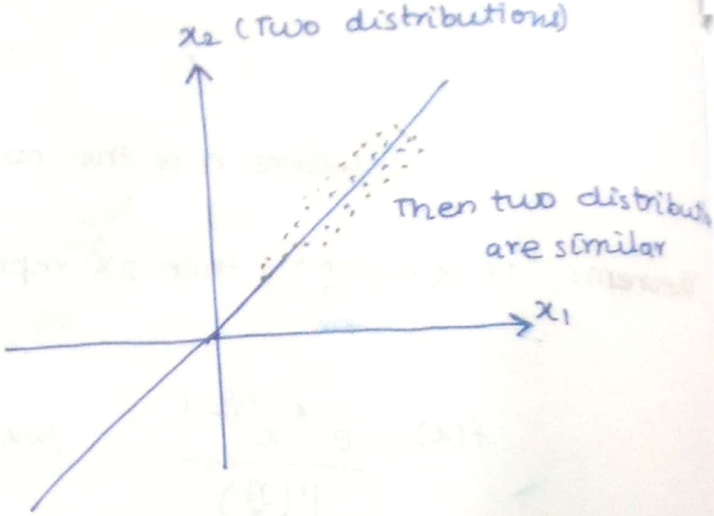
$$= \frac{e^{-y/2}\left(\frac{y}{2}\right)^{n/2-1}}{2\,\Gamma\left(\frac{n}{2}\right)}$$

Y represents $\chi^2(n)$

$$\boxed{Y \sim \chi^2(n)}$$

$$\int_0^\infty \frac{e^{-x/2}\left(\frac{x}{2}\right)^{\frac{n-1}{2}}}{2\,\Gamma\left(\frac{n}{2}\right)}\,dx$$

$$\frac{1}{2\,\Gamma\left(\frac{n}{2}\right)}\int_0^\infty e^{-x/2}\left(\frac{x}{2}\right)^{\frac{n-1}{2}}dx = 1$$



NOTE:

$$X \sim N(0,1)$$

$$\frac{1}{2}X^2 \sim \Gamma\left(\tfrac{1}{2}\right)$$

$$\frac{1}{2}X_1^2+\frac{1}{2}X_2^2 \dots + \frac{1}{2}X_n^2 \sim \Gamma\left(\tfrac{n}{2}\right)$$ $\left[{}_C\Gamma(m+n)\right]_{X+Y\text{dist}}$

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

$$\therefore\quad X_1^2+X_2^2 \dots + X_n^2 \sim \chi^2(n)$$

Pocchammes symbol $\alpha_k \longrightarrow k^{th}$ moment of Γ-distribution

$$= \frac{\Gamma(n+k)}{\Gamma(n)}$$

$$(\alpha)_n = \alpha(\alpha+1)\dots(\alpha+n-1)$$

For fractions α

$$= \frac{\Gamma(n+k)}{\Gamma(n)}$$

Scanned by CamScanner

$$\alpha_k(y) = E(Y^k)$$

$$= E((2X)^k)$$

$$= 2^k E(X^k)$$

$$= 2^k \frac{\Gamma(\frac{n}{2}+k)}{\Gamma(\frac{n}{2})}$$

$$= 2^k \left[\frac{n}{2}\left(\frac{n}{2}+1\right)\cdots\left(\frac{n}{2}+k-1\right)\right]$$

$$= n(n+2)(n+4)\cdots(n+2k-2)$$

$\alpha_1 = n$

$\alpha_2 = n(n+2)$

variance $= \alpha_2 - \alpha_1^2$

$= 2n$

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

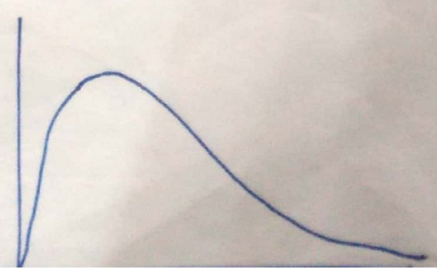
$$f'(x) = \frac{\frac{1}{2}e^{-x/2}\left(\frac{n}{2}-1\right)\left(\frac{x}{2}\right)^{n/2-2} + \left(\frac{x}{2}\right)^{\frac{n}{2}-1} e^{-x/2}\left(-\frac{1}{2}\right)}{2\Gamma(\frac{n}{2})}$$

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

$f'(x) = 0$

$\frac{1}{\cancel{2}}\left(\frac{n}{2}-1\right) = \frac{x}{\cancel{4}\,2}$ → Unimodal

$\boxed{n-2=x}$ → mode



Wed next 12th october $Y = 2X$

$\hookrightarrow (1-2it)^{-n/2}$

20/09/2017

$X \sim \Gamma(\frac{n}{2})$

$Y = 2X \sim \chi^2(n)$

$X \sim \chi^2(n)$

$\frac{X}{2} \sim \Gamma(\frac{n}{2})$

t- distribution
(Gosset)
Pseudo name → (student distribution)

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

continuous distribution

$n=1$

$$f(t) = \frac{1}{\sqrt{1}.\pi\,(1+t^2)^1}, \; -\infty < t < \infty$$

for $n=1 \rightarrow$ Cauchy Distribution
whose $\hookrightarrow$ does not exist mean, variance
~~A≥1~~
$n>1 \rightarrow$ mean, variance exist

Ex-1 If $X \sim N(0,1)$ and $\chi^2 \sim \chi^2(n)$, then

$\frac{X}{\sqrt{\chi^2(n)}}$ represents t distribution

Solution: Let $Y = \frac{X}{\sqrt{\chi^2/n}}$

$$\frac{Y^2}{n} = \frac{{}^{1/2}X^2}{{}^{1/2}\chi^2}$$

$$= \frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{n}{2})}$$

$$\Rightarrow \frac{Y^2}{n} \sim \beta_2\left(\frac{1}{2}, \frac{n}{2}\right) \qquad \left[\because \frac{X}{Y} \frac{\to \Gamma(m)}{\Gamma(n)} \quad \beta(m, n)\right]$$

The distribution is

$$dF = f\left(\frac{y^2}{n}\right) d\left(\frac{y^2}{n}\right) \qquad f(y)dy = f(x)dx$$

$$= \frac{\left(\frac{y^2}{n}\right)^{\frac{1}{2}-1}}{\left(1+\frac{y^2}{n}\right)^{\frac{1}{2}+\frac{n}{2}} \beta_2\left(\frac{1}{2}, \frac{n}{2}\right)} \, 2y \, \frac{dy}{n}$$

$$= \frac{2}{\sqrt{n}\, \beta_2\left(\frac{1}{2}, \frac{n}{2}\right)\left(1+\frac{y^2}{2}\right)^{\frac{1}{2}+\frac{n}{2}}} dy, \quad 0 < \frac{y^2}{n} < \infty$$

$$= \frac{1}{\sqrt{n}\, \beta_2\left(\frac{1}{2}, \frac{n}{2}\right)\left(1+\frac{y^2}{2}\right)^{\frac{1+n}{2}}} dy, \quad -\infty < y < \infty$$

Theorem: If $X \sim N(0,1)$, $\chi^2 = \chi^2(n)$, then $\frac{X}{\sqrt{\chi^2/n}}$ represents t distribution with (n) degrees of freedom.

$$t = \frac{\bar{X} - m}{s/\sqrt{n}}$$

$\downarrow$ small s

$$s^2 = \frac{n}{n-1} S^2$$

Next class

F - distribution

$$f(x) = \frac{m^{m/2}\, n^{n/2}\, x^{m/2-1}}{\beta(\frac{m}{2},\frac{n}{2})(mx+n)^{\frac{m}{2}+\frac{n}{2}}} \quad 0 < x < \infty$$

$= 0$, elsewhere

Q1. If $x_1^2 \sim x^2(m)$, $x_2^2 \sim x^2(n)$ then P.T. $\frac{x_1^2/m}{x_2^2/n}$ represents F(m, n) distribution.

Q2 If $X \sim F(m,n)$ then $\frac{1}{X} \sim F(n,m)$

Ans 1

$$Y = \frac{x_1^2/m}{x_2^2/n}$$

$$Y = \left(\frac{n}{m}\right)\frac{\frac{1}{2}x_1^2}{\frac{1}{2}x_2^2}$$

$$Y = \left(\frac{n}{m}\right)\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{n}{2})}$$

$$\frac{Y}{(\frac{n}{m})} = \cancel{\left(\frac{n}{m}\right)}\beta_2\left(\frac{m}{2},\frac{n}{2}\right)$$

$$\frac{Y}{(\frac{n}{m})} \sim \beta_2\left(\frac{m}{2},\frac{n}{2}\right)$$

$$dF = f\left(\frac{y}{\frac{n}{m}}\right) d\left(\frac{my}{n}\right)$$

$$= \frac{m}{n} f\left(\frac{my}{n}\right) dy$$

$$= \frac{m}{n}\frac{\left(\frac{my}{n}\right)^{\frac{m}{2}-1}}{\left(1+\frac{my}{n}\right)^{\frac{m}{2}+\frac{n}{2}}\beta_2\left(\frac{m}{2},\frac{n}{2}\right)} dy$$

Ans 2

Next class 12th October

$$= \frac{m}{n} \frac{\left(\frac{m}{n}\right)^{\frac{m}{2}-1} (y)^{\frac{m}{2}-1} \times n^{\frac{m+n}{2}}}{(n+my)^{\frac{m}{2}+\frac{n}{2}} \beta_2\left(\frac{m}{2},\frac{n}{2}\right)}$$

$$= \left(\frac{m}{n}\right)^{\frac{m}{2}} (n)^{\frac{m+n}{2}} \frac{(y)^{\frac{m}{2}-1}}{(n+my)^{\frac{m+n}{2}} \beta_2\left(\frac{m}{2},\frac{n}{2}\right)}$$

$$= \frac{m^{m/2} (n)^{n/2} (y)^{\frac{m}{2}-1}}{(n+my)^{\frac{m}{2}+\frac{n}{2}} \beta_2\left(\frac{m}{2},\frac{n}{2}\right)}$$

$\dfrac{x_1^2/m}{x_2^2/m}$ represents $F(m,n)$ distribution

Ans 2

$X \sim F(m,n)$

$y = \frac{1}{x}$

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

This is a monotonic function

$f(y) = f(x) \left| \frac{dx}{dy} \right|$

$$= x^2 f(x)$$

$$= x^2 \times \frac{m^{m/2} n^{n/2} (x)^{\frac{m}{2}-1}}{\beta\left(\frac{m}{2},\frac{n}{2}\right)(mx+n)^{\frac{m}{2}+\frac{n}{2}}}, \quad 0<x<\infty$$

$$= \frac{m^{m/2} n^{n/2} (x)^{\frac{m}{2}}}{\beta\left(\frac{m}{2},\frac{n}{2}\right)(mx+n)^{\frac{m+n}{2}}}, \quad 0<x<\infty$$

$$= \frac{m^{m/2} n^{n/2} \left(\frac{1}{y}\right)^{m/2}}{\beta\left(\frac{m}{2},\frac{n}{2}\right)\left(\frac{m}{y}+n\right)^{\frac{m+n}{2}}}, \quad 0<y<\infty$$

$$= \frac{m^{m/2} n^{n/2} (y)^{n/2-1}}{\beta\left(\frac{m}{2},\frac{n}{2}\right)(m+ny)^{\frac{m}{2}+\frac{n}{2}}}, \quad 0<y<\infty$$

$\therefore$ Y represents $F(n,m)$ distribution

12/10/2017

1. Tchebycheff's inequality

2. Convergence in probability

3. Tchebycheff's theorem

4. Binomial Theorem

5. Law of large number

Tchebycheff's inequality

If X be a random variable having a finite variance, then

$$P[\,|X-m| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

where m and σ^2 are mean and variance respectively.

Proof:

$$P(\,|X-m| \geq \epsilon) = \int_{|x-m|\geq \epsilon} f(x)\,dx$$

$$\frac{(x-m)^2}{\epsilon^2} \geq 1$$

$P(a<X<b) = \int_{x\in(a,b)} f(x).dx$ or $\int_{a<x<b} f(x).dx$

$$P(\,|X-m| \geq \epsilon) = \int_{|x-m|\geq \epsilon} 1.\,f(x)\,dx$$

$$\leq \int_{-\infty}^{\infty} \frac{(x-m)^2}{\epsilon^2} f(x)\,dx$$

$$\text{i.e.} \leq \frac{\sigma^2}{\epsilon^2}$$

If X,
and

This
mass

If
each
as r

Proof
NOTE:

NOTE: Th
en
th

Convergence in probability

If $X_1, X_2, \ldots X_n, \ldots$ be a sequence of random variables

and $\lim_{n \to \infty} P(|X_n - a| \geq \epsilon) = 0$

$$P(|X_n - a| < \epsilon) = 1$$

then X_n is said to converge to a constant a.

$$X_n \xrightarrow[\text{in } p]{} a \quad \text{as } n \to \infty.$$

This means that as n becomes large, the probability mass of X_n accumulates more and more to a.

Tchebycheff's Theorem

If $X_1, X_2, \ldots X_n, \ldots$ be the sequence of random variables each having mean m_n and variance σ_n and $\sigma_n \to 0$ as $n \to \infty$, then $X_n \xrightarrow[\text{in } p]{} m_n$ as $n \to \infty$

$$P(|X_n - m_n| \geq \epsilon) \to 0$$

Proof : By Tchebycheff's inequality,

$$P(|X_n - m_n| \geq \epsilon) \leq \frac{\sigma_n^2}{\epsilon^2}$$

~~NOTE~~ :

and as $\lim_{n \to \infty} P(|X_n - m_n| \geq \epsilon) = 0$

i.e. $P(|X_n - m_n| < \epsilon) = 1$

then X_n is said to converge to a constant m_n

$$X_n \xrightarrow[\text{in } p]{} m_n \quad \text{as } n \to \infty$$

NOTE: The importance of Tchebycheff's inequality is that they enable to derive bounds on probabilities when only the mean or both the mean and variance are known. But the distribution is not given. Valid for all distributions

Binomial Theorem

If $X_n \sim B(np, \sqrt{npq})$, then $\frac{X_n}{n} \to p$ as $n \to \infty$

$E(X_n) = np$ $\qquad \sigma(X_n) = \sqrt{npq}$

$$\sigma\left(\frac{X_n}{n}\right) = \sqrt{\frac{pq}{n}}$$

as $n \to \infty$
$\sigma\left(\frac{X_n}{n}\right) \to 0$

$E\left(\frac{X_n}{n}\right) = p$

$\frac{X_n}{n} \xrightarrow{\text{in } p} p$ as $n \to \infty$ (Tchebycheff's inequality)

Law of large number

If $X_1, X_2, \ldots X_n, \ldots$ be a sequence of random variables and

$S_n = X_1 + X_2 \cdots + X_n$ has mean m_n and standard deviation Σ_n then if $\Sigma_n = o(n)$

i.e. $\frac{\Sigma_n}{n} \to 0$ as $n \to \infty$, then

$$\frac{S_n - m_n}{n} \to 0 \text{ as } n \to \infty$$

Order of x

$o(x)$	$O(x)$
$\frac{o(x)}{x} \to 0$ as $x \to 0$	$\frac{O(x)}{x} \to$ finite as $x \to 0$
$(1-x)^{-1} = 1 + x + x^2 + \ldots$	$(1-x)^{-1} = 1 + x + x^2 + x^3 + \cdots$
$= 1 + x + o(x)$	$= 1 + x + O(x^2)$
also $o(x) = x^2 + x^3 + \ldots$	where $O(x^2) = x^2 + x^3 + \cdots$
$\frac{o(x)}{x} = \frac{x^2 + x^3 + \ldots}{x}$	$\frac{O(x^2)}{x^2} \to 1$ as $x \to 0$
$\frac{o(x)}{x} \to 0$ as $x \to 0$	

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

$X_1 \downarrow N(m_1, \Sigma_1)$, $X_2 \downarrow N(M_2, \Sigma_2)$, $X_n \downarrow N(m_n, \Sigma_n)$

case of equal components

$$S_n \sim N(nm, \sqrt{n}\Sigma) \qquad \frac{\sigma}{n} = \frac{\Sigma}{\sqrt{n}}$$

$\sqrt{n}\Sigma = \sigma$

$$\frac{S_n}{n} - \frac{M_n}{n} \to 0 \quad \text{as } n \to \infty$$

$$\bar{X} - m \to 0 \quad \text{as } n \to \infty$$

$$\bar{X} \to m \quad \text{as } n \to \infty$$

NOTE: only having a finite mean is sufficient. Σ/n is not necessary

Weak law of large numbers

Let $X_1, X_2, \dots X_n$ be a sequence of independent and identically distributed random variables each having finite mean $E(X_i) = \mu$

Then for any $\varepsilon > 0$

$$P\left\{ \left| \frac{X_1 + X_2 \dots + X_n}{n} - \mu \right| \geq \varepsilon \right\} \to 0 \text{ as } n \to \infty$$

Proof: Assumption each having finite variance σ^2

$$E\left(\frac{X_1 + X_2 \dots + X_n}{n} \right) = \mu$$

$$\text{var}\left(\frac{X_1 + X_2 \dots + X_n}{n} \right) = \frac{\sigma^2}{n}$$

By Chebyschev's inequality

$$P\left\{ \left| \frac{X_1 + X_2 \dots + X_n}{n} - \mu \right| \geq \varepsilon \right\} \leq \frac{\sigma^2}{n\varepsilon^2}$$

$\downarrow 0$ as $n \to \infty$

$$\cancel{\frac{X_1 + X_2 + \dots + X_n}{n} = \mu}$$

16/10/2017

Poisson distribution

Characteristic function $\chi_n(t) = e^{\mu(e^{it}-1)}$

Binomial distribution

Characteristic function $\chi_n(t) = (pe^{it}+q)^n$

$$\chi_n(t) = (pe^{it}+1-p)^n$$

$$= (1+p(-1+e^{it}))^n$$

Let $p = \frac{\mu}{n}$ as $n \to \infty$, $p \to 0$

$$= \left(1+\frac{\mu}{n}(e^{it}-1)\right)^n$$

as $n \to \infty$

$$= e^{\lim_{n\to\infty} \frac{\mu}{\cancel{n}}(e^{it}-1)\cancel{n}}$$

$$= e^{\mu(e^{it}-1)}$$

$\therefore$ Poisson process is a limiting case of binomial.

Characteristic function uniquely determine distribution function.

$$\left|\int_a^b f(x).dx\right| \le \int_a^b |f(x)|.dx$$

$$\left|\int_{-\infty}^{\infty} e^{itx} f(x).dx\right| \le \int_{-\infty}^{\infty} |e^{itx}||f(x)|\,dx = 1$$

So characteristic function doesn't pose the problem of convergency while m.g.f does.

If X_n

vari

an

Proof:

De Moivre-Laplace Theorem

If $X_n \sim B(n,p)$, then the corresponding standard normal variant is

$$X_n^* = \frac{X_n - np}{\sqrt{npq}} \sim N(0,1)$$

and for any two fixed numbers a and b : b>0,

$$\lim_{n\to\infty} P_n(a < X_n^* < b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2}\, dt$$

Proof: $X_n^* = \dfrac{X_n - np}{\sqrt{npq}}$ where $X_n \sim B(n,p)$

and characteristic function is $\phi_{X_n} = (pe^{it}+q)^n$

$$\phi_{X_n^*}(t) = E(e^{itX_n^*})$$

$$= E\left(e^{it\left(\frac{X_n - np}{\sqrt{npq}}\right)}\right)$$

$$= e^{\frac{-itnp}{\sqrt{npq}}} \{ E(e^{itX_n/\sqrt{npq}})$$

$$= e^{\frac{-itnp}{\sqrt{npq}}} (pe^{it/\sqrt{npq}} + q)^n$$

$$\boxed{\phi_{X_n^*}(t) = (pe^{itq/\sqrt{npq}} + qe^{-itp/\sqrt{npq}})^n}$$

$$pe^{itq/\sqrt{npq}} = p\left(1 + \frac{itq}{\sqrt{npq}} + \frac{(itq)^2}{(\sqrt{npq})^2}\frac{1}{2!} + \cdots\right)$$

$$= p\left(1 + \frac{itq}{\sqrt{npq}} - \frac{t^2q^2}{2(npq)} \cdots\right)$$

$$= p + \frac{it}{\sqrt{n}}\sqrt{pq} - \frac{t^2q}{2n} + O\left(\frac{1}{n}\right)$$

$$q e^{\frac{-itp}{\sqrt{npq}}} = q\left(1 - \frac{itp}{\sqrt{npq}} - \frac{p^2t^2}{2!(npq)} + o\left(\frac{1}{n}\right)\right)$$

$$= q - \frac{it\sqrt{pq}}{\sqrt{n}} - \frac{pt^2}{2!\,n} + o\left(\frac{1}{n}\right)$$

Adding

$$pe^{\frac{itq}{\sqrt{npq}}} + qe^{\frac{-itp}{\sqrt{npq}}} = p+q - \frac{t^2(p+q)}{2!\,n} + o\left(\frac{1}{n}\right)$$

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

$$\left(pe^{itq/\sqrt{npq}} + qe^{-itp/\sqrt{npq}}\right)^n = \left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n$$

as $n \longrightarrow \infty$

so $X_n^*(t) \longrightarrow e^{-\frac{t^2}{2}}$

$e^{-t^2/2}$ is the characteristic function of standard normal distribution.

$$X_n = X_n^*\sqrt{npq} + np$$

$$X_n \sim N(np, \sqrt{npq})$$

Central Limit Theorem (For equal components)

Let $X_1, X_2, \ldots X_n$ are the sequence of random variables having finite variance and $S_n = X_1 + X_2 \ldots + X_n$ with equal mean and variance, then

$$\bar{X} \sim N\left(m, \frac{\sigma}{\sqrt{n}}\right)$$

Proof: $S_n = X_1 + X_2 \quad \ldots \quad + X_n \sim N(nm, \sqrt{n}\sigma)$

$$\frac{S_n - nm}{\sqrt{n}\,\sigma} \sim N(0,1)$$

$$\frac{\frac{S_n}{n} - m}{\sigma/\sqrt{n}} = \frac{\bar{X} - m}{\sigma/\sqrt{n}} \sim N(0,1)$$

$\bar{X}$ = avg. of n random variables

so $\bar{X} \sim N(m, \sigma/\sqrt{n})$

(using De Moivre's Laplace (reverse))

Statistics

Let a sample $(x_1, x_2, \ldots x_n)$ of size n is chosen from the populatio[n] of a random variable X

$y = y(x)$
$Y = y(X)$

Any function of $(x_1, x_2, \ldots x_n)$ is called statistic

$a = a(x_1, x_2, \ldots x_n)$ is a statistic which is the observed values of the random variable $A = a(X_1, X_2, \ldots X_n)$

mean, variance etc. are statistics.

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$A_K = \frac{1}{n}\sum_{i=1}^{n} X_i^K$$

$$M_K = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^K$$

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

	Sample	Population
mean	$\bar{X}$	m or μ
K^{th} moment about mean	A_K	α_K
	M_K	μ_K
variance	S^2	σ^2

The s[...]
popul[...]
The [...]
As [...]
esti[...]
The [...]
of [...]
NOTE: T[...]

The statistic $a = a(x_1, x_2, \dots x_n)$ is consistent estimate of a population parameter if

$$A_K \xrightarrow[\text{in } P]{} \alpha_K \quad \text{as} \quad n \to \infty$$

The precision increases with the size of the sample.

As the sample size become large, the accuracy of the estimate increases.

The statistic $a = a(x_1, x_2, \dots x_n)$ is an unbiased estimate of the population parameter if

$$E(A_K^{(*)}) = \alpha_K$$

If $E(A_K) \neq \alpha_K$, the estimate is biased. It is negatively or positively biased if $E(A_K) - \alpha_K < 0 > 0$ resp.

NOTE: The estimate is said to be good estimate if it is both consistent and unbiased.

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

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i - m - (\bar{X} - m))^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2 - 2(\bar{X} - m)\frac{1}{n}\sum_{i=1}^{n}(X_i - m) + (\bar{X} - m)^2 \frac{1}{\cancel{n}}\cancel{n}$$

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2 - 2(\bar{X} - m)^2 + (\bar{X} - m)^2$$

$$\left[\because \frac{1}{n}\sum_{i=1}^{n}(X_i - m) = \bar{X} - m\right]$$

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2 - (\bar{X} - m)^2$$

$$E(S^2) = \frac{1}{n}\sum_{i=1}^{n}E(X_i - m)^2 - E(\bar{X} - m)^2$$

$$= \sigma^2 - \sigma^2/n$$

$E(Y-m)^2 = \sigma^2$

$E(\bar{X}) = m$

$E(\bar{X}-m)^2 = \frac{\sigma^2}{n}$

$\bar{X} \sim N(m, \frac{\sigma}{\sqrt{n}})$

$E(\bar{X}-m)^2 = \frac{\sigma^2}{n}$

S^2 is negatively biased $(\cancel{\sigma^2} - \frac{\sigma^2}{n} - \cancel{\sigma^2})$

$E(S^2) - \sigma^2 < 0$

as $n \to \infty$

~~$E(S^2)$~~

$S^2 \xrightarrow[\text{in } p]{} \sigma^2$

Consistent.

$\therefore$ Sample variance is consistent but biased estimate to ~~with~~ population variance

$E(S^2) = \frac{n-1}{n}\sigma^2$

$E\left(\frac{nS^2}{n-1}\right) = \sigma^2$

$E(s^2) = \sigma^2$

as $n \to \infty$

$s^2 \longrightarrow \sigma^2$

$\downarrow$

it is both consistent and unbiased.

$\therefore$ it is a good estimate

$$\boxed{s^2 = \frac{n}{n-1} S^2}$$

25 October, 2017

$A_k \xrightarrow[\text{in } P]{} \alpha_k$ as $n \to \infty$

↳ convergence in probability

$\lim_{x \to a} f(x) = l$

$|f(x) - l| < \varepsilon$ whenever $|x - a| < \delta(\varepsilon)$

$\lim_{x \to \infty} f(x) = l$

$|f(x) - l| < \varepsilon$ whenever $n \geq m$, $m \in \mathbb{N}$

$X_i \rightarrow$ sample $\rightarrow$ population

$\bar{X}, S^2$ m, σ^2

$E(X_i) = \bar{X}$ $E(X_i) = m$

$E(X_i - \bar{X})^2 = S^2$ $E(X_i - m)^2 = \sigma^2$

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

$$= \frac{1}{n}\sum_{i=1}^{n}\left((X_i - m)^2 - (\bar{X} - m)\right)^2$$

$$\frac{nS^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i - m}{\sigma}\right)^2 - \frac{n}{\sigma^2}\left(\frac{\bar{X} - m}{\sigma}\right)^2$$

$$= \sum_{i=1}^{n}\left(\frac{X_i - m}{\sigma}\right)^2 - \frac{n}{\sigma^2}\left(\frac{X_1 + X_2 \cdots + X_n}{n} - m\right)^2$$

$$= \sum_{i=1}^{n}\left(\frac{X_i - m}{\sigma}\right)^2 - \frac{1}{\sqrt{n}}\left(\left(\frac{X_1 - m}{\sigma}\right) + \left(\frac{X_2 - m}{\sigma}\right) + \cdots \left(\frac{X_n - m}{\sigma}\right)\right.$$

$$= \sum_{i=1}^{n}\left(\frac{X_i - m}{\sigma}\right)^2 - \left\{\frac{1}{\sqrt{n}}\left(\frac{X_1 - m}{\sigma}\right) - \cdots \frac{1}{\sqrt{n}}\left(\frac{X_n - m}{\sigma}\right)\right\}^2$$

$$\left[\begin{array}{l} x_i \sim N(m,\sigma) \\ \dfrac{x_i - m}{\sigma} \sim N(0,1) \end{array}\right]$$

$$= Y_1^2 + Y_2^2 + \dots + Y_n^2 - \left\{ \frac{1}{\sqrt{n}} Y_1 + \frac{1}{\sqrt{n}} Y_2 + \dots + \frac{1}{\sqrt{n}} Y_n \right\}^2$$

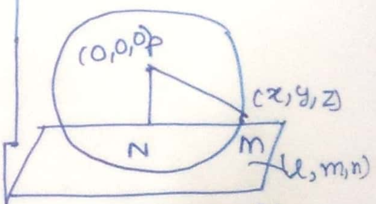
$$\boxed{\frac{ns^2}{\sigma^2} \sim \chi^2(n-1)}$$

$Y_i \sim N(0,1)$

$\frac{1}{2} Y_i^2 \sim \Gamma(\frac{1}{2})$

$\frac{1}{2} Y_1^2 + \frac{1}{2} Y_2^2 \dots + \frac{1}{2} Y_n^2 \sim \Gamma(\frac{n}{2})$

$Y_1^2 + \dots + Y_n^2 \sim \chi^2(n)$



$PN^2 = (x^2+y^2+z^2) - (lx+my+nz)^2$

PN dof = 2

$lx + my + nz$ → plane

$l^2 + m^2 + n^2 = 1$

$l_1 Y_1 + l_2 Y_2 + \dots + l_n Y_n$

$l_1^2 + l_2^2 + \dots + l_n^2 = 1$

26/10/2017

(i) $z = \frac{\bar{X}-m}{\sigma/\sqrt{n}} \sim N(0,1)$ (central limit theorem for equal components)

(ii) $\chi^2_{(n-1)} = \frac{nS^2}{\sigma^2}$

(no. of degrees of freedom)

(iii) $t = \left(\frac{\bar{X}-m}{s/\sqrt{n}}\right)$ with $n-1$ degrees of freedom

(sample is a part of population)

$$t = \frac{\bar{X}-m}{s/\sqrt{n}}$$

$$s^2 = \frac{n}{n-1}S^2 = \frac{n}{n-1} \times \frac{\chi^2\sigma^2}{n} \quad \text{or} \quad s = \frac{\chi\,\sigma}{\sqrt{n-1}}$$

$\frac{s}{\sigma} = \frac{\chi}{\sqrt{\nu}}$ where $\nu = n-1$

$$t = \frac{\bar{X}-m}{s/\sqrt{n}}$$

$$= \frac{\frac{\bar{X}-m}{\sigma/\sqrt{n}}}{\frac{s}{\sigma}}$$

$$= \frac{\frac{\bar{X}-m}{\sigma/\sqrt{n}}}{\sqrt{\frac{\chi^2}{\nu}}}$$

$\frac{\bar{X}-m}{\sigma/\sqrt{n}} \sim N(0,1)$

$$\cancel{\frac{\bar{X}-m}{\sigma/\sqrt{n}}}$$

$\frac{\bar{X}-m}{s/\sqrt{n}}$ is a t-distribution with $\nu = n-1$ degrees of freedom.

(using theorem If $X \sim N(0,1)$, $\chi^2 = \chi^2(n)$ then $\frac{X}{\sqrt{\chi^2/n}}$ represents t distribution with n degrees of freedom).

30/10/2017

Estimation of parameters

Maximum likelihood function

Suppose the distribution function of a random variable X has no functional form that contains a no. of unknown parameters.

Every functional is function that maps from ~~which maps~~ vector ~~field~~ space to underlying scalar field

~~underlying~~

Our aim will be to find a estimates of these parameters on the basis of sample $x_1, x_2, \ldots x_n$ drawn from population. There are several methods by which such estimations can be done, of which one of the methods is maximum likelihood

$$L(x_1, x_2 \ldots x_n; \theta_1, \theta_2 \ldots \theta_n)$$

Let X denotes a random variable such that

$$P(X = x_i) = f_{x_i}(\theta_1, \theta_2, \ldots \theta_n)$$

$f(x) = (x-2)^{2/3}$

$f'(x) = \frac{2}{3(x-2)^{1/3}}$

The event that the sample $x_1, x_2, \ldots x_n$ is drawn $(X_1 = x_1, X_2 = x_2, \ldots X_n = x_n)$

The probability of this event which is a function of sample values and unknown parameters is defined by the maximum likelihood function

$$L(x_1, x_2, \ldots x_n; \theta_1, \ldots \theta_n)$$
$$= P(X_1 = x_1, X_2 = x_2, \ldots X_n = x_n)$$

If $x_1, x_2, \ldots x_n$ are linearly independent then

$$L(x_1, x_2 \ldots x_n; \theta_1, \theta_2, \ldots \theta_n) = f_{x_1}(\theta_1, \theta_2, \ldots \theta_n) \cdot f_{x_2}(\theta_1, \theta_2, \ldots \theta_n) \ldots f_{x_n}(\theta_1, \theta_2, \ldots \theta_n)$$

$\because L>0$, maximising L amounts to maximising $\log L$

$$\frac{\partial}{\partial \theta_1}(\log L) = 0,$$

$$\frac{\partial}{\partial \theta_2}(\log L) = 0, \; \ldots \; \frac{\partial}{\partial \theta_n}(\log L) = 0$$

which are called likelihood equation by measuring this we get max likelihoodness of $\theta_1, \theta_2, \ldots \theta_n$ provided they exist.

$$\hat{\theta}_1 = \theta_1(x_1, x_2, \ldots x_n)$$

$$\hat{\theta}_2 = \theta_2(x_1, x_2, \ldots x_n)$$

$$\vdots$$

$$\hat{\theta}_n = \theta_n(x_1, x_2, \ldots x_n)$$

size ~~sample~~ of population

$$P(X=x_i) = {}^N C_{x_i}\, p^{x_i} (1-p)^{N-x_i}, \quad i = 1,2,3,\ldots n$$

$$L(x_1, x_2, \ldots x_n; p) = f_{x_1}(p)\, f_{x_2}(p) \ldots f_{x_n}(p)$$

$$= \left({}^N C_{x_1}\, {}^N C_{x_2} \ldots {}^N C_{x_n}\right) (p)^{\sum_{i=1}^{n} x_i} (1-p)^{nN - \sum_{i=1}^{n} x_i}$$

c is independent of parameter

$$\log L = \log c + \sum_{i=1}^{n} x_i \log p + \left(nN - \sum_{i=1}^{n} x_i\right) \log(1-p)$$

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

$$\Rightarrow \sum_{i=1}^{n} x_i \frac{1}{p} - \frac{\left(nN - \sum_{i=1}^{n} x_i\right)}{1-p}$$

$$\sum_{i=1}^{n} x_i = pnN$$

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

$$\boxed{\hat{p} = \frac{\bar{x}}{N}} \quad \therefore \text{ estimation of } p \text{ is } \hat{p}.$$

Scanned by CamScanner

$\therefore \bar{x}$ is consistent and unbiased estimator of population mean.

$\hat{p} =$

H.W. Poisson

Normal distribution. $P(X=x_i)\, f(x_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-m)^2}{2\sigma^2}}$

$L(x_1, x_2, \ldots x_n; m, \sigma)$

$= f(x_1)\, f_{x_1}(m,\sigma), f_{x_2}(m,\sigma) \ldots, f_{x_n}(m,\sigma)$

$$= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_1-m)^2}{2\sigma^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_2-m)^2}{2\sigma^2}} \cdots \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_n-m)^2}{2\sigma^2}}$$

$$= \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\left\{\frac{(x_1-m)^2+(x_2-m)^2 \ldots +(x_n-m)^2}{2\sigma^2}\right\}}$$

$$L(x_1, x_2, \ldots x_n; m, \sigma) = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\sum_{i=1}^{n} \frac{(x_i-m)^2}{2\sigma^2}}$$

$\log L = \cancel{\log(c) - \Sigma}$

$$L = (2\pi)^{-n/2} \sigma^{-n} e^{-\sum_{i=1}^{n} \frac{(x_i-m)^2}{2\sigma^2}}$$

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

$$\frac{\partial L}{\partial m} = \cancel{\frac{2}{2\sigma^2}\sum_{i=1}^{m}}(x_i - m)$$

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

$$\log L = -\left\{\frac{(x_1-m)^2}{2\sigma^2} + \frac{(x_2-m)^2}{2\sigma^2} \ldots + \frac{(x_n-m)^2}{2\sigma^2}\right\} - \frac{n}{2}\log(2\pi)$$

$$\frac{1}{L}\frac{\partial L}{\partial m} = +\left\{\frac{2(x_1-m)}{2\sigma^2} + \frac{2(x_2-m)}{2\sigma^2} \ldots + \frac{2(x_n}{2\sigma}\right.$$

$$\frac{1}{L}\frac{\partial L}{\partial m} = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} x_i - nm\right)$$

Tomorrow no class

$\bar{x} = m$

$\frac{1}{L}\frac{\partial L}{\cancel{\partial \sigma}} \cancel{=} \left((x_1 - m)^2 \left(\frac{-2}{\sigma^3}\right) + \right.$

$$\frac{1}{L}\frac{\partial L}{\partial \sigma} = -\left\{ \frac{(x_1-m)^2}{2} \times \frac{(-2)}{\sigma^3} + \frac{(x_2-m)^2}{2}\frac{(-2)}{\sigma^3} \cdots \frac{(x_n-m)^2}{2}\frac{(-2)}{\sigma^3} \right\} - \frac{n}{\sigma}$$

$$= \cancel{\frac{(x_1-m)^2}{}} \quad \sum_{i=1}^{n} \frac{(x_i-m)^2}{\sigma^2} = \frac{n}{\sigma}$$

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

$$\hat{\sigma}^2 = s^2$$

This is not a good estimator

Tomorrow
no class

$-\frac{n}{2}\log(2\pi) - n\log\sigma$

$\cdots + \frac{2(x_n - m)}{2\sigma^2}\}$

In general any numerical characteristic of a population is called a parameter and any quantity computed from a random sample is called a statistic. Statistics are used to estimate parameters.

02 November, 2017

Interval estimation

Let α be a population parameter and ϵ be a given number $(0<\epsilon<1)$

If there exists two statistics

$a = a(x_1, x_2, \dots x_n)$

$b = b(x_1, x_2, \dots x_n)$

such that

$P(A < \alpha < B) = 1-\epsilon$

$A = a(X_1, X_2, \dots X_n)$, $B = b(X_1, X_2, \dots, X_n)$

are the random variables corresponding to statistics a & b resp.

Then interval (a, b) is called an interval estimate or confidence interval

For the parameter α with confidence coeff $1-\epsilon$.

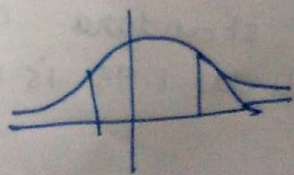
The statistics a & b are resp. the lower and upper confidence limits of α.

NOTE : The probability that Interval (a, b) covers fixed point α is $1-\epsilon$.

A practical interpretation is that if a long sequence of random samples is drawn, under uniform conditions. statistics a and b are computed each time, then the ratio of the number of times the interval (a, b) includes the population parameter, value α to the number of samples drawn is approx. equals to $1-\epsilon$.

If $\epsilon = 0.05$, then the confidence coeff. is 0.95.

The confidence interval is 95%



Application to Normal (m, σ) population

Confidence interval for m

Ans

Three statistics

$\bar{X}=$

$U = \dfrac{\bar{X}-m}{\sigma/\sqrt{n}}$

$t = \dfrac{\bar{X}-m}{S/\sqrt{n}}$

$\chi^2 = \dfrac{n s^2}{\sigma^2} \rightarrow$ smalls

Case 1: σ known

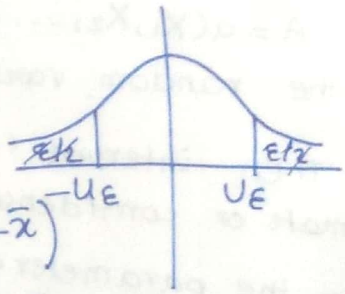
Let us consider the statistic

$$U = \frac{\bar{X}-m}{\sigma/\sqrt{n}}$$

Let $P(-U_\varepsilon < u < U_\varepsilon) = 1-\varepsilon$

$$P\left(-U_\varepsilon < \frac{\bar{X}-m}{\sigma/\sqrt{n}} < U_\varepsilon\right) = 1-\varepsilon$$

$$P\left(-\frac{U_\varepsilon \sigma}{\sqrt{n}} < \bar{X} - m < U_\varepsilon \right)$$



$$P\left(-\frac{U_\varepsilon \sigma}{\sqrt{n}} - \bar{X} < -m < \frac{U_\varepsilon \sigma}{\sqrt{n}} - \bar{x}\right) = 1-\varepsilon$$

$$P\left(\bar{X} - \frac{U_\varepsilon \sigma}{\sqrt{n}} < m < \bar{X} + \frac{U_\varepsilon \sigma}{\sqrt{n}}\right) = 1-\varepsilon$$

Confidence interval for m is $\left(\bar{x} - \frac{U_\varepsilon \sigma}{\sqrt{n}}, \bar{x} + \frac{U_\varepsilon \sigma}{\sqrt{n}}\right)$

$P(u > U_\varepsilon) = \varepsilon/2$

Q. The variable X is a normal distributed with mean 68 cm and $\sigma = 2.5$ cm. What should be size of sample whose mean shall not differ from the population mean by more than 1cm with probability 0.95 (Given that the area under standard normal curve to the right of the ordinate 1.96 is 0.025 cm^2.

Ans $m = 68cm, \sigma = 2.5cm$

$$U = \frac{\bar{x}-m}{\sigma/\sqrt{n}}$$

$$P(|\bar{x}-m| < 1) = 1-\varepsilon$$

$$P(-1 \leq \bar{x}-m \leq 1) = 1-\varepsilon$$

$$P\left(\frac{-1}{\sigma/\sqrt{n}} < \frac{\bar{x}-m}{\sigma/\sqrt{n}} < \frac{1}{\sigma/\sqrt{n}}\right) = 1-\varepsilon$$

$$P\left(\left|\frac{\bar{x}-m}{\sigma/\sqrt{n}}\right| \leq \frac{\sqrt{n}}{\sigma}\right) = 1-\varepsilon$$

$$P\left(\underbrace{\cancel{\ }}_{|U|} \leq \frac{\sqrt{n}}{\sigma}\right) = 1-\varepsilon = 0.95$$

~~P(|U≤)~~ = 0.95

$$P\left(|U| > \frac{\sqrt{n}}{\sigma}\right) = 0.05$$

$$P\left(U > \frac{\sqrt{n}}{\sigma}\right) = 0.025$$

$$P(U > 1.96) = 0.025$$

$$\frac{\sqrt{n}}{\sigma} = 1.96$$

$$\frac{\sqrt{n}}{2.5} = 1.96$$

$$n = (1.96 \times 2.5)^2$$

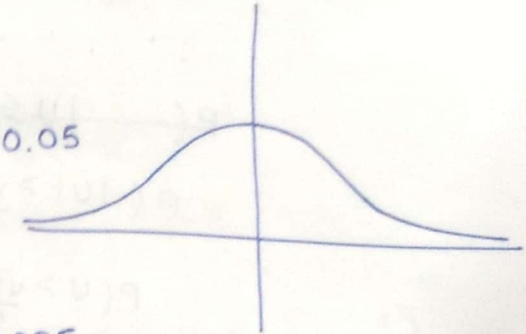
Scanned by CamScanner

06/11/2017

Q. The population of scores of 10 yr. children in a test is known to have standard deviation 5.2. If a random sample of size 20, chose with mean $\bar{X} = 16.9$,

find 95% confidence interval (limits) for the mean score of the population assuming that population mean $(m = ?)$ $\frac{\bar{X}-m}{\sigma/\sqrt{n}}$ is normal. given that $\frac{1}{\sqrt{2\pi}}\int_{1.96}^{\infty} e^{-t^2/2}\,dt = 0.025$

$P(-u_\varepsilon < U < u_\varepsilon) = 0.95$



$P(|U| > u_\varepsilon) = \cancel{0.025}\ 0.05$

$P(U > u_\varepsilon) = 0.025$

$\frac{1}{\sqrt{2\pi}}\int_{u_\varepsilon}^{\infty} e^{-t^2/2}\,dt = 0.025$

$u_\varepsilon = 1.96$

$P(-1.96 < U < 1.96) = 0.95$

$P\left(-1.96 < \frac{\bar{X}-m}{\sigma/\sqrt{n}} < 1.96\right) = 0.95$

$P\left(-1.96 \times \frac{\sigma}{\sqrt{n}} < \bar{X}-m < 1.96 \times \frac{\sigma}{\sqrt{n}}\right) = 0.95$

$P\left(-1.96 \times \frac{\sigma}{\sqrt{n}} - \bar{X} < -m < 1.96 \times \frac{\sigma}{\sqrt{n}} - \bar{X}\right) = 0.95$

$P\left(\bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}} < m < \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}}\right) = 0.95$

[14.62, 19.17]

Q. In a
be defe
true
of suc
F

Q. In a random sample of 400 articles, 40 are found to be defective. Obtain 95% confidence interval ~~for the~~ for the true population of the defectives in the population of such sample. Given that $\frac{1}{\sqrt{2\pi}}\int_0^{1.96} e^{-t^2/2}\,dt = 0.4750$

$p \rightarrow$ true value of the sample

$P(-u_\epsilon < U < u_\epsilon) = 0.95$

(confidence interval for proportions)

Q

If population is defined by the density function

$$f(x, \alpha) = \frac{x^{l-1} e^{-x/\alpha}}{\Gamma(l)\alpha^{l}}, \quad 0 < x < \infty$$

where l $(l>0)$ is a constant

so that approx confidence interval for $\alpha > 0$ (when sample size n is large) with confidence coeff is $1-\varepsilon$), $0<\varepsilon<1$

is $\left(\frac{\bar{x}}{l} - \frac{u_\varepsilon}{\sqrt{nl}} \times \frac{\bar{x}}{l}, \ \frac{\bar{x}}{l} + \frac{u_\varepsilon}{\sqrt{nl}} \frac{\bar{x}}{l}\right)$

Given that $\frac{1}{\sqrt{2\pi}} \int_0^{u_\varepsilon} e^{-t^2/2} \, dt = \varepsilon/2$

$$m = E(X) = \int x f(x) \, dx$$

Q. The marks

have a m

limit for

assuming

Let the

P

P(

Q. The marks obtained by 17 $\xrightarrow{\text{sample size}}$ candidates in an examination have a mean ~~57~~ and variance 64. Find 99% confidence limit for the mean of the population of the marks assuming it to be normal. Given that $\frac{1}{\sqrt{2\pi}}\int_{2.92}^{\infty} e^{-t^2/2}\,dt = 0.005$

$\bar{x} = 57$ $\quad \boxed{S^2 = 64}$

Let the fixed no. be chosen as t_ε

$$P(-t_\varepsilon < T < t_\varepsilon) = 1-\varepsilon$$

$$P(-t_\varepsilon < T < t_\varepsilon) = 0.99$$

$$P(T > t_\varepsilon) = 0.005$$

$$t_\varepsilon = 2.92$$

$$P\left(-t_\varepsilon < \frac{\bar{X}-m}{S/\sqrt{n}} < t_\varepsilon\right) = 0.99$$

$$P\left(-2.92 < \frac{\bar{X}-m}{S/\sqrt{n}} < 2.92\right) = 0.99$$

$$P\left(-2.92 \times \frac{S}{\sqrt{n}} < \bar{X}-m < 2.92 \times \frac{S}{\sqrt{n}}\right) = 0.99$$

$$P\left(-2.92 \times \frac{S}{\sqrt{n}} - \bar{X} < -m < 2.92 \times \frac{S}{\sqrt{n}} - \bar{X}\right) = 0.99$$

$$P\left(\bar{X} - 2.92 \times \frac{S}{\sqrt{n}} < m < \bar{X} + 2.92 \times \frac{S}{\sqrt{n}}\right) = 0.99$$

$$\bar{X} - 2.92 \times \frac{S}{\sqrt{n}}, \ \bar{X} + 2.92 \times \frac{S}{\sqrt{n}}$$

$$s^2 = \frac{n}{n-1} S^2$$

$$\left(57 - 2.92 \times \frac{\cancel{8}}{\sqrt{17}}, \ 57 + 2.92 \times \frac{\cancel{8}}{\sqrt{17}} \right)$$

$$s^2 = \frac{17}{\cancel{16}} \times \cancel{64}^{\,4}$$

$$s = \cancel{8}^{\,2} \sqrt{17}$$

$$\left(57 - 2.92 \times \frac{2\sqrt{17}}{\sqrt{17}}, \ 57 + 2.92 \times \frac{2\sqrt{17}}{\sqrt{19}} \right)$$

$$(51.16, \ 62.84)$$

Scanned by CamScanner

CORRELATION & REGRESSION

Correlation

$$\rho_{xy} = \frac{\overset{\text{covariance}}{cov(X,Y)}}{\sigma_x \sigma_y}$$

$$cov(X,Y) = E((X-\bar{X})(Y-\bar{Y}))$$

variance is a particular case of covariance if $X = Y$

$$= E(XY - X\bar{Y} - \bar{X}Y + \bar{X}\bar{Y})$$

$$= E(XY) - \bar{Y}E(X) - \bar{X}E(Y) + \bar{X}\bar{Y}$$

$$= E(XY) - \bar{X}\bar{Y} - \cancel{\bar{X}\bar{Y}} + \cancel{\bar{X}\bar{Y}}$$

$$= E(XY) - E(X)E(Y)$$

$cov(X,Y)$

$$\hookrightarrow \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$$

$$= \frac{1}{n}\sum x_i y_i - \bar{x}\bar{y}$$

Q. If (x,y) and (u,v) are two sets of bivariate data such that $u = ax+b$, $v = cy+d$

(Remember formula) Then P.T. $\rho_{uv} = \frac{ac}{|a||c|}\rho_{xy}$

Solution:

$u = ax+b$, $v = \cancel{cx}\, cy+d$

$$u_i = ax_i + b$$

$$\sum_{i=1}^{n} u_i = a\sum x_i + nb$$

$$\frac{1}{n}\sum_{i=1}^{n} u_i = a\frac{1}{n}\sum x_i + b$$

$$\bar{u} = a\bar{x} + b$$

Similarly $\bar{v} = c\bar{y} + d$

$$u_i - \bar{u} = a(x_i - \bar{x})$$

$$v_i - \bar{v} = c(y_i - \bar{y})$$

$$\rho_{uv} = \frac{cov(U,V)}{\sigma_u \sigma_v}$$

$$cov(U,V) = \frac{1}{n}\sum_{i=1}^{n}(u_i - \bar{u})(v_i - \bar{v})$$

$$= \frac{1}{n}\sum_{i=1}^{n} a(x_i - \bar{x})\, c(y_i - \bar{y})$$

$$= ac\ cov(X,Y)$$

$$\sigma_u^2 = \frac{1}{n}\sum (u_i - \bar{u})^2$$

$$= \frac{1}{n} \times a^2 \sum (x_i - \bar{x})^2$$

$$\sigma_u^2 = a^2 \sigma_x^2$$

$$\sigma_v^2 = c^2 \sigma_y^2$$

$$= \frac{ac\ cov(X,Y)}{|a|\sigma_x |c| \sigma_y}$$

Q. $var(X+Y) = \sigma_x^2 + \sigma_y^2 + 2\sigma_x\sigma_y \rho_{xy}$

$var(X-Y) = \sigma_x^2 + \sigma_y^2 - 2\sigma_x\sigma_y$

$$var(X+Y) = \frac{1}{n}\sum_{i=1}^{n}\left((x_i + y_i) - (\bar{x} + \bar{y})\right)^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left((x_i - \bar{x}) + (y_i - \bar{y})\right)^2$$

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

$$= \sigma_x^2 + \sigma_y^2 + 2\,cov(X,Y)$$

$$= \sigma_x^2 + \sigma_y^2 + 2\sigma_x\sigma_y\rho_{xy}$$

Theorem 3 :

$$u = \frac{x - \bar{x}}{\sigma_x} \qquad v = \frac{y - \bar{y}}{\sigma_y}$$

$$u_i = \frac{x_i - \bar{x}}{\sigma_x}$$

$$\frac{1}{n}\sum u_i = \frac{1}{n}\frac{1}{\sigma_x}\sum_{i=1}^{n}(x_i - \bar{x})$$

$$= 0$$

$$\bar{u} = 0 = \bar{v}$$

$$\sigma_u^2 = \frac{1}{n}\sum_{i=1}^{n}(u_i^2) - \bar{u}^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}\frac{(x_i - \bar{x})^2}{\sigma_x^2} - 0$$

$$= 1$$

$$\boxed{\sigma_u^2 = 1}$$

$$(u_i + v_i)^2 \geq 0$$

$$u_i^2 + v_i^2 + 2u_iv_i \geq 0$$

$$\frac{1}{n}\sum u_i^2 + \frac{1}{n}\sum v_i^2 + 2\frac{1}{n}\sum u_iv_i \geq 0$$

$$1 + 1 + 2\rho_{xy} \geq 0$$

$$\rho_{xy} \geq -1$$

$$(u_i - v_i)^2 \geq 0$$

$$\rho_{xy} \leq 1$$

$$-1 \leq \rho_{xy} \leq 1$$

If ρ_{xy} +ve $\rightarrow$ positively correlated

$\rho_{xy} = 0$ (no correlation)

Regression

From the bivariate data given by (x,y), the estimation or prediction of avg value of the variable say y corresponding to a specified value of the other variable x is called regression. The avg value of x may also be predicted corresponding to a specified value of y.

The regression line which is fixed and used to predict the avg value of y depending on x is called the regression line of y on x.

The regression line which is fixed and used to predict the avg value of x depending on y is called the regression line of x on y.

$$y = ax + b$$

$$y_i = ax_i + b$$

$$\bar{y} = a\bar{x} + b$$

$$\cancel{y - y_i = a(x_i - \bar{x})}$$

$$y - \bar{y} = a(x - \bar{x})$$

$$y_i = ax_i + b$$

$$x_i y_i = ax_i^2 + bx_i$$

$$\frac{1}{n}\sum_{i=1}^{n} x_i y_i = a\frac{1}{n}\sum x_i^2 + \frac{b}{n}\sum_{i=1}^{n} x_i$$

$$= a\frac{1}{n}\sum x_i^2 + (\bar{y} - a\bar{x})\bar{x}$$

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

$$a = \frac{\frac{1}{n}\sum x_i y_i - \bar{x}\bar{y}}{\frac{1}{n}\sum x_i^2 - \bar{x}^2} = \frac{cov(X,Y)}{\sigma_x^2} = \frac{\sigma_x \sigma_y \rho_{xy}}{\sigma_x^2} = \frac{\sigma_y}{\sigma_x}\rho_{xy}$$

$$y - \bar{y} = \rho_{xy}\frac{\sigma_y}{\sigma_x}(x - \bar{x})$$