Dated: Lecture # 17

Random variable

Discrete Random variable: Random variable that assumes countable values.

continuous Random variable: value that exists in an interval.

Probability distribution

n	p(n)
0	0.19
1	0.11
2	.20

not representing probability distribution

x	p(n)
1	0.33
2	0.33
3	0.33

yes it represents

Ex number of break downs

(a)
(b)
(c)
(d)

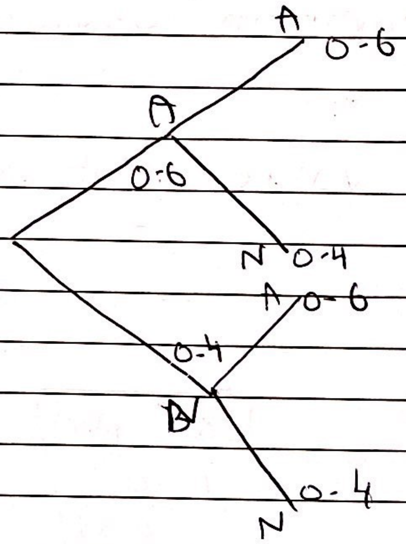
0 .19
1
2
3

0/10

Dated:

60 % students suffer from Math Anxiety

- Two students selected at random



$P(AA) = P(A) \cdot P(A) = 0.36$

$P(AN) = P(A) \cdot P(N) = 0.24$

$P(NA) = P(N) \cdot P(A) = 0.24$

$P(NN) = P(N) \cdot P(N) = 0.16$

10/10

X	P(x)
0	0.16
1	2(0.24)
2	0.36

Dated: Lecture # 18

Example

item 20

defected 3

select 2 items

$$F(0) = P(x=0) = \frac{{}^{3}C_0\,{}^{17}C_2}{{}^{20}C_2} = \frac{68}{95}$$

↑ zero defected

Non defected

pdf

x	$f(x)$
0	68/95
1	51/190
2	3/190

$$f(1) = P(x=1) = \frac{{}^{3}C_1\,{}^{17}C_1}{{}^{20}C_2} = \frac{51}{190}$$

$$f(2) = P(x=2) = \frac{{}^{3}C_2\,{}^{17}C_0}{{}^{20}C_2} = \frac{3}{190}$$

$f(x) \geq 0$ | $0 \leq P(x) \leq 1$

$\sum_x f(x) = 1$ | $\sum P(x) = 1$

$f(x) = P(x=2)$

At most 1 item is defected

$$P(x \leq 1) = f(0) + f(1)$$

At least

$$P(x \geq 1) = f(1) + f(2)$$

cumulative distribution function

$$F(x) = P(\cancel{X} \leq x)$$

$$= \sum_{t \leq x} f(t), \quad -\infty < x < \infty$$

$$F(1) = P(\cancel{X} \leq 1)$$

$$= P(X=0) + P(X=1)$$

$$f(0) + f(1)$$

Scanned with CamScanner

Dated:

$$f(x) = \begin{cases} 0, & x < 0 \quad \rightarrow \text{First} \\ f(0), & 0 \leq x < 1 \\ f(0)+f(1) & 1 \leq x < 2 \end{cases} \Big\} \text{intermediate}$$

$$\quad\quad 1 \quad\quad x \geq 2 \quad \text{last}$$

Mean discrete random variable
(Expected value)

$$\mu = \sum x\, P(x)$$

$$E(x) = \sum x P(x)$$

Dated:

Variance of a discrete random variable

$$\sigma^2 = \Sigma x^2 P(x) - \mu^2$$

$$\sigma = \sqrt{\Sigma x^2 P(x) - \mu^2}$$

x	$P(x)$	x^2	$x^2 P(x)$
0	0.15	0	0
1	0.20	1	0.2
2	0.35	4	1.4
3	0.30	9	2.7
			4.30

$$\sigma^2 = 4.30 - (1.8)^2 \rightarrow \Sigma x^2 P(x) - \mu^2$$

mean

$$\sigma = \sqrt{4.30 - (1.8)^2}$$

for expected value $\sqrt{E(x^2) - (E(x))^2}$ $\qquad \frac{\Sigma f(x)^2}{N} - \left(\frac{\Sigma f(x)}{N}\right)^2$

Example 5-8

x	$P(x)$	x^2	$x^2 P(x)$
4.5	0.32	20.25	6.48
1.2	0.51	1.44	0.7344
-2.3	0.17	5.29	0.8993
			8.11370

$\mu = (4.5)\,P(4.5) + (1.2)\,P(1.2) + (-2.3)\,P(-2.3)$

$\mu = 1.661$

Find $\sigma^2 \Rightarrow 9.7747$

$\sigma = 3.12645$

4. 8 ←

Scanned with CamScanner

Dated: # Lecture # 19

The binomial Experiment

1. There are n identical trials. Experiment repeated n times.
2. Each trial with two outcomes. Success and failure. (mutually exclusive events)
3. trials are independent.
4. p by success and q by Failure. $p+q=1$

To check if an experiment is binomial Experiment. The above properties must be satisfied

- 10 coins - 10 trials
- either head or tail
- does not affect eachother as trials
- Remains constant.

Example 5-10

- The probability doesn't remain constant throughout the experiment. if p represents "they use Insta" and q represents "they dont" ~~so~~ and it is mentioned that out of 12, 9 use instagram and the sample of 5 students can either be all insta users or a mixture.

Binomial Formula

$$P(x) = {}_nC_x\, p^x q^{n-x}$$

n = number of trials
p = prob of success
$q = 1-p$ = prob of Failure
x = num of success trials
$n-x$ = num. of failure trials

Dated:

At most = m

$$B(x) = \sum_{x=0}^{m} \text{form}$$

at least m

$$X = \sum_{x=m}^{n} \text{From}$$

$$B(x) = 1 - \sum_{x=0}^{m-1} \text{form}$$

for Exactly
the traditional Formula

Scanned with CamScanner

Lecture # 20

Dated:

Example 5-13

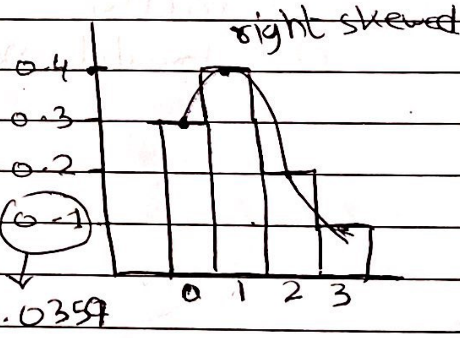
$n=3$
$p=0.33$
$q=1-0.33 = 0.67$

$P(0)={}^3C_0(0.33)^0(0.67)^{3-0} = 0.3008$
$= 0.444$
$= 0.2189$
$= 0.0359$

PDF

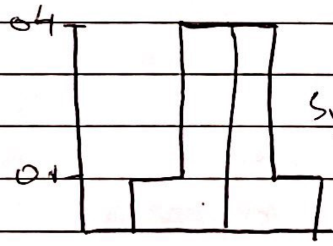
x	$f(x)$
0	0.3008
1	0.444
2	0.2187
3	0.0359

Histogram



Example 2
$n=3$
$p=0.5$
$q=0.5$

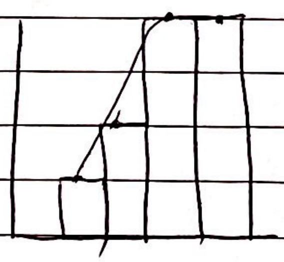
x	$f(x)$
0	0.125
1	0.375
2	0.375
3	0.125



Symmetric

$n=3$
$p=0.75$
$q=0.25$

${}^3C_0(0.75)^0(0.25)^{3-0} = 0.01563$
$= 0.14063$
$= 0.42188$
$= 0.42188$



Dated:

Mean and standard Deviation

$\mu = np$

$\sigma = \sqrt{npq}$

Ex 1 $n = 3$

$p = 0.33 < 0.5$

$q = 0.67$

left ~~right~~ skewed

$\mu = 0.99$

$\sigma = 0.81443$

Ex 2 $n = 3$

$p = 0.5$

$q = 0.5$

Symmetric

$\mu = 1.5$

$\sigma = 0.86603$

Ex 3 $n = 3$

$p = 0.75$

$q = 0.25$

right skewed

$\mu = 2.25$

$\sigma = 0.75$

Dated: **Lecture #21**

Hypergeometric prob. dist

$$P(x) = \frac{{}^{r}C_{x} \; {}^{N-r}C_{n-x}}{{}^{N}C_{n}}$$

\# of failure in population

\# of success in n trials

\# of success in population

\# of failures in n trials

\# of trials (sample size)

population size

Ex 5-17

12

7 F   5 M

(a) select 3, all ~~men~~ female

$N = 12$

$r = 7$

$x = 3$

$n = 3$

$$P(X=3) = \frac{{}^{7}C_{3} \; {}^{12-7}C_{3-3}}{{}^{12}C_{3}}$$

(b) at most 1 female

$$P(X \le 1) = P(X=0) + P(X=1)$$

$$\frac{{}^{7}C_{0} \; {}^{12-7}C_{3-0}}{{}^{12}C_{3}} + \frac{{}^{7}C_{1} \; {}^{12-7}C_{3-1}}{{}^{12}C_{3}}$$

0.36364

Dated:

① S-43 $N=11$, $r=4$, $n=4$

ⓐ $P(2) = \dfrac{{}^{4}C_{2}\ {}^{11-4}C_{4-2}}{{}^{11}C_{4}} =$

ⓑ $P(4) = \dfrac{{}^{4}C_{4}\ {}^{11-4}C_{4-4}}{{}^{11}C_{4}} =$

ⓒ $P(x \leq 1) = \dfrac{{}^{4}C_{0}\ {}^{11-4}C_{4-0}}{{}^{11}C_{4}} + \dfrac{{}^{4}C_{1}\ {}^{11-4}C_{4-1}}{{}^{11}C_{4}} =$

* The actual # of occurence within an interval is random and independent
* if the average # of occurence for a given interval is known

Conditions to apply Poisson Prob. distribution

1:- Random variable x is discrete
2. The occurence are random
3. The occurence are independent

$$P(x) = \frac{\lambda^{x} e^{-\lambda}}{x!}$$

Ex 5-18 $\lambda = 9.5$ call per week

Find the prob that exactly 6 call per week

Dated: Lecture # 22

Def: For continuous random variable X, p.d.f $f(x)$ is defined as

1) $f(x) > 0$ for all $x \in R$

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

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

Ex $f(x) = \begin{cases} \frac{x^2}{3} & -1 < x < 2 \\ 0 & \text{else when} \end{cases}$

Now $\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{-1} \overset{0}{f(x)}\,dx + \int_{-1}^{2} f(x)\,dx + \int_{2}^{\infty} \overset{0}{f(x)}\,dx$

$= ?$

(b) $P(0 < x \leq 1) = \int_{0}^{1} \frac{x^2}{3}\,dx = \frac{x^3}{9}\Big|_{0}^{1} = \frac{1}{9}$

Def (cdf)

$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$ for $-\infty < x < \infty$

Ex $F(x) = \int_{-1}^{x} \frac{t^2}{3}\,dt = \frac{x^3}{9}\Big|_{-1}^{x} -$

Dated:

Note: If we use cdf to compute

$$P(a < x < b) = F(b) - F(a)$$

Joint pdf $f(x,y)$ for Discrete case

1) $f(x,y) \geq 0$ for all $(x,y) \in R^2$

2) $\sum_x \sum_y f(x,y) = 1$

3) $P(X = x, X = y) = f(x,y)$

for continuous case

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

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

3) $P[(X, Y) \in A] = \iint f(x,y)\, dx\, dy$

Dated:

Ex $f(x,y) = \begin{cases} \frac{2}{5}(2x+3y), & 0 \le x \le 1 \\ & 0 \le y \le 1 \\ 0 & \text{, otherwise} \end{cases}$

continuous case

a) Is $f(x,y)$ joint pdf

b) Find $P[(X,Y) \in A]$, where $A = \{(x,y) \mid 0 < x < \frac{1}{2}, \frac{1}{2} < y < \frac{1}{2}\}$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy \quad \int_{x=0}^{1}\int_{y=0}^{1} \frac{2}{5}(2x+3y)$$

Ex Box Select 2

3B 2R 3G

X = # of B

Y = # of R

	0	1	2
0			
1			
2			

$$f(x) = \frac{{}^{3}C_{x}\,{}^{2}C_{y}\,{}^{3}C_{2-x-y}}{{}^{8}C_{2}}$$

$x = 0, 1, 2$

$y = 0, 1, 2$

4.2 wp.

Dated: **Lecture#23**

The mean or expected value of X is (F(x) is a pdf)

$\mu = E(X) = \sum_{x} x f(x)$ if X is discrete

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

Ex

7

4G (Good) 3D (Defected)

select 3

X: # of G

$$f(x) = \frac{C^4_x \; C^3_{3-x}}{C^7_3}$$

$x = 0, 1, 2, 3$

$$\mu = E(X) = \sum_{x=0}^{3} x f(x)$$

$$= 0 \cdot f(0) + 1 \cdot f(1) + 2 \cdot f(2) + 3f(3) = 1.7$$

$$\frac{1.7}{3} \times 100$$

Dated: who?

Ex $f(x) = \begin{cases} \dfrac{20{,}000}{x^3}, & x > 100 \\ 0, & \text{elsewhere} \end{cases}$

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

$$= \int_{-\infty}^{100} 0 \cdot dx + \int_{100}^{\infty} x \cdot \frac{20000}{x^3}\, dx$$

$X = x$

$\downarrow$

$g(x)$

$$\mu_{g(x)} = E[g(x)] = \sum_{x} g(x) f(x)$$

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

Scanned with CamScanner

Dated:

x	4	5	6	7	8	9
$f(x)$	$1/12$	$1/12$	$1/4$	$1/4$	$1/6$	$1/6$

$$g(X) = 2x - 1$$

$$\mu_{g(x)} = E[g(x)] = \sum_{x=4}^{9} g(x) f(x)$$

$$g(4) \cdot f(4) + g(5) \cdot f(5) \; \ldots \; g(9) f(9)$$

For joint p.d.f $f(x,y)$

$$\mu_{g(X,Y)} = E[g(X,Y)] = \sum_{x} \sum_{y} g(x,y)$$

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

$f(x,y)$

		x 0	1	2	Row total
	0	$3/28$	$9/28$	$3/28$	$15/28$
y	1	$3/14$	$3/14$	0	$3/7$
	2	$1/28$	0	0	$1/28$
column sum	$5/14$	$15/28$	$3/28$		

If $g(X,Y) = XY$, find the expected value of $g(x,y)$

Dated:

$$M_{g(x,y)} = \sum_x \sum_y g(x,y) \cdot f(x,y)$$

$$= g(0,0)\, f(0,0) + \dots$$

$$= g(0,1)\, f(0,1) + \dots$$

$$= 0 \cdot (3/28) + 0 + 0 + 0$$

Date: F

Compare variance among A and B

Company A

x	$f(x)$
1	0.3
2	0.4
3	0.3

Company B

x	$f(x)$
0	0.2
1	0.1
2	0.3
3	0.3
4	0.1

$\sigma^2 = npq$

$(1)^2 = npq$

$npq = 1$

for A

$\mu_A = \sum_{x=1}^{3} x f(x)$

$= 2$

$\sigma^2 = E(X^2) - \mu^2$

$= \sum x^2 f(x) - (2.0)^2$

$\sigma_A^2 = 0.6$

for B

$\mu_B = \sum_{B}^{3} x f(x)$

$= 0.2$

$\sigma_B^2 = E(X^2) - \mu_B^2$

$= \sum_{x=0}^{4} x^2 f(x) - (0.2)^2$

$\sigma_B^2 = 1.6$

F

Date:

Discrete

$$\mu_{g(x)} = \sum g(x)\, f(x)$$

$$\sigma^2_{g(x)} = E\left[(g(x) - \mu_{g(x)})^2\right]$$

$$= \sum (g(x) - \mu_{g(x)})^2 f(x)$$

Continuos:-

$$\sigma^2_{g(x)} = \int_{-\infty}^{\infty} \left[g(x) - \mu_{g(x)}\right]^2 f(x)\, dx$$

NOW Covariance:-

$$\sigma_{xy} = E\left[(X - \mu_x)(Y - \mu_y)\right]$$

$$= \sum \sum (x - \mu_x)(y - \mu_y)\, f(x,y)$$

$$\sigma_{xy} = \sum \sum \left[xy - x\mu_y - y\mu_x + \mu_x \mu_y\right] f(x,y)$$

$$= \sum\sum xy\, f(x,y) - \mu_y \overset{\mu_x}{\boxed{\sum\sum xf}} - \mu_x \sum\sum yf$$

$$+ \mu_x \mu_y \sum\sum f = 1$$

$$\mu_{xy} - \mu_x \mu_y - \mu_x \mu_y + \mu_x \mu_y$$

$$\boxed{\sigma_{xy} = E(XY) - \mu_x \mu_y}$$

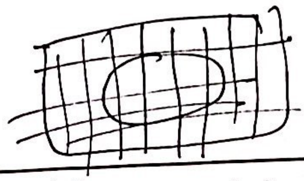
Use this

Date:

Correlation coefficient :-

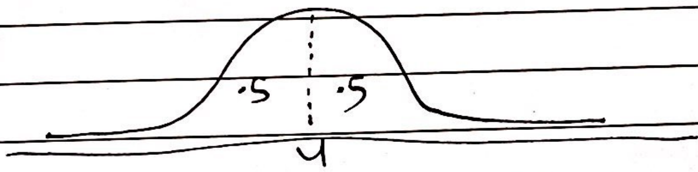
$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}.$$

x ———— x ———— x ————



Dated:

Normal probability distribution
conditions
gives a bell-shaped curve
1) The total area under the curve is 1.0
2) The curve is symmetric about the mean.
3) The two tails of the curve extends indefinitely



Find the area under the curve
$Z = 1.95$

Sol

Z	0.05
-1.4	
0	
0.1	
1.4	
1.9	✓

Area

Dated:

Ex 1 Left to $z=a$ | $P(Z \leq a)$
Area $= T(a)$

From $z=a$ to $z=b$ | $P(a < Z < b)$
Area $= T(b) - T(a)$

right to $z=a$ | $P(Z > a)$ $z=a$ $z=b$
Area $= 1 - T(a)$ | $= 1 - P(Z \leq a)$

Standardization

$$Z = \frac{x - \mu}{\sigma}$$

Dated: Mid 2 Last lecture

Example

$f(x,y)$	0	1	
0	0.5	0.2	$h(0)=0.7$
1	0.2	0.1	$h(1)=0.3$
	$g(0)=0.7$	$g(1)=0.3$	

(column heading: x)

Find correlation coefficient ρ_{xy}

Sol

since $\rho_{xy} = \dfrac{\sigma_{xy}}{\sigma_x \sigma_y}$ — ①

$\sigma_{xy} = E(XY) - \mu_x \mu_y$

$= \Sigma xy\, f(x,y) - \sqrt{a} \cdot \sqrt{b}$

$= 0.1 - \sqrt{a}\sqrt{b}$

$(0)(0)f(0,0) + (0)(1)f(0,1) + (1)(0)f(1,0) + (1)(1)f(1,1)$

$\boxed{\sigma_{xy} = c}$

plug in eq ①

$\rho_{xy} = \dfrac{c}{\sqrt{a}\sqrt{b}}$

$E(Z) = \Sigma Zf$

$E(g'(x)) = \Sigma g'(x) \cdot g(x)$

$= g'(0) \cdot g(0) + g'(1)\, g(1)$

$\sigma^2_{g'(x)} = E[E(g'(X))^2] - \mu^2_{g'(x)}$

$= [g'(0)]^2 \cdot g(0) + [g'(1)]^2 \cdot g(1)$

Here marginal distribution are

x	0	1
$g(x)$	0.7	0.3
y	0	1
$h(y)$	0.7	0.3

$\sigma_x^2 = E(x^2) - \mu_x^2$ ③

$\mu_x = \Sigma x\, g(x)$

$= 0 \cdot g(0) + 1 \cdot g(1)$

$\boxed{\mu_x = 0.3}$

$E(x^2) = \Sigma x^2 g(x)$

$= (0)^2 \cdot g(0) + (1)^2 \cdot g(1)$

$\boxed{E(x^2) = 0.3}$

plug in ③ $\sigma_x^2 = 0.3 - (0.3)^2 = a$

$\boxed{\sigma_x = \sqrt{a}}$

$\sigma_y^2 = E(y^2) - \mu_y^2$ ④

$\mu_y = \Sigma y\, h(y)$

⋮

$E(y^2) =$

plug

$\sigma_y^2 = b$

$\boxed{\sigma_y = \sqrt{b}}$

Scanned with CamScanner

Dated:

Q $f(x,y) = \begin{cases} 4xy & 0 < x, y < 1 \\ 0 & \text{elsewhen} \end{cases}$

$\rho_{xy} = \dfrac{\sigma_{xy}}{\sigma_x \sigma_y}$ ——— 1

$g(x)$