MSE

DATE : / /

PAGE NO. :

03/09/19

Continuous Probability distribution

$f(x) \rightarrow$ pdf

iff:

(i.) $f(x) \geq 0 \quad \forall x$

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

$\Sigma f(x) = 1 \rightarrow$ Discrete

(iii.) $\int_{a}^{b} f(x)\, dx = P(a < x < b)$

→ Cumulative Distribⁿ Funcⁿ (cdf)

$F(x) = P(X \leq x)$

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

Properties of $F(x)$

(i.) Domain is $(-\infty, \infty)$

Range $[0, 1]$

(ii) $F(x)$ is non-↓ng funcⁿ of x in the right

ie, $F'(x) = f(x) \quad \forall \quad x \geq 0$

(iii.) $F(x)$ is continuous on the right

(iv.) $F(-\infty) = 0$ (Prob. always defined on an interval)

$F(\infty) = 1$

(v) $P(a \leq x \leq b) = P(a \leq x < b) = P(a < x \leq b) = P(a < x < b)$

$= \int_{a}^{b} f(x)\, dx$

in discrete → $P(a \leq x \leq b) = F(b) - F(a)$

$P(a < x \leq b) = F(b) - F(a-1)$

Ex. $f(x) = ke^{-3x}, \quad x \geq 0$

$0, \quad x < 0$

Determine k & hence, compute cdf of x

DATE : / /

PAGE NO. :

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \rightarrow \int_{-\infty}^{0} 0 + \int_{0}^{\infty} k e^{-3x}\,dx = 1$$

$$= k\left(\frac{e^{-3x}}{-3}\right)\Big|_0^{\infty} = k\left[\frac{+1}{3}\right] = 1 \rightarrow \boxed{k = 3}$$

$$F(x) = \int_0^x 3e^{-3x}\,dx = \frac{3}{-3}\left[e^{-3x}\right]_0^x = 1 - e^{-3x}, \; x \geq 0$$

0 ew

$$F(2) = P(X \leq 2) = 1 - e^{-3(2)} = 1 - e^{-6}$$

$$F(4) = 1 - e^{-3(4)} = 1 - e^{-12}$$

Ex. Pdf of random variable X :

$$f(x) = \begin{cases} x & 0 \leq x < 1 \\ 2 - x & 1 \leq x < 2 \\ 0 & x \geq 2 \end{cases}$$

4 - 2

Compute cdf of x:

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

$$F(x < 1) = \int_0^x x\,dx = \frac{x^2}{2}\Big|_0^1 = \frac{x^2}{2}$$

$$F(x < 2) = \int_0^1 x\,dx + \int_1^x (2-x)\,dx = \frac{1}{2} + 2x - \frac{x^2}{2}\Big|_1^x$$

$$= \frac{1}{2} + 2x - \frac{x^2}{2} - \frac{3}{2} = 2x - \frac{x^2}{2} - 1$$

$$F(x \geq 2) = \int_0^1 x\,dx + \int_1^2 (2-x)\,dx = \frac{1}{2} + 2 - \frac{3}{2} = \underline{\underline{1}}$$

$$F(x) = \begin{cases} 0 & x < 0 \\ x^2/2 & 0 \leq x < 1 \\ 2x - \frac{x^2}{2} - 1 & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

DATE: / /

PAGE NO.:

(i) $P(-1 \le x \le 3) = F(3) - F(-1) = 1 - 0 = \boxed{1}$

(ii) $P(1 \le x \le 1.5) = 2x - \frac{x^2}{2} - 1 = 2\left(\frac{3}{2}\right) - \left(\frac{9}{8}\right) - 1$

$\Downarrow$

$F(1.5) - F(1) \qquad = 3 - \frac{17}{8} = \frac{7}{8} - \left(\frac{1}{2}\right)$

$= \boxed{\frac{3}{8}}$

Expectation :

$\mu = E(x) = \sum_{\forall x} x\, P(x)$ if x is discrete

$= \int_{-\infty}^{\infty} x\, f(x)\, dx$ if x is continuous

Let $g(x)$ = func^n of x, so,

$E[g(x)] = \sum_{\forall x} g(x)\, p(x)$

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

(Always +ive) (Unitless) Variance :

$\sigma^2 = E[(x-\mu)^2] = \sum_{\forall x} (x-\mu)^2\, p(x)$ discrete

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

(Has unit) $S.D.(\sigma) = +\sqrt{E(x-\mu)^2}$

→ Computational formula for σ^2 :

$\sigma^2 = \sum x^2 p(x) - \left[\sum x\, p(x)\right]^2$

$\boxed{\sigma^2 = E[x^2] - (E[x])^2}$

$\sigma^2 = E[(x-\mu)^2] = \sum_{\forall x} (x-\mu)^2 p(x)$

$= \sum_{\forall x} x^2 p(x) - 2\mu x\, p(x) + \mu^2 p(x)$

DATE : / /

PAGE NO. :

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

$$= \sum x^2 p(x) - 2\mu(\mu) + \mu^2 (1)$$

$$= \sum x^2 p(x) - \mu^2$$

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

continuous : $\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx$

Properties of Expectations & Variance

(i.) $E(c) = c$ where c: constant $= \sum c p(x) = c \sum p(x) = c$

(ii.) $E(cx) = c\, E(x)$

(iii.) $E(c+x) = c + E(x)$

(iv.) $Var(c) = 0$ $= E[c^2] - (E[c])^2 = c^2 - c^2 = 0$

(v.) $Var(cx) = c^2\, var(x) = E[x^2 c^2] - [E(cx)]^2$

$= c^2 E[x^2] - (c^2 E(x)^2) = c^2\, var($

Ex $E[X] = 5$

$var(X) = 2$

$E[X-5] = ?$ $Var(-2X) = ?$

$var(-2X) = 4\, Var(X) = 4 \times 2 = 8$

$E[X-5] = E[x] - 5 = 0$

The r^{th} moment about origin :

$\mu'_r = E[X^r] = \sum_{\forall x} x^r p(x)$: x is discrete

$\int_{-\infty}^{\infty} x^r f(x)\, dx$ ~ cont.

$r=1 \rightarrow E[X]$: 1^{st} moment about origin is Mean. $= \underline{\underline{\mu}}$

(μ'_1)

r^{th} moment about Mean :

$\mu_r = E[(x-\mu)^r]$: $\sum_{\forall x} (x-\mu)^r p(x)$

$\int_{-\infty}^{\infty} (x-\mu)^r f(x)\, dx$

~~$M[X(t)] = E[$~~

DATE: / /

PAGE NO.:

If $r = 2$

$\mu_2 = E[(x-\mu)^2] = Var(x) = \sigma^2$ → 2^{nd} Moment about Mean is Variance

$$\boxed{\sigma^2 = \mu_2' - \mu^2}$$

$$\sigma^2 = E[(x-\mu)^2] = \Sigma (x-\mu)^2 p(x)$$
$$= E[x^2] - (E(x))^2$$
$$= \mu_2' - \mu^2$$

→ **Moment Generating func^n : MGF** — func^n is used to generate moments

$M_x(t) = E[e^{tx}] = \sum_{\forall x} e^{tx} p(x)$ discrete

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

$$\mu_r' = \frac{d^r}{dt^r}[M_x(t)]\Big|_{t=0}$$

diff. r times

$\mu = \mu_1' = \frac{d}{dt}[M_x(t)]\Big|_{t=0}$ $\sigma^2 = \mu_2' - \mu^2$ → calculate from $M_x(t)$

→ **Mean Deviation about Mean**:

$$M.D. = \sum_{\forall x} |x-\mu|\, p(x)$$

- sometimes, moments generated are complex, but we want only real func^{n}s, so we multiply it with i:

→ **Characteristic Func^n**:

$\phi_x(t) = E[e^{itx}]$

Engineering Mathematics - C.B. Gupta

DATE: / /

PAGE NO.:

Probability Distributions

① Bernoulli :

Assump's :

- 2 outcomes of each experiment
- P(success) is const. (p) , P(failure) : (1-p) = q

$$\cancel{p(x) = \begin{cases} p & x=1 \rightarrow \text{success} \\ q & x=0 \rightarrow \text{failure} \end{cases}}$$

$P(X=1) = p \rightarrow$ success $\quad P(X=0) = q \rightarrow$ failure

$$p(x) = \begin{cases} p^x q^{1-x} & , x = 0,1 \\ 0 & \text{e.w.} \end{cases}$$

$E(X) = p \quad , \quad Var(X) = pq$

$M_X(t) = q + pe^t$

13/9/19

Joint Distributions

~~Discrete~~ Joint Mass Func^n (discrete joint densities)

$f_{XY}(x,y) = P(x,y) = P[X=x \text{ and } Y=y]$

- $f_{XY}(x,y) \geq 0 \quad \forall (x,y)$
- $\sum_{\forall x} \sum_{\forall y} f_{XY}(x,y) = 1$

} → Necessary & sufficient

* Necessary / Sufficient / Necessary & sufficient

Marginal distrib^n (individual ki prob.)

$f_X(x) = \sum_{\forall y} f(x,y)$ → Marginal density for x.

$f_Y(y) = \sum_{\forall x} f(x,y)$

Discrete : Random variable takes value countable.
Cont : interval

DATE : / /

PAGE NO. :

Continuous Joint Density

1. $f_{XY}(x,y) \geq 0 \quad ; \quad x, y \in \mathbb{R}$
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, dy\, dx = 1$
3. $P[a \leq X \leq b \text{ and } c \leq Y \leq d] = \int_a^b \int_c^d f(x,y)\, dy\, dx$

↳ Joint density for (X,Y)

Cont. Marginal Densities :

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\, dy \quad , \quad f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\, dx$$

Independence :

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$$

$$P(X = x, Y = y) = P[X = x] * P[Y = y]$$

or $$f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$$

Multivariant : $f(x_1, x_2, \ldots x_n) = f(x_1) \cdot f(x_2) \cdots f(x_n)$

Ex-1 1^{st} task → 2 joins 2^{nd} task → 3 bolts

X → no. of defective welds

Y → no. of improperly tightened bolts produced per car

x/y	0	1	2	3	$f_X(x)$
0	.840	.030	.020	.010	.900
1	.060	.010	.008	.002	.080
2	.010	.005	.004	.001	.020
$f_Y(y)$	.910	.045	.032	.013	1.000

$f_X(x=0) = \sum_{\forall y} f(0,y) \Rightarrow 1^{st}$ row sum

DATE: / /

PAGE NO.:

checking independence here are different 12 values combination (check all combination)

↳ That's why checking independence in case of discrete is difficult

Question ~~Table~~ 2

(a) $P(X=2, Y=1) = 0.005$

(b) $P(X>=1, Y>=1) = \cancel{0.045 + 0.032 + 0.}$ 0.030

(c) $P(X<=1) = 0.980$

(d) $P(Y>=2) = 0.032 + 0.013 = 0.045$

Ques

$$f_{XY}(x,y) = \frac{1}{n^2}, \quad x = 1, 2, \ldots n, \quad y = 1, 2 \ldots n$$

(a) verify $f(x,y)$ satisfies conditional

(1) $f(x,y) \geq 0$

$n^2 \geq 0$

$\Rightarrow \frac{1}{n^2} \geq 0$

(2) $\sum_{\forall x} \sum_{\forall y} f(x,y) = 1$

$$\sum_{\forall x} \sum_{\forall y} f(x,y) = \sum_{x=1,2,\ldots n} \sum_{y=1,2\ldots n} \frac{1}{n^2}$$

$$= \sum_{x=1,\ldots n} \frac{1}{n^2} + \frac{1}{n^2} + \ldots \text{ n times} = \sum_{x=1,2..n} n\left(\frac{1}{n^2}\right) = \frac{1}{n}$$

$$= \boxed{1}$$

(b) find marginal density for X & Y

$$f_X(x) = \sum_{y=1}^{n} \frac{1}{n^2} = \frac{1}{n}, \quad f_Y(y) = \frac{1}{n} \quad y = 1, 2, \ldots n$$

$x = 1, 2, \ldots n$

0, elsewhere

(c) Are X & Y independent?

$f_X(x) \times f_Y(y) = f(x,y) \Rightarrow$ Independent

Ques 4

$$f_{XY}(x,y) = c(4x + 2y + 1) \quad 0 \leq x \leq 40; \; 0 \leq y \leq 2;$$

DATE : / /

PAGE NO. :

$f(x,y) = c(4x + 2y + 1) \quad 0 \le x \le 40 \quad , \quad 0 \le y \le 2$ Q5

(a) Find c

$$c\int_0^{40}\int_0^{2}(4x+2y+1)\,dy\,dx$$

$$c\int_0^{40}(4xy + y^2 + y)\Big|_0^2 dx = c\int_0^{40}(6+8x)\,dx = 1$$

$$4 + 2 + 8x \Rightarrow c \quad 6x + 4x^2\Big|_0^{40} = 1$$

6400
240

$$\Rightarrow \boxed{c = \frac{1}{6640}}$$

$$\Rightarrow f(x,y) \begin{cases} \frac{1}{6640}(4x+2y+1) & 0 \le x \le 40 \ , \ 0 \le y \le 2 \\ 0 & \text{e.w.} \end{cases}$$

(b) $f_X(x) = ?$ $\quad f_Y(y) = ?$

$$f_X(x) = c\int_0^2(4x+2y+1)\,dy = \frac{8x+6}{6640} \quad , \quad 0 \le x \le 40$$
$$0 \quad \text{e.w.}$$

3200
+ 40
3240

$$f_Y(y) = c\int_0^{40}(4x+2y+1)\,dx = c\left[2x^2 + (2y+1)x\right]\Big|_0^{40} = \frac{80y + 3240}{6640}$$
$$0 \ , \ \text{e.w.}$$

(c) $f(X > 20, y > 1) = c\int_{20}^{40}\int_1^2 (4x+2y+1)\,dy\,dx$

$$= c\int_{20}^{40}(4xy + y^2 + y)\Big|_1^2 dx = \cancel{c}\int_{20}^{40}\Big(\quad \frac{2480}{6640}$$

(d) $P(y > 1) = \int_1^2 \frac{80y + 3240}{6640}\,dy = 0.526$

(e) $P(X > 20) = c\int_{20}^{40} \frac{8x+6}{c}\,dx = 0.741$

(f) $f_X(x) \cdot f_Y(y) \ne f_{XY}(x,y) \rightarrow$ Not Independent

DATE : / /

PAGE NO. :

Q5. (n- dimensional discrete random variables)

$f(x_1, x_2, \ldots x_n) = P[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n]$

$P(x_1) = 0.9$ → non-def $\quad P(x_2) = 0.08$ → def but salvagable $\quad P(x_3) = 0.02$ → neither

n = 20 items

(a) $P[X_1 = 15, X_2 = 3, X_3 = 2] = ?$

(b) Find general formula for density for (X_1, X_2, X_3).

sol^n. $P(X_1 = x_1, X_2 = x_2, X_3 = x_3) = f(x_1, x_2, x_3)$

$$= \frac{n!}{x_1!\, x_2!\, x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3} \Rightarrow \boxed{\text{Multinomial distrib}^n}$$

↓ extension of binomial one.

$$= \frac{20!}{15!\, 3!\, 2!} (0.9)^{15} (0.08)^3 (0.02)^2$$

$$= \underline{\underline{0.0065}}$$

lecture 2 :

Expectation & Covariance

① $E[H(X,Y)] = \sum_{\forall x} \sum_{\forall y} H(x,y) f_{XY}$ ~~[illegible]~~

② $= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(x,y) f_{XY}\, dy\, dx$

Ex. $f(x,y) = \begin{cases} \frac{3}{2}(x^2+y^2) & 0 \leq x < 1, \quad 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$

$E[X^2 + Y^2]$

$$= \int_0^1 \int_0^1 (x^2+y^2) \frac{3}{2}(x^2+y^2)\, dy\, dx$$

$$= \frac{3}{2} \int_0^1 \left(\int_0^1 x^4 + 2x^2y^2 + y^4 \right) dy\, dx =$$

DATE : / /

PAGE NO. :

$$\frac{3}{2}\int_0^1 \left[x^4 y + \frac{2x^2y^3}{3} + \frac{y^5}{5}\right]\Big|_0^1 dx$$

$$= \frac{3}{2}\int_0^1 \left(x^4 + \frac{2x^2}{3} + \frac{1}{5}\right) dx$$

$$= \frac{3}{2}\left[\frac{1}{5} + \frac{2}{3}\times\frac{1}{3} + \frac{1}{5}\right] = \frac{3}{2}\left[\frac{1}{5} + \frac{2}{9} + \frac{1}{5}\right]$$

$$= \frac{3}{2}\left[\frac{2}{5} + \frac{2}{9}\right] = 3\left[\frac{1}{5} + \frac{1}{9}\right] \cdot \frac{3 \times 14}{5 \times 9_3} = \frac{14}{15}.$$

<u>Univariate</u> <u>Average</u> : calculate $E[X]$ from $f(x,y)$

$$E[X] = \sum_{\forall x} x \underbrace{f_X(x)}_{\text{Marginal density of } x.} = \sum_{\forall x}\sum_{\forall y} x\, f(x,y)$$

<u>Covariance</u> : → How X & Y are related

$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$ $\qquad E[X]=\mu_x,\ E[Y]=\mu_y$

* $X\downarrow$ & $Y\downarrow$ ⇒ $(X-\mu_x)(Y-\mu_Y) > 0$ } +tively correlated

$X\uparrow$ & $Y\uparrow$ ⇒ " " .

else < 0

if $Cov(X,Y) = 0$ → X & Y are independent

$$\boxed{Cov(X,Y) = E[XY] - E[X]E[Y]}$$

$$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$$
$$= E[XY - YE[X] - XE[Y] + E[X]E[Y]]$$
$$= E[XY] - E[Y]E[X] - \cancel{E[X]E[Y]} + \cancel{E[X]E[Y]}$$
$$= E[XY] - E[X]E[Y]$$

DATE : / /

PAGE NO. :

Thm If X & Y are independent :

$E[XY] = E[X]\,E[Y] \Rightarrow \underline{Cov(X,Y) = 0}$ But not vice versa

Proof $E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\, dx\, dy$

$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_X(x)\, f_Y(y)\, dx\, dy$

$= \int_{-\infty}^{\infty} x f_X(x)\, dx \int_{-\infty}^{\infty} y\, f_Y(y)\, dy = E[X]\,E[Y]$

Ex. $p(-1) = p(0) = p(1) = 1/3$ proof

Let $Y = X^2$

$E(X) = (-1+0+1)/3 = 0$

$E[Y] = (1+0+1)/3 = 2/3$

$E(XY) = (-1+0+1)/3 = 0$

$E[XY] = E(X)\,E(Y)$, but X & Y are <u>dependent</u>.

Proof $Cov(X,Y) = Var(X) = Var(Y)$ if $X = Y$

if $X = Y$, $\Rightarrow Cov(X,X) = E[X \cdot X] - E[X]\,E[X]$

$= E[X^2] - (E(X))^2 = Var(X)$

17/9/19 $Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$

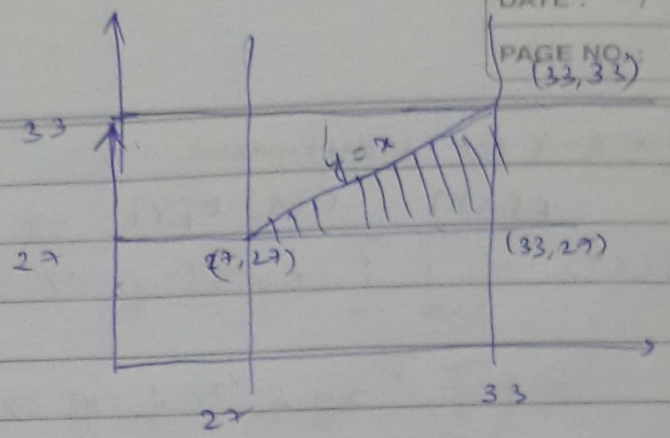
Exercise 19. The joint density is given by

$f_{XY}(x,y) = \frac{c}{x}$ $27 \le y \le x \le 33$

$c = \frac{1}{(6 - 27\ln 33/27)} = 1.72$

(a) Find $E[X]$, $E[Y]$, $E(XY)$, $Cov(X,Y)$

(b) $E[X-Y]$



$\cancel{\int \frac{c}{x} dx}$

To find c :

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

$$= \int_{27}^{33}\int_{27}^{x} \frac{c}{x}\,dy\,dx \Rightarrow \int_{27}^{33} \frac{c}{x}(x-27)\,dx$$

$$= c\int_{27}^{33}\left(1-\frac{27}{x}\right)dx = 1 \Rightarrow c\left[x - 27\ln x\Big|_{27}^{33}\right] = 1$$

$$\Rightarrow c = \frac{1}{6 - 27\ln 33/27} = \underline{1.72}$$

(a) ~~E[x]~~ $\Rightarrow$

$$f_x(x) = \int_{27}^{x} \frac{c}{x}\,dx \Rightarrow \frac{(x-27)}{x}c$$

$$f_Y(y) = \int_{y}^{33} \frac{c}{x}\,dx \Rightarrow c(\ln 33 - \ln y)$$

$$E[x] = \int_{27}^{33} x f_x(x)\,dx = \int_{27}^{33} c(x-27)\,dx = \frac{x^2}{2} - 6*27$$

$$= 30.96$$

$$E[Y] = \int_{27}^{33} c(\ln 33 - \ln y)\,y\,dy = 28.99$$

$$E[XY] = \int_{27}^{33}\int_{y}^{33} xy\frac{c}{x}\,dx\,dy = 897.84$$

~~(a) E[X-Y] =~~ $Cov(XY) = E[XY] - E(X)E[Y] = \underline{0.31}$

(b) $E[X-Y] = \int_{27}^{33}\int_{27}^{x} (x-y)\frac{c}{x}\,dy\,dy = 1.97$

DATE: / /

PAGE NO.:

* Prove $Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$

If, X & Y are independent:

$Var(X+Y) = Var(X) + Var(Y)$, $Var(X-Y) = Var(X) + Var(Y)$

$$var(X+Y) = E[(X+Y)^2] - (E[X+Y])^2$$
$$= E[X^2 + Y^2 + 2XY] - (E[X] + E[Y])^2$$
$$= E(X^2) + E[Y^2] + 2E[XY] - (E[X])^2 - (E[Y])^2 - 2E[X]E[Y]$$
$$= Var(X) + Var(Y) + 2\,Cov(X,Y)$$

Ex. $$f_{XY}(x,y) = \frac{1}{2(e-1)}\left[\frac{1}{x} + \frac{1}{y}\right) \qquad \begin{array}{l} 1 \le x \le e \\ 1 \le y \le e \end{array}$$

(a) show that it is valid joint density

(b) Find $E[X]$, $E[Y]$, $E[XY]$ $\quad \frac{3e-1}{4}, \frac{3e-1}{4}, \frac{e^2-1}{2}$

(c) X & Y Independent? No

(a) $$\int_1^e \int_1^e \frac{1}{2(e-1)}\left[\frac{1}{x} + \frac{1}{y}\right] dy\,dx$$

$$= \int_1^e \frac{1}{2(e-1)}\left(\frac{y}{x} + \ln y\right)\Big|_1^e dx = \frac{1}{2(e-1)}\int_1^e \left(\frac{(e-1)}{x} + 1\right) dx$$

$$= \frac{1}{2(e-1)}\left((e-1)\ln x + x\right]\Big|_1^e = \frac{1}{2(e-1)}\left[e-1 + e-1\right] = 1$$

(b) $E[X]$:

$$f_X(x) = \int_1^e \frac{1}{2(e-1)}\left[\frac{1}{x} + \frac{1}{y}\right] dy = \frac{1}{2(e-1)}\left[\frac{(e-1)}{x} + 1\right]$$

$$E(X) = \int_1^e x f_X(x)\,dx = \int_1^e \left(\frac{x}{2(e-1)}\left[\frac{e-1}{x} + 1\right]\right) dx$$

$$= \int_1^e \frac{1}{2} + \int_1^e \frac{x}{2(e-1)}\,dx \qquad \frac{x^2}{2}$$

$$= \frac{e-1}{2} + \frac{1}{2(e-1)}\left[\frac{e^2-1}{2}\right] \Rightarrow \frac{e-1 \times 2}{2 \times 2} + \frac{e+1}{4} = \frac{3e-1}{4}$$

20/9/19

DATE: / /

PAGE NO.:

Conditional Distribution

Conditional Densities.

$f_{X|y}$: $P(X=x \mid Y=y) = \frac{P[X=x \ \& \ Y=y)}{P[Y=y]} = \frac{f_{XY}(x,y)}{f_Y(y)}, \ f_Y(y)>0$

satisfies

1) ≥ 0

2) $\int_{-\infty}^{\infty} f_{X|y} \, dx = 1$

Cond^n al Mean & variance

$E(Y|x) = \sum y f_{Y|x}(y) = \int y f_{Y|x}(y)\, dy$

$var(Y|x) = \sum_y y^2 f_{Y|x}(y) - \mu^2_{Y|x}$

Ex. $f(x,y) = x^2 + \frac{xy}{3}, \quad 0 \le x \le 1, \ 0 \le y \le 2$

$\quad\quad 0 \quad\quad e.w.$

sol^n. Marginal density for x ?

$f_X(x) = \int_0^2 \left(x^2 + \frac{xy}{3}\right) dy = x^2 y + \frac{x}{6} y^2 \Big|_0^2$

$= 2x^2 + \frac{2}{3}x \quad 0 \le x \le 1$

$f_Y(y) = \int_0^1 \left(x^2 + \frac{xy}{3}\right) dx = \frac{1}{3} + \frac{y}{6}$

- $f_{X|y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{x^2 + xy/3}{1/3 + y/6} = \frac{6x^2 + 2xy}{2+y}, \ \ 0 \le x \le 1, \ 0 \le y$

- $f_{Y|x} = \frac{f_{XY}(x,y)}{f_X(x)}$

DATE: / /

PAGE NO.:

Mean & Variance of linear combⁿ of Random variables

$$Y = c_1 X_1 + c_2 X_2 + \dots + c_p X_p \qquad X_i \to r.v. \quad c_i \to \text{consts}$$

$$E[Y] = c_1 E[X_1] + c_2 E(X_2) + \dots + c_p E[X_p]$$

$$V(Y) = c_1^2 V(X_1) + c_2^2 V(X_2) + \dots + c_p^2 V(X_p) + 2 \sum_{i<j} \sum c_i c_j \, cov(X_i, X_j)$$

If $X_1, X_2, \dots, X_p$ are independent

$$V(Y) = c_1^2 V(X_1) + c_2^2 V(X_2) + \dots + c_p^2 V(X_p)$$

→ $V(Y) = E[(Y - E[Y])^2]$

$= E[(c_1X_1 + c_2X_2 + \dots + c_pX_p - (c_1\mu_1 + c_2\mu_2 + \dots + c_n\mu_n))^2]$

$= E[(c_1(X_1 - \mu_1) + c_2(X_2 - \mu_2) + \dots + c_p(X_p - \mu_p))^2]$

$= E[c_1^2(X_1-\mu_1)^2 + c_2^2(X_2 - \mu_2)^2 + \dots + - 2 \dots$

Ex. X & Y : independent r.v.

$\mu_x = 7 \qquad \mu_y = -5$

$V(X) = 9 \qquad V(Y) = 3$

(i) $E[4X - 2Y + 6] = 4E[X] - 2E[Y] + 6 = 4(7) - 2(-5) + 6$

$= 28 + 10 + 6$

$= 44$

(ii) $V(4X - 2Y + 6) = 16V(X) + 4V(Y) + 36$

$= 16(9) + 4(3) + \cancel{36}^{\,0} = 144 + 12 + 36$

$= 156$

DATE: / /

PAGE NO.:

Ex. Let X_1 & X_2 be 2 independent r.v. having normal distribⁿ having mean 5 & 6, var 2 & 3. Find mean of $Y = X_1 + X_2$. Find distribⁿ of Y

Solⁿ (Take help of Moment generating funcⁿ)

$M_x(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$

Given $X_1 \sim N(5, 2)$

$M_{X_1}(t) = e^{5t + \frac{1}{2} \cdot 2t^2}$

$M_{X_2}(t) = e^{6t + \frac{1}{2} \cdot 3t^2}$

$M_Y(t) = M_{(X_1+X_2)}(t) = M_{X_1}(t) \cdot M_{X_2}(t)$ [X_1 & X_2: independent]

$= (e^{5t+t^2}) \cdot e^{(6t + 3/2 t^2)}$

$= e^{11t + \frac{5}{2}t^2} \rightarrow N(11, 5)$

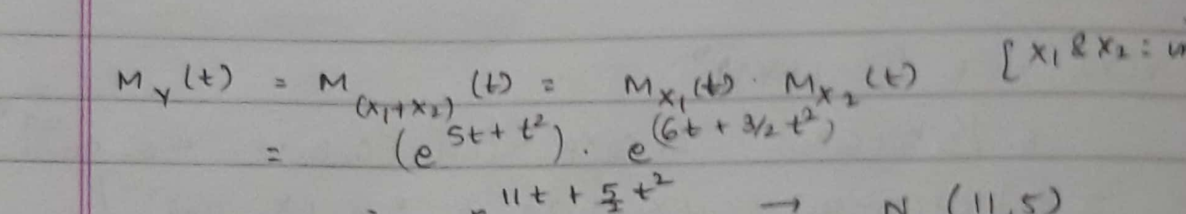
$Y \rightarrow$ also a Normal distribⁿ.

Correlation

$\rho \rightarrow$ Pearson Coeffecient

$$\boxed{\rho = \frac{Cov(X,Y)}{\sqrt{(Var(X))(Var(Y))}}}$$ $\rightarrow$ linear relⁿ b/w X & Y

If distribⁿ is given, use this



Cor : +ive   Cor : 0   Cor : -ive

$\rho = -1 \rightarrow$ Perfect -ive correlation

$\rho = 0 \Rightarrow$ Independent var (Not correlated) $\rightarrow$ Non-Linear relⁿship

$\rho = 0 \Rightarrow$ Perfect +ve correlⁿ.

→ Not possible for large popⁿ to calculate ρ, taken random samp of size n.

$\hat{Var}\ x = \sum_{i=1} (x_i - \bar{x})^2 / n = S_{xx}/n$. similarly for y.

DATE: / /

PAGE NO.:

$$\operatorname{Cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

$\Rightarrow$ $\hat{\operatorname{cov}}(X,Y) = \Sigma(X_i-\bar{X})(Y_i-\bar{Y})/n = S_{XY}/n$

$$\hat{\rho} = R = \frac{S_{XY}}{\sqrt{S_{XX}S_{YY}}}$$

Computational formula :

$$\hat{\rho} = \frac{n\Sigma xy - \Sigma x\,\Sigma y}{\left(\left[(n\Sigma x^2-(\Sigma x)^2)\right]\left[n\Sigma y^2-(\Sigma y)^2\right]\right)^{1/2}}$$

$\Rightarrow$ Use when data is given

Q. Find sample correlⁿ coeff.

(i) X : 1, 2, 3, 4, 5

Y = 2, 4, 6, 8, 10

solⁿ : ~~$\hat{\rho}$ = 0~~

	x	y	x^2	y^2	xy
	1	2	1	4	2
	2	4	4	16	8
	3	6	9	36	18
	4	8	16	64	32
	5	10	25	100	50
Σ :	15	30	55	220	110

$\times 5$

1100
900

$$\hat{\rho} = r = \frac{5(110)-(15)(30)}{\sqrt{[5\times 55-225][5\times 220-900]}}$$

$$= \frac{550-450}{\sqrt{50 \times 200}} = \frac{100}{\sqrt{10000}} = \frac{100}{100} = \boxed{1}$$

4

(4

DATE : / /

PAGE NO. :

2.) $f_{XY}(x,y) = \frac{1}{6640}(4x+2y+1)$ $\quad 0 \le x \le 40$, $0 \le y \le 2$

Find $E[X]$, $E[Y]$, $E[X^2]$, $E[Y^2]$, $Var(X)$, $var(Y)$, $Cov(XY)$, ρ

Soln $E[X,Y] = \int_0^{40}\int_0^2 xy \frac{1}{6640}(4x+2y+1)\,dy\,dx = 26.586$

$E[X] = \int_0^{40}\int_0^2 x f_{X}(x)\,dx\,dy = 26.426$ $\quad E[X^2] = 700.361$

$E[Y^2] = 1.3494$

$E[Y] = 1.008$

$cov(X,Y) = E[XY] - E[X]E[Y] = 26.586 - (26.426)(1.008)$

$= -0.051$

$Var(X) = 700.361 - (26.426)^2 = 2.027$

$Var(Y) = 1.3494 - (1.008)^2 = 0.333$

$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}} = -0.009 \Rightarrow$ -ively correlated

24/9/19

Chi-square distribution

Let $Z \sim N(0,1)$

$\Rightarrow Z^2 \sim \chi^2$ → chi-square

Let $X_1, X_2, \ldots, X_n$ be n independent variables with mean & variance μ_i & σ_i^2 $(i=1,2,\ldots n)$ then,

$\sum_{i=1}^{n}\left(\frac{X_i-\mu_i}{\sigma_i}\right)^2$ is a χ^2-variate with n.d.f.

Let X be χ^2 variate with n.d.f. n

$E[X] = n$

$Var(X) = 2n$

χ χ χ

DATE:
PAGE NO.:

→ This distribution is used for Goodness of fit
↓
Observed & Expected frequency → calc → con-spm
↓
check if theoretical value is same as observed

Chi-square (χ^2) Test of goodness of fit:

Let O_i ($i=1,2,\dots n$) be a set of observed frequencies & e_i ($i=1,2,\dots n$) corresponding set of expected (theoretical) frequencies. Then

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - e_i)^2}{e_i}$$ has a χ^2 distribⁿ with $(n-1)$ d.f.
↓
degree of freedom

given, all $e_i \geq 5$. (Thumb rule)

if $\chi^2_{calc} < \chi^2_{Tab}$ ⇒ We have made correct assumption. If no significant difference, we can accept it.

Ex. The demand of a spare part in a factory was found to vary from day to day. In a sample study, foll. info was obtained.

Days	M	T	W	Th	F	S	Total
No. of parts demanded (O_i)	1124	1125	1110	1120	1126	1115	6720 = n

⟹ Observed frequencies

Test the hypothesis, that the no. of parts doesn't depend on the day.
⇓
Solⁿ P(demand) is same for all day

① find total of all days = 6720 = n

H_0: The demand of no. of parts doesn't depend on day

H_1: H_0 is not true

$p_i = \frac{1}{6}$, $i = 1, 2, \dots, n$ $e_i = n \times p_i$

→ No. of parts (e_i) : 1120 1120 1120 1120 1120 1120

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i}$$

$$= \frac{(1124-1120)^2}{1120} + \frac{(1125-1120)^2}{1120} + \dots$$

$$= 0.179$$

No. of cells = 6

$\Rightarrow$ Degree of freedom here = 5

$\chi^2_{\text{tab with 5 d.o.f.}} = 11.07$ (from table)

$0.179 < 11.07 \Rightarrow$ no significant difference

$\Rightarrow H_0$ is accepted.

$\Rightarrow$ If any $e_i < 5 \rightarrow$ can add it with any preceding or succeeding cell (add both e_i & o_i). Thus, d.o.f. --;

Student t-distribution

→ It is a sample distrib^n

→ Can be used wherever z-distrib^n is used, only cond^n is $n < 30$

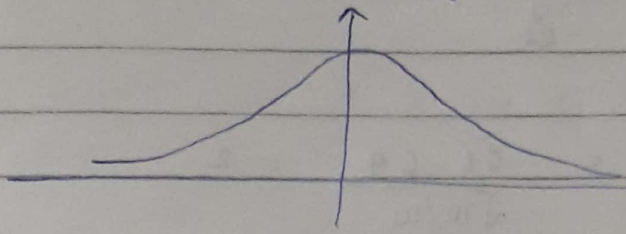
→ ~~If~~ let $x_1, x_2, \dots x_n$ be random sample of small size n ~~with mean μ_i and variance σ_i^2 respectively. ($i=1,2,\dots n$)~~ from a normal population with mean μ & variance σ^2. Let $\bar{x}$ be mean of sample & s^2 be variance of sample.

ie, $$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

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

In sample, only $n-1$ samples are linearly independent then $t = \dfrac{\bar{x} - \mu}{S/\sqrt{n}}$ has a t-distrib^n with $(n-1)$ d.f

$f(-t) = f(t)$: symmetric



⇒ ~~mean, mode = 0~~

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

$\mu_{2k} = \dfrac{n^k (2k-1)(2k-2) \dots 3.1}{(n-2)(n-4)\dots(n-2k)}$, $\dfrac{n}{2} > k$

μ_2 (Var) $= \dfrac{n}{n-2}$, $n > 2$, $\boxed{\text{mean} = 0}$ $(\mu_1 = 0)$

Significance Test for Sample Mean

Let $x_1, x_2, \dots x_n$ be a random sample of size n from a normal population with a specified popul^n mean μ ~~and~~.

To test $\mu = \mu_0$ or there is no significant difference between sample mean $\bar{x}$ and ~~sam~~ popl^n mean μ_0, cal compute

$t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$ where $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$ & $s^2 = \dfrac{1}{n-1} \sum (x_i - \bar{x})^2$

- if $t_{cal} < t_{tab}$ ⇒ no significant diff. between $\bar{x}$ & μ_0

~~Quest^n~~ If $n > 30$ → then t will become z.

Quest^n: The heights of 10 men of a locality are found to be:
70, 67, 62, 68, 61, 68, 70, 64, 64, 66
Is it reasonable to believe that avg. height is > 64 inches?
Test at 5% level of significance

$t > t_\alpha$
$t < t_\alpha$
$|t| < t_{\alpha/2}$

DATE: / /
PAGE NO.:

Solⁿ $\bar{x} = \frac{\sum x_i}{n} = \frac{660}{10} = 66$ $\alpha = 0.05$

$s^2 = \frac{1}{10-1} \sum_{i=1}^{10} (x_i - \underset{66}{\bar{x}})^2 = 10$

$\mu_0 = 64$

$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{66-64}{\sqrt{10/10}} = 2$

t_{tab} with 9 d.f. $= 2.26$

$t_{calc} < t_{tab}$ ⇒ Avg. height is not greater than 64 inches

F- distribution

Let X & Y be two independent chi-square variates with ν_1 & ν_2 d.f. Then,

$$F = \frac{X/\nu_1}{Y/\nu_2} = F(\nu_1, \nu_2)$$

Appⁿs :

↳ entire analysis of variance of
↳ Agriculture, Medical

- Let $x_1, x_2 \ldots x_n$ & $y_1, y_2 \ldots y_n$ be 2 independent random samples drawn from same normal population with same ~~mean~~ varian
Let $\bar{x}$ & $\bar{y}$ be the sample mean & s_x^2 & s_y^2 be sample variance of 2 samples.

⇒ $\bar{x} = \frac{1}{n_1} \sum x_i$ $\bar{y} = \frac{1}{n_2} \sum y_i$

$s_x^2 = \frac{1}{n_1 - 1} \sum (x_i - \bar{x})^2$ $s_y^2 = \frac{1}{n_2 - 1} \sum (y_i - \bar{y})^2$

Then, $\boxed{F = \frac{s_x^2/n_1}{s_y^2/n_2}}$