Probability and Statistics, Lec 5.

Prob Consider a construction work in campud. Let 0.6 is the probability that there will be a strike, 0.85 is the prob. the work will be completed on time if there is no strike, and 0.35 is the probability that the work will be completed on time even if there is a strike.
What is the probability that the construction work will be completed on time?

Ans Let A be the event which describes that the work will completed on time.
Let B be the event that there will be a strike.

$P(B) = 0.6$, $P(A|B^c) = 0.85$, $P(A|B) = 0.35$

$$P(A) = P[(A\cap B)\cup(A\cap B^c)]$$
$$\overset{A3}{=} P(A\cap B) + P(A\cap B^c)$$
$$\overset{\text{Cond. prob.}}{=} P(B)P(A|B) + P(B^c)P(A|B^c)$$
$$= (0.60\times 0.35) + [(1-0.60)\times 0.85]$$
$$= **$$

Observation: If a partition of events / partition of the sample space is given along with the probabilities of happening of the events.

$A = (A\cap B_1)\cup(A\cap B_2)\cup\cdots\cup(A\cap B_k)$

Rule of total probability:

If there are events $B_1, B_2, \ldots, B_k$ which constitute partition of a sample space S and $P(B_i)\neq 0$, $i=1,2,\ldots,k$ then for any event A in S,

$$P(A) = \sum_{i=1}^{k} P(B_i)\,P(A|B_i)$$

Partition of a set: Let S be a set. Then $A_1, \ldots, A_k$ constitute a partition of S if $A_i \subseteq S$, $A_i\cap A_j = \Phi$ and $\bigcup_{i=1}^{k} A_i = S$

Prob. Let IIT Kgp rent cars from three rental agencies: 60 percent of time IIT Kgp rent cars from Agency 1, 30 percent from Agency 2, and 10 percent from Agency 3. If 9 percent of the cars from Agency 1 need an oil change, 20 percent from Agency 2 need oil change, and 6 percent from Agency 3 need oil change.
What is the probability that a rented car given to IIT Kgp needs an oil change?

Ans. Let A be the event that the car needs an oil change.
Let B_1, B_2, B_3 be the events that the car comes from Agency 1, 2 or 3 respectively.

$P(B_1) = 0.6$, $P(B_2) = 0.3$, $P(B_3) = 0.1$
$P(A|B_1) = 0.09$, $P(A|B_2) = 0.2$, $P(A|B_3) = 0.06$

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)$$
$$= **$$

Bayes' theorem.

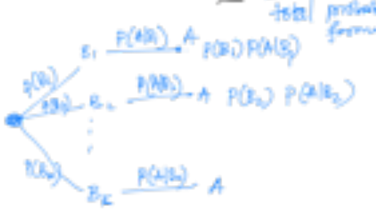
If $B_1, B_2, \ldots, B_k$ form a partition of a sample space S, and $P(B_i)\neq 0$, $1\le i\le k$ [illegible], Then for any event A with $P(A)\neq 0$,

$$P(B_r)\,P(A|B_r)$$

$$P(B_r \mid A) = \frac{\qquad}{\sum_{i=1}^{k} P(B_i)\, P(A \mid B_i)}$$

If $P(B_r \mid A) = \dfrac{P(B_r \cap A)}{P(A)}$

$= \dfrac{P(B_r)\, P(A \mid B_r)}{\text{total probability formula}}$



Prob. In the Agency problem: what is the probability that the rented car which needs an oil change ~~and~~ came from Agency 1?

Random variables:

Objective: Associate real numbers with possible outcomes and formulate the desired events through their association.
Explain the events through the number which correspond to the outcomes/sample points.

$X : S \longrightarrow \mathbb{R}$

Exp. $S = \{H, T\}$

$X : S \to \mathbb{R}$

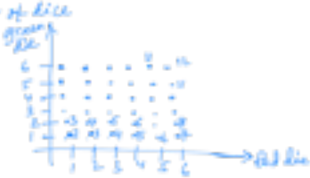
$X(H) = 2 \quad \Big| \quad X(H) = 0$

$X(T) = 1 \quad \Big| \quad X(T) = 1$

Exp. roll a pair of dice

$X : S \to \mathbb{R}$

$X(d_1, d_2) = d_1 + d_2$



Random variable: If S is the sample space corresponding to an experiment then a function $X : S \to \mathbb{R}$ is called a random variable.

Hint: measurable functions

$X^{-1}(x) = \{ s \in S \mid X(s) = x \}$
event

If $\boxed{x = 5}$, $X : S \to \mathbb{R}$

$\boxed{X = 5}$ is the short form of $\{ s \in S \mid X(s) = 5 \}$

$X = a \in \mathbb{R}$ corresponds to the event $\{ s \in S \mid X(s) = a \}$

In the previous example,

"$X = 5$" corresponds to the event $\{(1,4), (4,1), (2,3), (3,2)\}$

"$X = 9$" corresponds to the event $\{(3,6), (4,5), (5,4), (6,3)\}$

"$X = 0$" corresponds to the event Φ

... the random variable is a constant

If the ... $X : S \to \mathbb{R}$ is defined function, for example as $\underline{X(s) = 100}$ $\forall x \in S$, in the previous example of rolling a pair of dice.

Then for any $a \in \mathbb{R}$,

$\underline{X = a}$ $\underline{\text{corresponds to the event}}$ $\begin{cases} \emptyset & \text{if } a \neq 100 \\ S & \text{if } a = 100 \end{cases}$

The objective of defining a random variable is to explain the events in terms of the real valued function.

Next: How to find $P(X=a)$? ~~how~~

$P(X=x) = f(x)$ ← distribution function of the random variable.

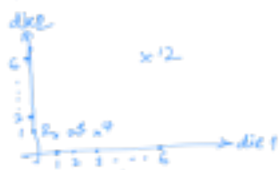
Lec-6

Test-1 on Feb 2 (12:10 – 1 pm)
45 mins
Topic: whatever we discussed before defining Random Variables.

Test-2 on Feb 28 (12:10 – 1 pm)
45 mins
Topic: Random variables and special distributions

Tests will be – "Fill in the blanks type"

Random Variables



$X : S \to \mathbb{R}$

The sample space corresponding to the example of rolling a pair of dice:

$$S = \{ (x, y) \mid 1 \leq x, y \leq 6 \}$$

$$X(s) = X((x,y)) = \boxed{x+y}, \quad s \in S.$$

$|S| = 36$. $\boxed{p(s) = \frac{1}{36}}$ $s \in S$.

Let $a \in \mathbb{R}$, $b \in \mathbb{R}$

"$X = a$" $\equiv \{ s \in S \mid X(s) = a \} \subseteq S$

"$X \leq a$" $\equiv \{ s \in S \mid X(s) \leq a \} \subseteq S$

"$X \geq a$" $\equiv \{ s \in S \mid X(s) \geq a \} \subseteq S$

"$a \leq X \leq b$" $\equiv \{ s \in S \mid a \leq X(s) \leq b \} \subseteq S$

If $a = 2$, $b = 5$ then

"$a \leq X \leq b$" $\equiv$ "$2 \leq X \leq 5$"

$\equiv \{ (1,1), (1,2), (2,1), (1,3), (1,4), (2,2), (2,3), (3,1), (4,1), (3,2) \} \subseteq S$

Q. $P(a \leq X \leq b) = ?$

Sample space – → discrete → discrete random variable
Sample space – → continuous → continuous random variable.

probability mass function (p.m.f.) / distribution function

For any $x \in \mathbb{R}$

$$P(X = x) = P(\{ s \in S \mid X(s) = x \})$$

$$= f(x) \leftarrow \text{p.m.f.}$$

Q. What is the probability of the random variable for the previous exp. of rolling a pair of dice.

x	$P(X=x)$
2	1/36
3	2/36
4	3/36
⋮	⋮

First look for those 'x' for which "$X = x$" is a non-trivial event

... $= \{ (1,1) \}$

$X = 2$ ...

"$X = 3$" $= \{(1,2),(2,1)\}$ 12 | $1/36$

"$X = 4$" $= \{(1,3),(3,1),(2,2)\}$

⋮

"$X = 12$" $= \{(6,6)\}$

$\boxed{f(x)} = P(X=x)$, $2 \le x \le 12$

$= \frac{6 - |x-7|}{36}$

Exp. Tossing of a 'unbiased' coin 4 times.

$S = \{xyzw \mid x,y,z,w \in \{H,T\}\}$

$X : S \to \mathbb{R}$

s	Prob.	x
HHHH	1/16	4
HHHT	1/16	3
		⋮

$X(s) = \#$ of 'heads' in 's'.

"$X = x$" $= \{s \in S \mid X(s) = x\}$

"$X = 0$" $= \{TTTT\}$

"$X = 1$" $= \{HTTT, THTT, TTHT, TTTH\}$

x	$P(X=x)$ = $f(x)$
0	1/16
1	4/16
2	6/16
3	4/16
4	1/16

Q. "$f(x)$" $= P(X=x) = \frac{\binom{4}{x}}{16}$, $x = 0,1,2,3,4$

$f(x) = 0$, $x \notin \{0,1,2,3,4\}$

Q. Can any function $f : \mathbb{R} \to [0,\infty)$ be a p.m.f. corresponding to some discrete random variable

Ans. No.

For finite sample space,
$S = \{s_1, s_2, \ldots, s_n\}$, $X : S \to \mathbb{R}$
$X(s_i) = x_i$, $1 \le i \le n$

$[f(x_1), f(x_2), \ldots, f(x_n)]$

$$\equiv \begin{bmatrix} P(X=x_1) \\ P(X=x_2) \\ \vdots \\ P(X=x_n) \end{bmatrix} \in \mathbb{R}^n_{\ge 0}$$

For infinite sample space $S = \{s_1, s_2, \ldots\}$, $X : S \to \mathbb{R}$, $X(s_i) = x_i$, $i \ge 1$

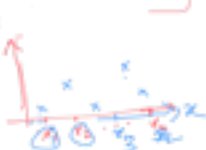
$$\text{p.m.f.} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \end{bmatrix}$$

Defn. A function $f : \mathbb{R} \to \mathbb{R}$ can represent a distribution function corresponding to a random variable if.

(1) $f(x) \ge 0$, $\forall x \in \mathbb{R}$

(2) $\sum_{x \in \mathbb{R}} f(x) = 1$

H.W. Why?



Q. Let $f(x) = \frac{x+2}{25}$, $x = 1,2,3,4,5$

... a pmf ?

Does $f(x)$ represe...

Lec-7.

Random variables (contd.)

(Exam next Tuesday, Feb 02, 2021)

Lecture videos are available in youtube now, the time is mentioned in MT chat.

— PS Ln, no. 1, 2, 3, 4, 5, 6, ...

$X : S \to \mathbb{R}$

↳ purpose is to associate numbers to sample points

explaining events as subsets of $\mathbb{R}$ through X

event

$P(\boxed{X = x}) = P(\{s \in S \mid X(s) = x\})$

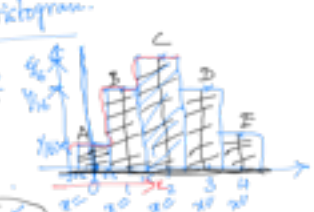
$f(x)$ = p.m.f of X

$f : \mathbb{R} \to \mathbb{R}$

① $f(x) \geq 0$ ② $\sum_{x \in \mathcal{X}} f(x) = 1$ ← range set of X

Probability Histogram.

Exp. tossing of an unbiased coin 4 times



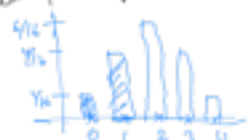
X = # of heads.

$P(X = x) = \boxed{f(x)}$

$f(0)$ = area of rectangle A

$f(4)$ = area of rectangle E.

Bar-chart:



Cumulative distribution of a random variable 'X'

$F : \mathbb{R} \to \mathbb{R}$

$f(x) = P(X = x)$

$F(x) = P(X \leq x) = P(\{s \in S \mid X(s) \leq x\})$

(c.d.f)

$= \sum_{t \leq x} f(t)$ → p.m.f.

$P(X > x) = 1 - P(X \leq x)$

Exp. $f(0) = \frac{1}{16}$, $f(1) = 4/16$

$f(2) = 6/16$, $f(3) = 4/16$

$f(4) = 1/16$

$F(0) = f(0) = 1/16$

$F(1) = f(0) + f(1) = \frac{5}{16}$

$F(2) = f(0) + f(1) + f(2) = 11/16$

$F(3) = 15/16$

$F(4) = 1$

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{16}, & 0 \le x < 1 \\ 5/16, & 1 \le x < 2 \\ 11/16, & 2 \le x < 3 \\ 15/16, & 3 \le x < 4 \\ 1, & x \ge 4 \end{cases}$$

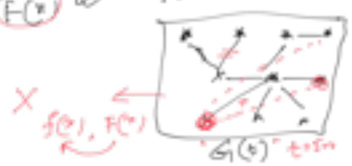
Properties of c.d.f.

① $F(\infty) = \lim_{x \to \infty} F(x) = 1$ (why)

② $F(-\infty) = 0$ (why?)

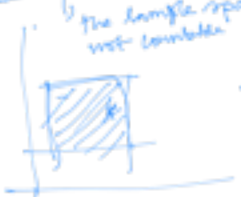
③ $a < b$. $F(a) \le F(b)$ (why?)

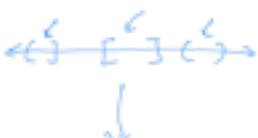
$$X \to \begin{cases} f(x) & ?? \\ F(x) & ?? \end{cases}$$



Continuous random variable

→ the sample space is not countable



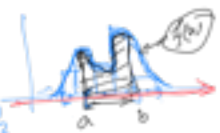


Probability density function (pdf)

Goal: Define prob. of $a < b$

For events →

$P(a \le X \le b) = P(\{\omega \mid a \le X(\omega) \le b\})$

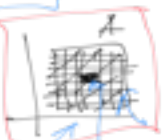


$$\boxed{P(a \le X \le b) = \int_a^b f(x)\,dx}, \quad \boxed{f : \mathbb{R} \to \mathbb{R}}$$

$X = c$, $\boxed{c \le X \le c}$

$\boxed{P(X = c) = 0}$

probability of an interval



Area of a point zero

it implies that

$P(a \le X \le b) = P(a \le X < b)$

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

$= P(a < x < b)$

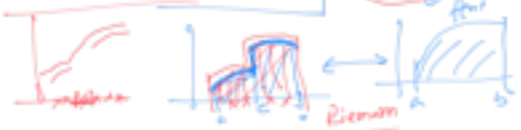
$f(x) : [a,b] \to \mathbb{R}$

$f(x) : (a,b] \to \mathbb{R}$

Q. Can any fn $f : \mathbb{R} \to \mathbb{R}$ act as a pdf for some continuous r.v.

Ans. No.

$\int_a^b \boxed{f(x)}\, dx$

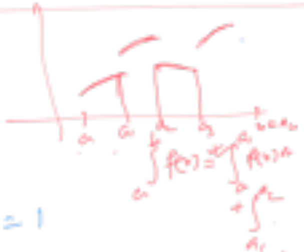


Riemann

$f(x)$ has to satisfy two conditions:

① $f(x) \geq 0$

② $\int_{-\infty}^{\infty} f(x) = 1$



Exp. $f(x) = \begin{cases} k e^{-3x}, & x > 0 \\ 0, & \text{else} \end{cases}$

Find the value of k for which $f(x)$ represent a pdf of a random variable X. Then find $\boxed{P(0.5 \leq X \leq 1)}$.

$f(x)$ X

$f(x) \to X$

Pf. $f(x) \geq 0$

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

$\int_{-A}^{A} f(x)\, dx = \int_0^{\infty} k e^{-3x}\, dx$

$= k \int_0^{\infty} e^{-3x}\, dx$

$= k \lim_{t \to \infty} \int_0^t e^{-3x}\, dx$

$= k \lim_{t \to \infty} \left. \frac{e^{-3x}}{-3} \right|_0^t = \frac{k}{3}$

$f(x)$ $g(x)$

X

$P(a \leq x \leq b)$

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

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

it implies $k = 3$.

$\boxed{P(0.5 \leq X \leq 1)} = \int_{0.5}^{1} 3 e^{-3x}\, dx$

$=$

Cumulative density / distribution fn:

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

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

$-\infty < x < \infty$

$F(-\infty) = 0, \quad F(\infty) = 1,$

$F(a) \leq F(b)$ if $a < b$.

Q. If we know $F(x)$ what can ... of certain

... are about probability ... events.

Observation-1 $\boxed{P(a \le X \le b) = F(b) - F(a)}$

$$\int_a^b f(x)\,dx \overset{?}{=} \int_{-\infty}^{b} f(x)\,dx - \int_{-\infty}^{a} f(x)\,dx$$

Observation-2 $f(x) = \frac{d}{dx}F(x)$ where the derivative exists.

Q. Calculate the c.d.f. of the pdf defined above.

$$f(x) = \begin{cases} 3e^{-3x}, & x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Ans.
$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_0^x 3e^{-3t}\,dt = -e^{-3t}\Big|_0^x = 1 - e^{-3x}.$$

$$\therefore \quad F(x) = \begin{cases} 0, & x \le 0 \\ 1 - e^{-3x}, & x > 0. \end{cases}$$

$$P(0.5 \le X \le 1) = F(1) - F(0.5)$$

Random Experiment (Sample space, set of events)

$\downarrow$

Random variable $\to$ discrete (sum), $\to$ continuous (integration)

$\downarrow$

pmf/pdf $\quad f(x) = P(X = x)$

cdf $\quad \int_a^b f(x)\,dx = P(a \le X \le b)$

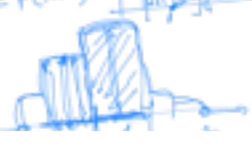
Next goal: Determine properties of random variables corresponding to these functions.

Reminder: Think of random variable as a measuring something

$$\boxed{\int_a^b f(x)\,dx = P(a \le X \le b)}$$

$a = b \quad P(c \le X \le c)$
$= P(X = c)$
$= \int_c^c f(x)\,dx = 0.$

$f(x) = P(X = x)$ 



Lec-8

Random Variables

discrete
- pmf (Histogram)
- cdf

$P(X=x) = f(x)$

$F(x) = \sum_{t \le x} f(t)$

continuous
- pdf
- cdf

$P(a \le X \le b) = \int_a^b f(x)\,dx$

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

Expected value of a random variable.

$X \rightarrow \{x_1, x_2, \ldots, x_n\}$

$P(X=x_1) = f(x_1)$
$P(X=x_2) = f(x_2)$
$\vdots$
$P(X=x_n) = f(x_n)$

Defn: If X is a discrete random variable and $f(x)$ is the pmf corresponding to X then the expected value of X is

$$E(X) = \sum_x x f(x) = \sum_x x\, P(X=x)$$

If X is a continuous random variable then

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

$$\mathrm{Avg}(x_1, x_2, \ldots, x_n) = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

$$= \left(\frac{1}{n}\right)x_1 + \left(\frac{1}{n}\right)x_2 + \cdots + \left(\frac{1}{n}\right)x_n$$

$$= P(X=x_1)\,x_1 + P(X=x_2)\,x_2 + \cdots + P(X=x_n)\,x_n$$

$Y = X^2 + .5$, for some x

$y = x^2 + 0.5$

Theorem If X is a discrete r.v. and $f(x)$ is the corresponding pmf then the expected value of $g(X)$ is

$$E(g(X)) = \sum_x g(x) f(x)$$

If X is continuous r.v. with pdf $f(x)$ then

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

Exp let X has pdf $f(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$
then find $E(g(X))$ where $g(X) = e^{3X/4}$.

Sol: $E\left(e^{3X/4}\right) = \int_0^{\infty} e^{3x/4} e^{-x}\,dx$

$$= \int_0^{\infty} e^{-x/4}\,dx$$

$$= 4$$

Formula of expectation of 'affine' function

Obs. ① $\boxed{E(aX+b) = a\,E(X) + b}$ where a and b are some constants.

Pf. $E(aX+b)$

$= \int_{-\infty}^{\infty} (ax+b) f(x)\,dx$

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

$= a\,E(X) + b$

② $E(aX) = a\,E(X)$

③ $E(b) = b$

$X : S \to \mathbb{R}$

$X(s) = b$

④ $E\left(\sum_{i=1}^{k} c_i\, g_i(X)\right) = \sum_{i=1}^{n} c_i\, E(g_i(X))$

⑤ $E\left((aX+b)^n\right) = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^{i} E(X^{n-i})$

Recall Math-II

$T : V \to W$

$T(\alpha u + v) = \alpha T(u) + T(v)$.

Obs. Expectation $E(X)$ is linear

Concept of "moments".

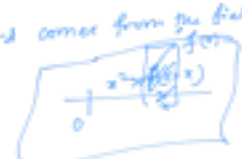
Def. The k-th moment of a discrete random variable X is given by

$$\mu_k' = E(X^k) = \sum_x x^k f(x)$$

If X is a continuous r.v. then

$$\mu_k' = E(X^k) = \int_{-\infty}^{\infty} x^k f(x)\,dx$$

— The 'moment' word comes from the field of physics



$$\mu_2' = \begin{cases} \sum x^2 f(x) \\ \int x^2 f(x) \end{cases}$$

$\mu_1' =$ does it relation with center of gravity

Note: μ_1' is also called mean of the r.v. X.

Moments "about" the mean:

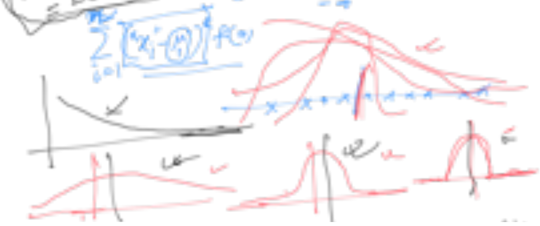
The k-th moment about the mean of a r.v. X is

the mean of X

$$\mu_k = E((X-\mu)^k) = \begin{cases} \sum_x (x-\mu)^k f(x) \\ \int_{-\infty}^{\infty} (x-\mu)^k f(x)\,dx \end{cases}$$

$\mu_1 = E(X-\mu) = E(X) - \mu = 0$

$\sum_{i=1}^{n} (x_i - \mu_1)^k f(x)$



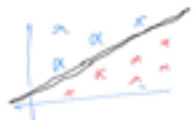
Measure of skewness / lack of symmetry

how much skew the pdf is μ_3 / σ^3

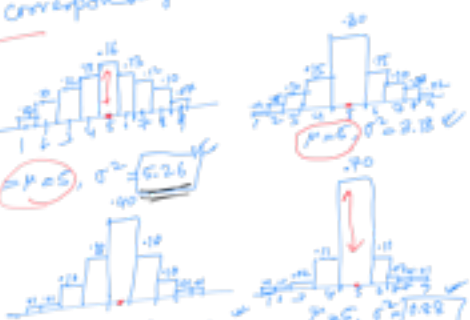
μ_4 / σ^4

Q. How to measure the spread of a pdf

Def: $\mu_2 \to$ the 2nd moment about the mean is called the variance of the random variable. It is denoted as $Var(X)$, $V(X)$ or $\boxed{\sigma^2,}$ σ_X^2



The positive square root of the variance is called standard deviation of the corresponding random variable.



$E(X) = \mu = 5$, $\sigma^2 = \boxed{5.26}$

$\mu = 6$, $\sigma^2 = 2.18$

$\mu = 5$, $\sigma^2 = 1.66$

$\mu = 5$, $\sigma^2 = \boxed{1.88}$

H.W. The small value of σ^2 means that there is a high probability the value you get is close to the mean, whereas for ~~small and~~ large value of σ^2 there is a greater probability of getting a value that is not close to the mean.

The asymmetry or skewness (lack of symmetry) of the distribution is ~~the~~ given by μ_3 / σ^3.

σ^3

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

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

Obs. $\sigma^2 = E[(X-\mu)^2]$

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

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

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

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

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

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

$E(aX) = aE(X)$

Obs. If X has variance σ^2

then $Var(aX+b) = a^2\sigma^2$

(H.H.)

If $b=0$ — $Var(aX) = a^2\sigma^2$

If $a=1$ — σ^2

$Var(aX+b) = Var(X+b) = \sigma^2$

$Var(aX) = \boxed{a^2\boxed{\sigma^2}}$

---

random variable → discrete, → continuous **Lec 9**

Expectation of a random variable

$E(X)$

Expectation of some "special" functions of a random variable:

Moments: $\mu'_k = E(\boxed{X^k})$, $k=1,2,\ldots$

Moments about the mean:

notation for "mean": μ

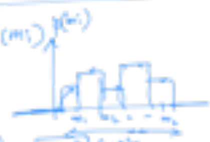
$\mu_k = E(\boxed{(X-\mu)^k})$, $k=1,2,\ldots$

Eg. Class Test-3.

$m_1, m_2, \ldots, m_k$

pmf → $\boxed{p(m_1),\ p(m_2)\ \ldots\ p(m_k)}$

$E(X) = \sum_{i=1}^{k} m_i\, p(m_i)$



variance $\sigma^2 = \mu'_2$

$= E((X-\mu)^2) = \boxed{E(X^2)} - E(X)^2$

Standard deviation: $\sigma = +\sqrt{\sigma^2}$

Imp. ① Expectation is a linear fn

$E : \boxed{\mathcal{X}} \to \mathbb{R}$

$E(aX+b) = aE(X)+b$, $a, b \in \mathbb{R}$

② Variance$(aX+b) = \boxed{a^2\sigma^2}$

Chebyshev's Theorem:

Let μ and σ be the mean and standard deviation of a random variable X. Suppose $k>0$. Then

$P(|X-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}$, $\boxed{\sigma \neq 0}$

$\mu - k\sigma < X < \mu + k\sigma$

Pf. $\sigma^2 = E((X-\mu)^2) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx$

$$= \int_{-\infty}^{\mu-k\sigma} (x-\mu)^2 f(x)\,dx + \dots + \int_{\mu+k\sigma}^{\infty} (x-\mu)^2 f(x)\,dx$$

This implies

$$\sigma^2 \geq \underbrace{\int_{-\infty}^{\mu-k\sigma} (x-\mu)^2 f(x)\,dx}_{} + \underbrace{\int_{\mu+k\sigma}^{\infty} (x-\mu)^2 f(x)\,dx}_{}$$

Since $\boxed{(x-\mu)^2 \geq (k\sigma)^2}$ when $\underbrace{x \geq \mu+k\sigma}_{}$ or $x \leq \mu-k\sigma$

Then $\boxed{\sigma^2} \geq \int_{-\infty}^{\mu-k\sigma} k^2\boxed{\sigma^2} f(x)\,dx + \int_{\mu+k\sigma}^{\infty} k^2\sigma^2 f(x)\,dx$

$$\Rightarrow \frac{1}{k^2} \geq \underbrace{\boxed{\int_{-\infty}^{\mu-k\sigma} f(x)\,dx} + \boxed{\int_{\mu+k\sigma}^{\infty} f(x)\,dx}}_{}$$

$$\boxed{\text{Remember} \quad P(a \leq X \leq b) = \int_a^b f(x)\,dx}$$

$$\Rightarrow P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

$$\Rightarrow \boxed{P(|X-\mu| < k\sigma) \geq 1-\frac{1}{k^2}}$$

$$\Rightarrow P(\mu-k\sigma < X < \mu+k\sigma) \geq 1-\boxed{\frac{1}{k^2}}$$

"Moment" generating function (mgf)

Let X be a random variable.

Then the mgf of X is defined as

$$M_X(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx} f(x), & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} e^{tx} f(x)\,dx, & \text{if } X \text{ is continuous.} \end{cases}$$

$P(X=x)$

$$e^{tx} = 1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \dots + \frac{t^kx^k}{k!} + \dots$$

If X is discrete

$$M_X(t) = \sum_x \left(1 + tx + \dots + \frac{t^kx^k}{k!} + \dots\right) f(x)$$

$$= \boxed{\sum_x f(x)} + t\boxed{\sum_x x f(x)} + \frac{t^2}{2!}\sum_x x^2 f(x) + \dots + \frac{t^k}{k!}\sum x^k f(x) + \dots$$

$$= 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \dots + \frac{t^k}{k!}E(X^k) + \dots$$

$$= 1 + t\mu + \frac{t^2}{2!}\mu_2' + \dots + \frac{t^k}{k!}\mu_k' + \dots$$

Then $\boxed{\left.\frac{d^k}{dt^k} M_X(t)\right|_{t=0} = \mu_k'}$

Exp Let $f(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$

be the pdf of a random variable X.

... of X?

The mgf ... $\int_{-\infty}^{\infty} e^{tx} e^{-x} dx$

Ex: $M_X(t) = E(e^{tX}) = \int_{0}^{\infty} e^{-x(1-t)} dx$

$= \boxed{\frac{1}{1-t}}$ for $t < 1$

Properties of mgf:

① $M_{X+a}(t) = E\left(e^{(X+a)t}\right) = e^{at} M_X(t)$

② $M_{bX}(t) = E\left(e^{btX}\right) = M_X(bt)$

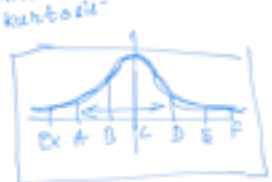
③ $M_{\frac{X+a}{b}}(t) = E\left(e^{\frac{X+a}{b}t}\right) = e^{\frac{a}{b}t} M_X\left(\frac{t}{b}\right)$.

Symmetry or lack of symmetry of a pdf can be measured by μ_3/σ^3

The extent up to which a pdf is flat or peaked can be measured by μ_4/σ^4

ζ — This measure is called "kurtosis"

H.W. How to decide grading once the total marks are known.



"Standard" pdfs. (Discrete)

① Discrete "uniform" distribution.

A random variable X is said to have discrete uniform distribution if the pmf is given by $P(X=x) = f(x) = \frac{1}{k}$, $x = x_1, \ldots, x_k$

, all distinct: $x_i \neq x_j$ $i \neq j$

② Bernoulli distribution.

A random variable X is said to have a Bernoulli distribution if the pmf is $f(x) = p^x (1-p)^{1-x}$, $\boxed{x = 0, 1}$

$0 \le p \le 1$

" $P(X=0) = 1-p$, $P(X=1) = p$ "

H.W. * Erdős–Rényi Model

$p = \frac{1}{2}$

↳ Bernoulli ~~trial~~ trial.

(3) The binomial distribution

A random variable X is said to have "binomial" distribution if

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n$$

$$\sum_{x=0}^{n} f(x) = 1$$

is called 'binomial' since the probabilities are given by the successive terms of the binomial expansion $[(1-p) + p]^n$

"X = # of successes in 'n' independent trials"

0 | 0 1 0 0 0 1

The mean or expectation of the r.v. associated with binomial distribution is — $\boxed{\mu = np}$ $\left(= \sum_{x=0}^{n} x\, f(x) = \sum x \binom{n}{x} p^x (1-p)^{n-x} \right)$

and the variance is $\boxed{\sigma^2 = np(1-p)}$.

The mgf of the binomial distribution is $M_X(t) = [1 + p(e^t - 1)]^n$

"k-th Bernoulli trial / k-th success"

0 1 0 1 0 1 0 1 0 0 0 1 0 1 0 1

'1' denotes success
'0' denotes failure

(4) "Negative" Binomial distribution.

A r.v. X has negative binomial distribution if pmf is

$$P(X=x) = f(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}, \quad x = k, k+1, k+2, \ldots$$

the parameters k, p.

probability of k-th success to occur on the x-th trial.

Explain

$X \equiv$ the # of trial on which k-th success occurs.

binomial

This is called "negative"

Look at the expansion of

$$\left[\frac{1}{p} - \frac{1-p}{p} \right]^{-k}$$

successive terms of this expansion are the probabilities.

Lec-10

Special discrete probability

Special ...

distribution functions $\quad f(x) = P(X=x)$

Uniform distribution: $\quad f(x) = \frac{1}{k},\ x \in \{x_1, \dots, x_k\}$

Bernoulli distribution: (Bernoulli trial) $\quad B(p) = f(x) = p^x (1-p)^{1-x},\ x \in \{0, 1\}$.

p = probability of success

Binomial distribution:
(independent Bernoulli trials)

$$B(n, p) = f(x) = \binom{n}{x} p^x (1-p)^{n-x},\ x \in \{0, 1, \dots, n\}$$

Negative binomial distribution:
(Pascal distribution)

$$\overline{B}(k, x, p) = f(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k},\ x \in \{k, k+1, k+2, \dots\}$$

$k \to$ # of successes

$x \to$ index for the trial

Geometric distribution ($\boxed{k=1}$ in negative binomial distribution)

$$G(x; p) = p(1-p)^{x-1},\ x \in \{1, 2, 3, \dots\}$$

$\boxed{0\ 0\ 1\ 0\ 0 - 0 \cdots}$

Hypergeometric distribution.

Assumption: the trials are Not independent

100 times

$0\,1\,0\,0\,1\,0\,0\,1 \cdots \cdots 0\,1 \cdots 0$

$\boxed{2^{100}}$

Consider a set of $\boxed{N}$ elements.
Suppose $\boxed{M}$ out of these N elements are considered as successes, which means $\boxed{N-M}$ are considered as failures.

?? [We are interested in the probability of getting $\boxed{x}$ successes in "n" trials.

From above we have the following observations:

$A = \boxed{\binom{M}{x}}$ = ways of choosing x of the M successes.

$B = \binom{\boxed{N-M}}{n-x}$ = ways of choosing $n-x$ of the $N-M$ failures.

(obs.) Thus $\boxed{\binom{M}{x}\binom{N-M}{n-x}}$ is the number of ways of x successes and $n-x$ failures.

$\boxed{n=2 \\ x=2}$

$\boxed{1} \to$ success

$\bigcirc \to$ failure

# of choose ways choosing n elements that of N elements $\binom{N}{n}$

Lec-11

Special discrete (time) random variables
- Uniform
- Bernoulli
- Binomial
- Negative binomial | Pascal
- Geometric
- Hypergeometric

Percentile for random variable

Defn The pth percentile of a random variable X is the value $x_p \in \mathbb{R}$ that satisfies

$$\boxed{P(X \le x_p) = \boxed{\frac{p}{100}}}$$

Which means: The $100p$th percentile is a measure of location for the probability distribution in a sense that x_p divides the distribution into two parts: one having probability mass/density 'p' and the other having probability mass $1-p$

$P(X \le x_p) \le p$

Exp Suppose the random variable X has the density function

$$f(x) = \begin{cases} e^{x-2} & \text{for } x < 2 \\ 0 & \text{otherwise} \end{cases}$$

What is the 75th percentile of X?

Ans. Using the definition

$$\boxed{\frac{75}{100}} = 0.75 = P(X \le x_p)$$

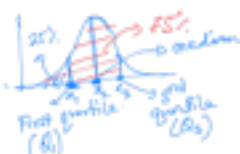
$$= \int_{-\infty}^{x_p} f(x)\,dx = \int_{-\infty}^{x_p} e^{x-2}\,dx$$

$$= e^{x-2}\Big|_{-\infty}^{x_p} = e^{x_p-2}$$

$$\Rightarrow \boxed{x_p = 2 + \ln\frac{3}{4}}$$

Defn The 25th and 75th percentiles of any distribution are called the first and the third quartiles respectively.

Q_1: is a number such that 25% of the observations are less than it

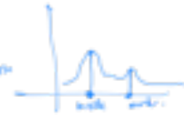


Q_3: is a number such that 75% observations are less than it.

Defn The 50th percentile of any distribution is called the '<u>median</u>'.

Which means: The median of the distribution divides the probability mass/density into two equal parts.

Another measure



A 'mode' of the distribution of a continuous random variable X is the value of x where the pdf attains a relative/ local maximum.

$f(x) =$ pdf.

Which means: A mode of a r.v. X is one of the most probable values. And a pdf can have infinitely many modes.

Syllabus for Class Test - 2, Tuesday Feb 22

Random variable, pmf, pdf, cdf moments and mgf, median and quantiles, Chebyshev's inequality, special discrete distribution (excluding Poisson)

"Stochastic process."

A strochastic process is a mathematical model of a probabilistic experiment that evolves in time and generate a sequence of numerical values.

For example,

1. the sequence of failure times of a machine
2. The sequence of daily price of a stock
3. the sequence of scores in a cricket match
4. the sequence of # of friends

Each numerical value in the seq. can be modeled as a random variable, and hence stochastic process is a sequence of random variables.

i.i.d. — independent and identically distributed random variables.

$\boxed{X_1}, \boxed{X_2}, \boxed{X_3}, \ldots, X_n, \ldots$

Stochastic Processes

- "Arrival"-Type Processes:

Interarrival times are distributed

Bernoulli Process

Poisson Process

- "Markov Processes:"

there is a probabilistic dependence on the past.

In Markov Process the value depends on past next value ... the current values through the current value

$X_1 \quad \boxed{X_2} \quad \boxed{X_3} \; \cdots\cdots$

<u>Bernoulli Process</u>

A sequence of Bernoulli trials.

$X_1, X_2, X_3, \cdots\cdots$

$X_i \sim Ber(p)$ $\qquad P(X_i = 1) = P(\text{success at } i\text{th trial}) = p$

$P(X_i = 0) = P(\text{failure at } i\text{th trial}) = 1-p.$

X_i's are <u>i.i.d.s</u>

X_i's are "Memoryless"

Consider two new random variables associated with Bernoulli Process

Y_K —— the time of K-th success

"T_K" —— K-th interarrival time which represents the number of trials following (K-1)st success until Kth success

$Y_K = \underbrace{T_1 + T_2 + \cdots + T_K}$

<u>Previously we observed:</u>

$T_1 \rightarrow$ the time until the first success is geometrically distributed

$T_2 \rightarrow$ geometrically distributed and it is independent of T_1

$\vdots$

$$\boxed{T_1, T_2, T_3, T_4, \cdots}$$

interarrival times: these are independent and geometrically distributed

<u>Suppose K-th arrival Time.</u>

$Y_K = \boxed{T_1} + \boxed{T_2} + \cdots + T_K$ which is a sum of <u>independent identically distributed random variables</u>.

$$E(Y_K) = E(T_1) + E(T_2) + \cdots + E(T_K) = \frac{1}{p} + \frac{1}{p} + \cdots + \frac{1}{p} = \frac{K}{p}$$

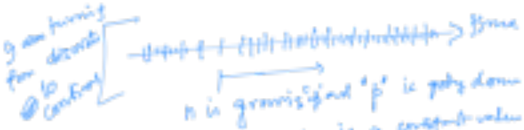
$$Var(Y_K) = K\frac{(1-p)}{p^2}$$

$$P_{Y_K}(t) = \binom{t-1}{K-1} p^K (1-p)^{t-K}, \quad t = K, K+1, \ldots$$

negative binomial / Pascal

Recall binomial: the # of success in n independent Bernoulli trials with mean "np".

Focus: "n is large but p is small"
the mean 'np' → constant value.



n is growing and 'p' is going down such that np is a constant value

Let $np = '\lambda'$

Poisson distribution.

The distribution of the Poisson random variable X with parameter $\boxed{'\lambda'}$ is given by

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

$E(X) = \lambda$, $Var(X) = \lambda$

Lec-12

Stochastic Processes

Arrival-type $\quad$ $\boxed{\text{Markov}}$

↳ Bernoulli process

↳ Poisson process
(continuous-time analog of the Bernoulli process)

Exp. traffic accidents in a city
Success – at least one traffic accident in a minute

Assumption: traffic intensity is constant over time and successive time are independent.

When we increase the # of observations 'n' and 'p' → 0

$\boxed{\lambda = np}$ → constant, $\boxed{n \to \infty, p \to 0}$
↳ "rate of arrival"

$p = \lambda/n$

$$B(k; n, p) = \frac{n!}{(n-k)!\,k!} p^{k} (1-p)^{n-k}$$

$$= \frac{n(n-1)\cdots(n-k+1)}{k!} \cdot \frac{\lambda^{k}}{\boxed{n^{k}}} \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{(n-k+1)}{n} \cdot \boxed{\frac{\lambda^{k}}{k!}} \left(1 - \frac{\lambda}{n}\right)^{\boxed{n-k}}$$

Observe that 'k' is fixed. $n \to \infty$

$1 \quad 1-\frac{1}{n} \quad 1-\frac{k-1}{n}$

$$\left(1-\frac{\lambda}{n}\right)^{n-k} = \left(1-\frac{\lambda}{n}\right)^{-k} \left(1-\frac{\lambda}{n}\right)^{n}$$

$$\left(1-\frac{\lambda}{n}\right)^{-k} \to 1, \quad \left(1-\frac{\lambda}{n}\right)^{n} \to \boxed{e^{-\lambda}}$$

$$\therefore \quad B(k; n, p) = f(n,p) \longrightarrow e^{-\lambda}\frac{\lambda^k}{k!}$$

$k = 0, 1, 2, \ldots$

~~Bernoulli~~
Poisson distribution can be approximated by Binomial distribution when $n \to \infty, p \to 0$.

Given, n, p and k

$B(k; n, p) =$ Binomial $\approx f(k; \lambda) =$ Poisson

Rule of thumb:

$$e^{-\lambda}\frac{\lambda^k}{k!} \approx \frac{n!}{(n-k)!\,k!} p^k (1-p)^{n-k}, \quad \boxed{k = 0, 1, \ldots, n}$$

if $n \geq 100$, $p \leq 0.01$, and $\lambda = np$.
Then Binomial closely approximates the Poisson.

MGF of Poisson.

$$M_{Poisson}(t) = e^{\lambda(e^t - 1)},$$

$$\mu = \lambda, \quad \sigma^2 = \lambda$$

	Poisson	Bernoulli
Time of arrival	Continuous	Discrete
PMF of number of arrivals	Poisson	Binomial
Interarrival/waiting time CDF	$\boxed{\text{Exponential}}$	Geometric
Arrival rate	λ per unit time	p per unit time.

Definition of Poisson Process

Let $p(k; \tau) = P$(there are exactly k arrivals during an interval of length τ)

$$\frac{T_1}{\tau} \quad \frac{T_2}{\tau} \quad \frac{T_3}{\tau} \quad \frac{T_4}{\tau} \quad \cdots$$

$\lambda \longrightarrow$ arrival rate or intensity

Then the arrival process is said to be a Poisson process with rate λ if it follows the following:

ⓐ (Time-homogeneity) The probability $P(k, \tau)$ of k arrivals is same for all intervals of same length τ

ⓑ (Independence) The number of ... interval in ...

arrivals ... the history of independent of arrivals outside this interval

(2) (Small interval probabilities) The probabilities $\boxed{P(k,\tau)}$ satisfy

$$\begin{cases} P(0,\tau) = 1 - \lambda\tau + \boxed{o(\tau)} \\ P(1,\tau) = \boxed{\lambda\tau} + \boxed{o_1(\tau)} \\ P(k,\tau) = \boxed{o_k(\tau)} \quad k = 2,3,\ldots \end{cases}$$

where $o(\tau)$ and $o_k(\tau)$ are functions of τ that satisfy

$$\boxed{\lim_{\tau \to 0} \frac{o(\tau)}{\tau} = 0}, \quad \boxed{\lim_{\tau \to 0} \frac{o_k(\tau)}{\tau} = 0}$$

~~Class Test — Saturday~~, Mandatory. Tomorrow 25th Feb. NO class Test.

Probability density functions of special continuous random variables.

$$P(a \le X \le b) = \boxed{\int_a^b \boxed{f(x)}\,dx}$$

density.

$\boxed{X} : S \to \mathbb{R}$

$\{ w \in S \mid a \le X(w) \le b \}$

(1) Uniform distribution ~~Suppose~~ A continuous r.v. X has a uniform distribution if

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

where $a, b \in \mathbb{R}$.

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \frac{a+b}{2}$$

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx = E(X^2) - \left[E(X)\right]^2 = \frac{1}{12}(b-a)^2.$$

(2) Now we are interested in the pdfs which are of the following type:

$$f(x) = \begin{cases} k\,x^{\alpha-1} e^{-x/\beta} & \text{for } x > 0 \\ 0, & \text{elsewhere } x \le 0 \end{cases}$$

where $\alpha > 0$, $\beta > 0$ and we can find the value of k such that this happens.

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

Then $\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{\infty} K x^{\alpha-1} e^{-x/\beta}\,dx$

$= K\beta^{\alpha} \underbrace{\int_0^{\infty} y^{\alpha-1} e^{-y}\,dy}$ where $y = \frac{x}{\beta}$

<u>Recall:</u> Gamma function

$$\boxed{\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y}\,dy} \quad \boxed{\alpha > 0}$$

and also recall,

$$\boxed{\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)} \text{ when } \boxed{\alpha > 1}$$

and $\boxed{\Gamma(1) = \int_0^{\infty} e^{-y}\,dy = 1}$

$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$$
$$= (\alpha-1)(\alpha-2)\Gamma(\alpha-2)$$
$$= (\alpha-1)! \quad \text{when } \alpha \text{ is a positive integer}$$

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

Then $\int_{-\infty}^{\infty} f(x)\,dx = \int_0^{\infty} K\beta^{\alpha} x^{\alpha-1} e^{-x/\beta}\,dx$ $\quad y = x/\beta$

$= K\beta^{\alpha} \int_0^{\infty} x^{\alpha-1} e^{-y}\,dy$

$= K\beta^{\alpha}\Gamma(\alpha).$

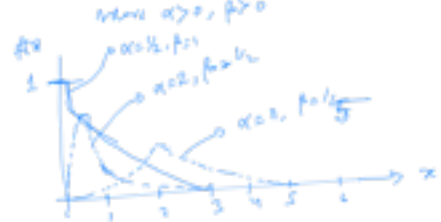
$= 1$

$$\Rightarrow \boxed{K = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}}$$

<u>Gamma distribution:</u> A r.v. X is said to have Gamma distribution if

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

where $\alpha > 0$, $\beta > 0$



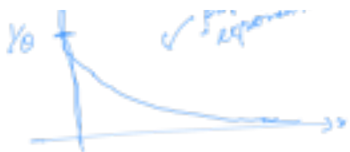
(3) Setting $\alpha = 1$, "$\beta = \theta$"

Then the Gamma is called the exponential distribution:

Def: A r.v. X is said to have an exponential distribution if

$$f(x;\theta) = \begin{cases} \dfrac{1}{\theta} e^{-x/\theta}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

where $\theta > 0$



Suppose the probability of getting $\boxed{x}$ successes during a time interval of length $\boxed{t}$ $\left(= p(x,t) \right)$ ~~...~~ denotes the dimension of Poisson probability

(1) the prob. of a success

Lec-13

Gamma distribution.

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, & \boxed{x>0} \\ 0 & \text{otherwise} \end{cases}$$

$\alpha > 0$, $\beta > 0$.

Exponential. $\alpha = 1$, $\beta = \theta$

$$f(x;\theta) = \begin{cases} \dfrac{1}{\theta} e^{-x/\theta} & \boxed{x>0} \\ 0, & \text{otherwise} \end{cases}$$

Chi-square. $\alpha = \dfrac{\nu}{2}$, $\beta = 2$

$$f(x;\nu) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\frac{\nu-2}{2}} e^{-x/2}, & x>0 \\ 0, & \text{otherwise} \end{cases}$$

ν = the number of "degrees of freedom"

Thm. The rth moment about origin of Gamma distribution

$$\mu_r' = \frac{\beta^r \Gamma(\alpha+r)}{\Gamma(\alpha)}$$

Pf.

$$\mu_r' = \int_0^\infty x^r \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}\,dx$$

Let $y = \frac{x}{\beta}$

$$= \int_0^\infty \frac{\beta^r}{\Gamma(\alpha)} \boxed{y^{\alpha+r-1}} \boxed{e^{-y}}\, dy.$$

$$= \frac{\beta^r}{\Gamma(\alpha)} \Gamma(r+\alpha) \quad \text{--- } (*)$$

From (*), setting $r=1$,

$$\mu_1' = \frac{\beta \boxed{\Gamma(\alpha+1)}}{\Gamma(\alpha)} = \alpha\beta$$

$$\boxed{\mu_2'} = \frac{\beta^2}{\Gamma(\alpha)} \Gamma(2+\alpha)$$

$$= \alpha(\alpha+1)\beta^2.$$

$$\sigma^2 = \alpha(\alpha+1)\beta^2 - (\alpha\beta)^2 = \alpha\beta^2$$

$$\boxed{\mu = \alpha\beta}$$

For exponential dist.

$\mu = \theta$, $\sigma^2 = \theta^2$

Chi-square $\mu = \nu, \ \sigma^2 = 2\nu$

Mgf. of gamma

$$M_X(t) = (1-\beta t)^{-\alpha}$$

Beta distribution. pdf is

$$f(x;\alpha,\beta) = \begin{cases} \boxed{\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)}}\, x^{\alpha-1}(1-x)^{\beta-1}, & 0<x<1 \\ 0, & \text{otherwise} \end{cases}$$

$\alpha > 0, \ \beta > 0$

Why Beta.

$$\int_{-\infty}^{\infty} f(x;\alpha,\beta) = 1$$

$$\Rightarrow \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\,dx = 1$$

$$\Rightarrow \underbrace{\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx}_{\text{Beta function } B(\alpha,\beta)} = \boxed{\frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)}}$$

$$\mu = \frac{\alpha}{\alpha+\beta}$$

$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Normal distribution

$$N(\mu,\sigma^2) = f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty<x<\infty$$

Mgf.

$$M_X(t) = \int_{-\infty}^{\infty} e^{xt} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \boxed{\int_{-\infty}^{\infty}} e^{-\frac{1}{2\sigma^2}\left[-2xt\sigma^2 + (x-\mu)^2\right]} dx$$

$$-2xt\sigma^2 + (x-\mu)^2 = \left[x-(\mu+t\sigma^2)\right]^2 - 2\mu t\sigma^2 - t^2\sigma^4$$

$$M_X(t) = e^{\mu t + \frac{1}{2}t^2\sigma^2}\left[\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[\frac{x-(\mu+t\sigma^2)}{\sigma}\right]^2} dx\right]$$

pdf. $N(\mu + t\sigma^2, \sigma)$

$$= e^{\mu t} \cdot e^{\frac{1}{2}\sigma^2 t^2}$$

Then $M_X'(t) = (\mu + \sigma^2 t) M_X(t)$

$M_X''(t) = [(\mu + \sigma^2 t)^2 + \sigma^2] M_X(t)$

For $t = 0$ $\boxed{M_X'(0) = \mu}$

$M_X''(0) = \mu^2 + \sigma^2$.

$$\text{Var}(X) = M_X''(0) - (M_X'(0))^2$$
$$= \mu^2 + \sigma^2 - \mu^2$$
$$= \sigma^2$$

Standard Normal distribution.

$N(\mu = 0, \sigma = 1) = N(0, 1)$.

$Z \to$ random variable.

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

Thm. If X has normal distribution with mean μ & s.d. σ i.e. $X \sim N(\mu, \sigma^2)$

Then 'Z' $= \boxed{\frac{X - \mu}{\sigma}} = \boxed{\frac{X}{\sigma}} - \boxed{\frac{\mu}{\sigma}}$

has standard normal dist.

Pf. $z_1 = \frac{x_1 - \mu}{\sigma}$, $z_2 = \frac{x_2 - \mu}{\sigma}$.

Claim. $P(x_1 < X < x_2)$
$= P(z_1 < Z < z_2)$

$$P(x_1 < X < x_2) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{x_1}^{x_2} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{1}{2} z^2} dz$$

$$= \int_{z_1}^{z_2} N(z; 0, 1)\, dz$$

$$= P(z_1 < Z < z_2).$$

Prob. 'X' $\sim N(\boxed{3}, \boxed{16})$ then what is $P(\underbrace{4 \le X \le 8}) $?

Soln. $P(\underbrace{4 \le X \le 8})$

$$= P\left(\frac{4-3}{4} \le \boxed{\frac{X-3}{4}} \le \frac{8-3}{4}\right)$$

$$= P\left(\tfrac{1}{4} \le Z \le \tfrac{5}{4}\right)$$

$$= \boxed{P(Z \le 1.25)} - \boxed{P(Z \le 0.25)}$$

$$= 0.8944 - 0.5987$$

$$= \boxed{0.2957}$$

<u>Thm.</u> $X \sim N(\mu, \sigma^2)$ then

$$\left(\frac{X-\mu}{\sigma}\right)^2 \sim \chi^2(1)$$

<u>Pf</u> $\quad W = \left(\frac{X-\mu}{\sigma}\right)^2$

$$w = \left(\frac{x-\mu}{\sigma}\right)^2$$

<u>Claim.</u> $g(w) = \begin{cases} \frac{1}{\sqrt{2\pi w}} e^{-\frac{1}{2}w}, & 0<w<\infty \\ 0 & \text{otherwise} \end{cases}$

$$\begin{aligned} G(w) &= P(W \le w) \\ &= P\left(\left(\frac{X-\mu}{\sigma}\right)^2 \le w\right) \\ &= P\left(-\sqrt{w} \le \boxed{\frac{X-\mu}{\sigma}} \le \sqrt{w}\right) \\ &= P\left(-\sqrt{w} \le \boxed{Z} \le \sqrt{w}\right) \\ &= \int_{-\sqrt{w}}^{\sqrt{w}} f(z)\, dz \end{aligned}$$

$f(z)$ → pdf. of standard normal.

$$\begin{aligned} g(w) &= \frac{d}{dw} G(w) \\ &= \frac{d}{dw} \int_{-\sqrt{w}}^{\sqrt{w}} f(z)\, dz \\ &= f(\sqrt{w}) \frac{d}{dw}\sqrt{w} - f(-\sqrt{w}) \frac{d(-\sqrt{w})}{dw} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}w} \frac{1}{2\sqrt{w}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}w} \frac{1}{2\sqrt{w}} \\ &= \boxed{\frac{1}{\sqrt{2\pi w}} e^{-\frac{1}{2}w}} \end{aligned}$$

<u>Lognormal distribution.</u>

this is the distribution of a random variable whose logarithm ... distribution.

is normally distributed

— distribution of stars in the universe

— dist. of size of income

(Cobb - Douglas distribution)

p.d.f.

$$f(x) = \begin{cases} \dfrac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2} \\ 0, & \text{otherwise} \end{cases}$$

$-\infty < \mu < \infty$, $0 < \sigma^2 < \infty$ are parameters.

$$X \sim \Lambda(\mu, \sigma^2)$$

Thm. $E(X) = e^{\mu + \frac{1}{2}\sigma^2}$

$Var(X) = [e^{\sigma^2} - 1] e^{2\mu + \sigma^2}$

Weibull distribution.

$$f(x) = \begin{cases} k\, \alpha\, x^{\beta-1} e^{-\alpha x^{\beta}}, & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

$\alpha > 0$, $\beta > 0$.

$\beta = 1$ then it becomes exponential

$\mu = \alpha^{-1/\beta}\, \Gamma\left(1 + \frac{1}{\beta}\right)$.

Cauchy distribution.

$$f(x) = \boxed{\frac{\beta/\pi}{(x-\alpha)^2 + \beta^2}} \quad -\infty < x < \infty$$

Here μ, σ^2 undefined.

Multivariate distributions. Lec-14

(bivariate)

Sample space, Probability measure.

univariate $P(X \leq x)$, $P(a \leq X \leq b)$

$S = \{(a,b) \mid 1 \leq a \leq 6,\ 1 \leq b \leq 6\}$

$X : S \rightarrow \mathbb{R}$

$(a,b) = a + b$

$X((a,b)) =$

$Y((a,b)) = a-b$

X = total marks in this course

Y = — —

Y/X → bivariate

$\begin{bmatrix} X(s) \\ Y(s) \end{bmatrix}$, $[X(s)\ Y(s)]$

"(X, Y)" → bivariate

Joint probability distribution.

When X and Y are discrete random variables then the 'joint' probability distribution is defined

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

If X, Y are continuous random variables, the $f(x,y)$ is a joint density fn. of (X,Y) iff

$$P((X,Y) \in A) = \iint_A f(x,y)\, dx\, dy$$

Conditions for a function $f(x,y)$ to be pdf.

Discrete:
① $f(x,y) \geq 0$
② $\sum_x \sum_y f(x,y) = 1$

Continuous:
① $f(x,y) \geq 0$
② $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, dx\, dy = 1$

Joint distribution fn.

$$F(x,y) = \begin{cases} P(X \leq x, Y \leq y) = \sum_{s \leq x} \sum_{t \leq y} f(s,t), & -\infty < x < \infty,\ -\infty < y < \infty \quad \text{for discrete } (X,Y) \\ \int_{-\infty}^{y}\int_{-\infty}^{x} f(s,t)\, ds\, dt, & -\infty < x < \infty,\ -\infty < y < \infty \end{cases}$$

Exp. Suppose

$$f(x,y) = \begin{cases} \frac{6}{5}x(y+x), & 0<x<1,\ 0<y<1 \\ 0, & \text{otherwise.} \end{cases}$$

is the joint pdf of (X,Y).

Find $P\left((X,Y)\in A\right)$

where $A = \{(x,y) \mid 0<x<\tfrac{1}{2},\ 1<y<2\}$

Sol:

$$P\left((X,Y)\in A\right) = P\left(0<X<\tfrac{1}{2},\ 1<Y<2\right)$$

$$= \int_1^2 \int_0^{1/2} \frac{3}{5}\, x(y+2x)\, dx\, dy$$

$$= \int_1^2 \left(\frac{3x^2 y}{10} + \frac{2x^3}{5} \Big|_0^{1/2} \right) dy$$

$=$

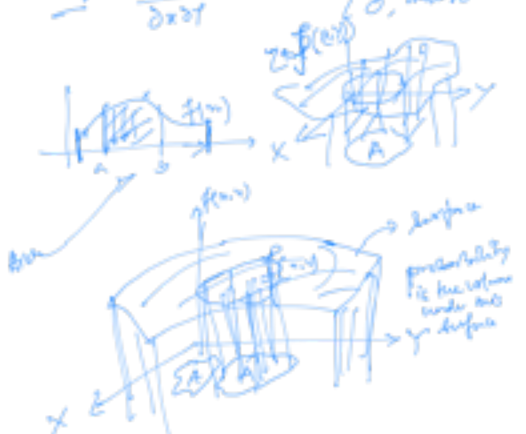
$$f(x) = \frac{d\, F(x)}{dx}$$

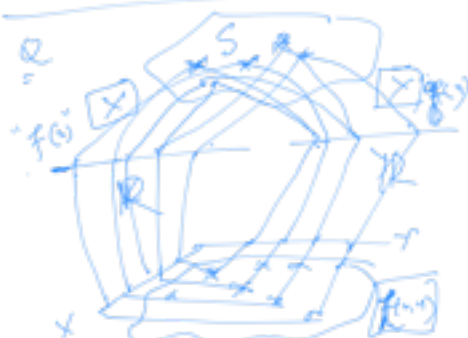
$$f(x,y) = \boxed{\frac{\partial^2}{\partial x \partial y}} F(x,y) \quad \text{if partial derivatives exist.}$$

Exp. $F(x,y) = \begin{cases} (1-e^{-x})(1-e^{-y}), & x>0,\ y>0 \\ 0, & \text{otherwise} \end{cases}$

Find p.d.f. and determine $P(1<X<3,\ 1<Y<2)$.

Sol. $\dfrac{\partial^2}{\partial x \partial y} F(x,y) = \begin{cases} e^{-(x+y)}, & x>0, y>0 \\ 0, & \text{otherwise} \end{cases}$

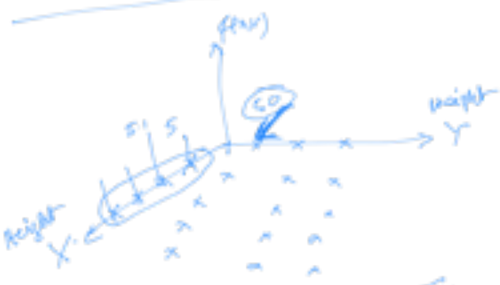




How $h(x,y)$ is related to $f(x)$ and $g(y)$

if I know $f(x)$ and $g(y)$, can we determine $h(x,y)$

→ if I know $h(x,y)$ can we determine f & g



Joint pdf / $\boxed{\text{marginal.}}$

Conditional distribution.

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$

$P(A)$

$$P(X = x \mid \boxed{Y = y}) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

$$= \frac{f(x,y)}{h(y)}$$

Conditional distribution:

Let $f(x,y)$ be the joint pdf for discrete (x,y) at (x,y), $h(y)$ is the marginal $\boxed{\text{distribution}}$ (density) of Y at y. Then

$$f(x|y) = \frac{f(x,y)}{h(y)}, \quad h(y) \neq 0$$

For each x, is known as the conditional $\boxed{\text{distribution}}$ (density) of X given $\boxed{Y = y}$

$$\text{Illy,} \quad u(y|x) = \frac{f(x,y)}{g(x)}, \quad g(x) \neq 0$$

$f(x,y) = \begin{cases} \frac{2}{3}(x+2y), & 0 < x < 1, \\ & 0 < y < 1 \end{cases}$

Exp. $f(x,y) = \begin{cases} \ldots \\ 0, & \text{otherwise} \end{cases}$

find $P(X \le \frac{1}{2} \mid Y = \frac{1}{2})$

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

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

$h(y) = \int_0^1 \frac{2}{3}(x+2y)\,dx = \frac{1}{3}(1+4y)$.

$g(x) = \begin{cases} \frac{2}{3}(x+1), & 0<x<1 \\ 0, & \text{otherwise} \end{cases}$

$h(y) = \begin{cases} \frac{1}{3}(1+4y), & 0<y<1 \\ 0, & \text{otherwise} \end{cases}$

$f(x|y) = \frac{f(x,y)}{h(y)} = \frac{\frac{2}{3}(x+2y)}{\frac{1}{3}(1+4y)}$

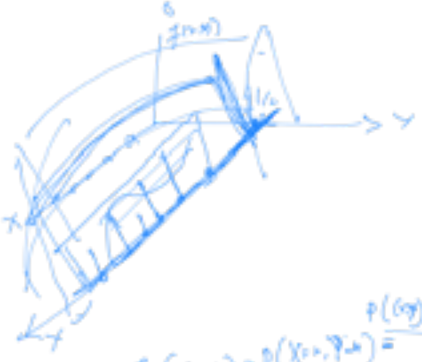
$= \frac{2x+4y}{1+4y}, \quad 0<x<1$

Now $f(x|\frac{1}{2}) = \frac{2x+4\cdot\frac{1}{2}}{1+4\cdot\frac{1}{2}}$

$= \frac{2x+2}{3}$

$P(X \le \frac{1}{2} \mid y = \frac{1}{2})$

$= \int_0^{1/2} \frac{2x+2}{3}\,dx = \frac{5}{12}$



$P((x,y)=(x,y))$

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

$g(x) = P(X=x)$, $h(y) = P(Y=y)$

$\boxed{f(X|y)}$, $f(Y|x)$

$= P(X=x \mid Y=y)$ $\quad = P(Y \mid X=x)$.

Independence of random variable.

$(X,Y) \longrightarrow$ pair of r.v.

$f(x,y)$, $g(x)$, $h(y)$

Then X, Y are independent if and only.

$g(x)\; h(y)$

$f(x,y) = g(x, \dots)$

$\downarrow$

$\underbrace{P(X=x, Y=y)} = P(X=x)\, P(Y=y)$

Suppose, $(X_1, X_2, \dots, X_n)$ is an n-tuple of random variables. Then $X_1, X_2, \dots, X_n$ are independent if $f(x_1, x_2, \dots, x_n) = f_1(x_1) \cdots f_n(x_n)$

Expectation of (a function of a pair of random variables) a pair of r.v.s.

$$E(X) = \sum_x x f(x)$$

$$E(g(X)) = \sum_x g(x) f(x)$$

$$E[g(X,Y)] = \sum_x \sum_y g(x,y) f(x,y)$$

is the joint probability distribution fn.

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

Moments (about the origin)

$\mu_r' = E(X^r) = \sum x^r f(x)$

$(X,Y) = \{(X(\omega), Y(\omega)) \mid \omega \in S\}$

The rth and sth 'product' moment about origin of the r.v. X & Y

$$\mu_{r,s}' = E(X^r Y^s)$$

$$= \begin{cases} \sum_x \sum_y x^r y^s f(x,y) & \text{discrete pair } (X,Y) \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^r y^s f(x,y)\, dx\, dy & \text{when } (X,Y) \text{ continuous.} \end{cases}$$

As a special case,

$\mu_{1,0}' = E(X) = \boxed{\mu_X}$, $\quad \mu_{0,1}' = E(Y) = \mu_Y$

Product moment about the mean

The rth & sth product moment about the means of the r.v. X and Y, is

$$\mu_{r,s} = E\left[\underbrace{(X-\mu_x)^r (Y-\mu_y)^s}\right]$$

$$= \sum_x \sum_y (x-\mu_x)^r (y-\mu_y)^s f(x,y)$$

$r = 0, 1, 2, \ldots$

$s = 0, 1, 2, \ldots$

when X, Y are discrete

$$\mu_{r,s} = E\left[(X-\mu_x)^r (Y-\mu_y)^s\right]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_x)^r (y-\mu_y)^s f(x,y)\,dx\,dy$$

$\boxed{\text{When } r = 1, s = 1}$

Then $\mu_{1,1}$ is called the covariance of X and Y denoted by $\operatorname{cov}(X,Y)$, σ_{xy}, $\sigma(X,Y)$

Note:

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

$$= E\left[XY - X\mu_y - Y\mu_x + \mu_x\mu_y\right]$$

$$= E(XY) - \mu_y E(X) - \mu_x E(Y) + \mu_x\mu_y$$

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

$$= \boxed{\mu'_{1,1} - \mu_x\mu_y}$$

Exp.

		X: 0	1	2	
Y	0	1/6	1/3	1/12	7/12
	1	2/9	1/6		7/18
	2	1/36			1/36
		7/12	1/2	1/12	

Find the $\operatorname{cov}(X,Y)$ // σ_{xy}

Sol^n: $\boxed{E(XY) = \sum_x \sum_y xy\, f(x,y)}$

$$= 0.0.\tfrac{1}{6} + 1.0.\tfrac{1}{3} + 2.0.\tfrac{1}{12} + 0.1.\tfrac{2}{9} + 1.1.\tfrac{1}{6} + 0.2.\tfrac{1}{36}$$

$$= 1/6$$

$$\mu_X = 0.\tfrac{5}{12} + 1.\tfrac{1}{2} + 2.\tfrac{1}{12} = \tfrac{2}{3}$$

$$\mu_Y = -$$

Obs. If X and Y are independent then $E(XY) = E(X)\,E(Y) = \mu_X \mu_Y$

and hence $Cov(X,Y) = \sigma_{XY} = 0$

Pf. $E(XY) = \sum_x \sum_y xy\, \boxed{f(x,y)}$

However the converse need not be true i.e. the zero covariance does not necessarily imply that the corresponding r.v.s are independent.

Ex.

		X			
		-1	0	1	
Y	-1	$1/6$	$1/3$	$1/6$	$2/3$
	0	0	0	0	0
	1	$1/6$	0	$1/6$	$1/3$
		$1/3$	$1/3$	$1/3$	

Then $\mu_X = 0$, $\mu_Y = -1/3$

$$E(XY) = (-1)(-1)\tfrac{1}{6} + 1.(-1).\tfrac{1}{6} + (-1).1.\tfrac{1}{6} + 1.1.\tfrac{1}{6} = 0$$

$$\therefore \sigma_{XY} = E(XY) - \mu_X \mu_Y = 0 - 0 = 0$$

But $f(x,y) \neq g(x)\, h(y)$ when $x=-1, y=-1$.

Ob. If $X_1, X_2, \dots, X_n$ are

Obs. If $X_1, \dots, X_n$ independent then

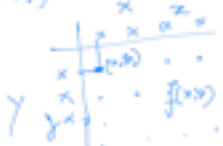
$$E(X_1 X_2 \dots X_n) = E(X_1) E(X_2) \dots E(X_n).$$

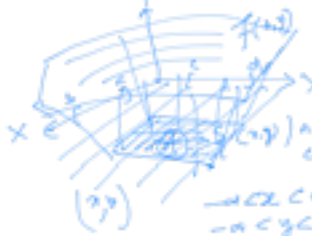
Lec-15

Pair of random variables.

(X, Y) $\quad X: S \to \mathbb{R}$
$\quad Y: S \to \mathbb{R}$

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





$\{(x,y) \mid a \le x \le b,\ c \le y \le d\}$

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

$$P((x,y) \in A) = \iint_A f(x,y)\,dx\,dy.$$

$(X, Y) \rightarrow$

marginal pdf.

Conditional pdf.

$f(X \mid Y=y) = f(x \mid y)$ → conditional pdf.

$f(Y \mid X=x) = g(y \mid x)$

Expectation

$E(X, Y)$

moments / $E(X^r Y^s)$

moments about the mean

$$E\left[(X-\mu_X)^r (Y-\mu_Y)^s\right]$$

$\boxed{r=1, s=1}$ covariance.

$$\sigma_{XY} = E\left[(X-\mu_X)(Y-\mu_Y)\right]$$

$= E(XY) - \mu_x \mu_y$

" $f(x,y)$ " are related with $g(x)$, $h(y)$

(X, Y) independent. $\boxed{f(x,y) = g(x)\, h(y)}$

$f(x_1, \ldots, x_n) = f_1(x_1) f(x_2) \cdots f(x_n)$

$\sigma_{XY} = 0$ since $E(XY) = E(X)E(Y) = \mu_x \mu_y$

Pearson product-moment Correlation coefficient

$$\boxed{\rho = \frac{\operatorname{Cov}(X,Y)}{\sigma_x \sigma_y}}$$

$\operatorname{Cov}(X,Y) = E(XY) - \mu_x \mu_y$

$\Rightarrow \boxed{E(XY) = \mu_x \mu_y + \boxed{\rho}\, \sigma_x \sigma_y}$

$Z(w) = X(w)\, Y(w)$

$$\rho = \frac{\operatorname{Cov}(X,Y)}{\sigma_x \sigma_y}$$

$$= \frac{E[(X-\mu_x)(Y-\mu_y)]}{\sigma_x \sigma_y}$$

$$= \frac{\sum_x \sum_y (x-\mu_x)(y-\mu_y) f(x,y)}{\sigma_x \sigma_y}$$

$\rho > 0$ or $\rho < 0$.

Obs. $-1 \leq \rho \leq 1$

-vely correlated. +vely correlated

□ Linear Combinations of random variables.

Let $X_1, X_2, \ldots, X_n$ be r.v.s.

Then ... $Y = \sum_{i}^{n} a_i X_i$, $a_i \in \mathbb{R}$

$$\boxed{E(Y) = \sum_{i=1}^{n} a_i E(X_i)}$$

$$Var(Y) = E\left([Y - E(Y)]^2\right)$$

$$= E\left(\left[\sum_{i=1}^{n} a_i X_i - \sum_{i=1}^{n} a_i E(X_i)\right]^2\right)$$

$$= E\left(\left(\sum_{i=1}^{n} a_i (X_i - \mu_i)\right)^2\right)$$

recall. $(a+b+c+d)^2$

$= a^2+b^2+c^2+d^2 + 2ab + 2ac + \dots + 2cd$

$$= \sum_{i=1}^{n} a_i^2 E(X_i - \mu_i)^2 + 2 \sum_{i<j}\sum a_i a_j E[(X_i-\mu_i)(X_j-\mu_j)]$$

$$\boxed{= \sum_{i=1}^{n} a_i^2 Var(X_i) + 2\sum_{i<j}\sum a_i a_j Cov(X_i, X_j)}$$

$Cov(X_i, X_j) = Cov(X_j, X_i)$

If $X_1, X_2, \dots, X_n$ are independent random variables.

$$Y = \sum_{i=1}^{n} a_i X_i$$

$$\boxed{Var(Y) = \sum_{i=1}^{n} a_i^2 Var(X_i)}$$

Exp. X, Y, Z.

$\mu_X = 2$, $\mu_Y = -3$, $\mu_Z = 4$

$\sigma_X^2 = 1$, $\sigma_Y^2 = 5$, $\sigma_Z^2 = 2$

$Cov(X, Y) = -2$,

$Cov(X, Z) = -1$

$Cov(Y, Z) = 1$

Find the mean & variance of $W = 3X - Y + 2Z$.

Suppose $X_1, \dots, X_n$.

$$Y_1 = \sum_{i=1}^{n} a_i X_i, \quad Y_2 = \sum_{i=1}^{n} b_i X_i$$

$$\mathrm{Cov}(Y_1, Y_2)$$

$$= \boxed{\sum_{i=1}^{n} a_i b_i \,\mathrm{Var}(X_i)} + \sum_{i<j}\sum (a_i b_j + a_j b_i)\,\mathrm{Cov}(X_i, X_j)$$

$$\boxed{\mathrm{Var}(Y) = \mathrm{Cov}(Y, Y)}$$

$$(X, Y)$$

Conditional random variable.

$$\boxed{Z = X \mid Y = y}$$

$f(x \mid y)$

Conditional expectation.

Let X be a discrete r.v. and $f(x \mid y)$ is the condition prob. distribution of X given $Y = y$.

Then the conditional expectation "$u(X)$ given $Y = y$"

$$E\left(u(X) \mid y\right) = \sum_x u(x) \cdot f(x \mid y)$$

If X is continuous.

$$E\left[\left(u(X) \mid y\right)\right] = \int_{-\infty}^{\infty} u(x)\, f(x \mid y)\, dx$$

If $u(X) = X$

'Conditional mean'

$$\mu_{X|y} = E(X \mid y)$$

'Conditional variance'

$$\sigma^2_{X|y} = E\left[(X - \mu_{X|y})^2 \mid y\right]$$

$$= E(X^2 \mid y) - \mu_{X|y}^2$$

$$\left(E(X^2) - (E(X))^2\right)$$

$2(x+y),\ 0<x<1$

Exp. $f(x,y) = \begin{cases} \frac{2}{3}(\ldots), & 0<\ldots \\ 0, & \text{otherwise} \end{cases}$

Find the mean & the variance of X given $Y = 1/2$.

Solⁿ. $f(x|y) = \begin{cases} \frac{2x+4y}{1+4y}, & 0<x<1 \\ 0 & \text{otherwise} \end{cases}$

Then $f(x|y=\tfrac{1}{2}) = \begin{cases} \frac{2}{3}(x+1), & 0<x<1 \\ 0, & \text{otherwise} \end{cases}$

$$\mu_{X|\frac{1}{2}} = E(X|\tfrac{1}{2}) = \int_0^1 \frac{2}{3}x(x+1)\,dx = \frac{5}{9}$$

$$E(X^2|\tfrac{1}{2}) = \int_0^1 \frac{2}{3}x^2(x+1)\,dx = \frac{7}{18}$$

Then $\sigma^2_{X|\frac{1}{2}} = \frac{7}{18} - \left(\frac{5}{9}\right)^2 = \frac{13}{162}$

$(X, Y) \longrightarrow$ $\boxed{f(x,y)}$ joint $\quad$ $\boxed{g(x),\ h(y)}$ marginals.

Bivariate Normal Distribution.

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

A pair of random variables X & Y said to have a bivariate normal distribution if the joint pdf is given by

$$f(x,y) = \frac{e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

$-\infty < x < \infty,\ -\infty < y < \infty,\ -1 < \rho < 1$

coefficient x, y

Find the marginal dist. of x.

$$g(x) = \frac{e^{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}\right)^2}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{y-\mu_2}{\sigma_2}\right)^2 - 2\rho\ldots\right]} dy$$

$u = \frac{x-\mu_1}{\sigma_1}, \quad v = \frac{y-\mu_2}{\sigma_2}$

$$g(x) = \frac{e^{-\frac{1}{2(1-\rho^2)}x^2}}{2\pi\sigma_1\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}(v^2-2\rho u v)}\,dv$$

$$v^2 - 2\rho u v = (v-\rho u)^2 - \rho^2 u^2$$

$$g(x) = \frac{e^{-\frac{1}{2}u^2}}{\sigma_1\sqrt{2\pi}} \left\{ \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{v-\rho u}{\sqrt{1-\rho^2}}\right)^2} dv \right\}$$

$$= \frac{e^{-\frac{1}{2}u^2}}{\sigma_1\sqrt{2\pi}} = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2}$$

$$= N(\mu_1, \sigma_1)$$

Obs. Suppose X, Y have bivariate normal distribution. Then the condition mean & variance for $X \mid Y = y$ are given by

$$\boxed{\begin{aligned} \mu_{X|y} &= \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2) \\ \sigma^2_{X|y} &= \sigma_1^2(1-\rho^2) \end{aligned}}$$

The conditional mean & variance $Y \mid X = x$ are

$$\boxed{\begin{aligned} \mu_{Y|x} &= \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1) \\ \sigma^2_{Y|x} &= \sigma_2^2(1-\rho^2) \end{aligned}}$$

Obs. If two r.v.s have bivariate normal distribution, they are independent if and only if $\rho = 0$

Lec-16

Functions of random variables.

$X_1, X_2, \ldots, X_n$

Suppose we know the joint probability distribution / density function

Q. $\boxed{Y} = u(X_1, \ldots, X_n)$

pdf of $Y = ?$

Two methods

1. distribution function technique
2. transformation technique

Distribution function technique.

$$F(y) = P(Y \le y)$$
$$= P(u(X_1, \ldots, X_n) \le y)$$

Then differentiating $F(y)$ we obtain

$$f(y) = \frac{dF(y)}{dy}$$

Exp. suppose pdf of X

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

Find the pdf of $Y = X^3$.

Ans.
$$G(y) = P(Y \le y)$$
$$= P(X^3 \le y)$$
$$= P(X \le y^{1/3})$$
$$= \int_0^{y^{1/3}} 6x(1-x)\,dx$$
$$= 3y^{2/3} - 2y.$$

$$\therefore g(y) = \begin{cases} 2(y^{-1/3} - 1), & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Exp. X — random variable.

$$\boxed{Y = |X|}$$

If X is standard normal. what is the pdf of Y?

Ans.
$$G(y) = P(Y \le y)$$
$$= P(|X| \le y)$$
$$= P(-y \le X \le y)$$
$$= F(y) - F(-y)$$

upon differentiation

$$g(y) = f(y) + f(-y)$$

... then

Since $|x|$ is non-negative.

$g(y) = 0$ when $y \le 0$.

Then $g(y) = \begin{cases} f(y) + f(-y), & y > 0 \\ 0 & \text{otherwise} \end{cases}$

If X is standard normal

$X \sim N(0,1)$

$g(y) = N(y; 0,1) + N(-y; 0,1)$

$= 2N(y; 0,1)$

when $y > 0$

otherwise $g(y) = 0$.

Exp. Let the joint pdf of X_1 and X_2 be

$$f(x_1, x_2) = \begin{cases} 6e^{-3x_1 - 2x_2}, & x_1 > 0, x_2 > 0 \\ 0, & \text{otherwise} \end{cases}$$

Find the pdf of $Y = X_1 + X_2$

$y = x_1 + x_2$

Ans. $F(y) = \int_0^y \int_0^{y - x_2} 6e^{-3x_1 - 2x_2}\, dx_1\, dx_2$

$= 1 + 2e^{-3y} - 3e^{-2y}$

Then upon differentiation

$f(y) = \frac{dF(y)}{dy}$

$= 6(e^{-2y} - e^{-3y})$

when $y > 0$

Otherwise $f(y) = 0$.

Transformation technique.

(a) One variable.

Given pdf of X then find pdf of $Y = u(X)$.

Exp Suppose X denotes the # of heads in four tosses of a balanced coin. Find the pdf of $Y = \frac{1}{1+X}$.

Find the pdf of "Z" $= (X-2)^2$.

Ans. $X \sim \text{Binomial}(n=4, p = \frac{1}{2})$

x	0	1	2	3	4
$f(x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

$$\boxed{Y = \frac{1}{1+X}}$$

$y = \frac{1}{1+x} \Rightarrow x = \frac{1}{y} - 1$.

x ... y_5

y	1	$\frac{1}{2}$	$\frac{1}{3}$	[illegible]	[illegible]
$g(y)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

$$f(x) = \binom{4}{x}\left(\frac{1}{2}\right)^4, \quad x = 0,1,2,3,4$$

$$g(y) = f\left(\frac{1}{y} - 1\right)$$

$$= \binom{4}{\frac{1}{y}-1}\left(\frac{1}{2}\right)^4, \quad y = 1, \tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, \tfrac{1}{5}$$

For Z, suppose $h(z)$ denotes the pdf.

$$h(0) = f(2) = \frac{6}{16}$$

$$h(1) = f(1) + f(3) = \frac{4}{16} + \frac{4}{16} = \frac{8}{16}$$

$$h(4) = f(0) + f(4) = \frac{2}{16}$$

For continuous. $Y = u(x)$

$y = u(x)$ is differentiable and either increasing or decreasing for all values of x for which $f(x) \neq 0$.

$\boxed{x = w(y)}$ exist, for all 'y'

Thm. Let $f(x)$ be the pdf of X. If $\boxed{y = u(x)}$ is differentiable and either increasing or decreasing for all values within the range of X for which $f(x) \neq 0$, then for these values of x, the eqn $y = u(x)$ can be uniquely solved for x, to give $\boxed{x = w(y)}$ & for the corresponding y the pdf of $\boxed{Y = u(X)}$ is

$$g(y) = f[w(y)]\,|w'(y)|, \quad \text{provided } u'(x) \neq 0$$

elsewhere $g(y) = 0$.

Exp. Suppose X is following exponential distribution

$$f(x) = \begin{cases} e^{-x}, & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Find pdf of $Y = \sqrt{X}$.

Soln. $y = \sqrt{x}$, $\Rightarrow x = y^2 \equiv w(y)$

$$w'(y) = \frac{dx}{dy} = 2y.$$

$$\therefore g(y) = e^{-y^2}\,|2y|$$

$$= 2y\,e^{-y^2}, \quad \text{for } y > 0$$

otherwise $g(y) = 0$

otherwise.

Exp Let X be standard normal. Find the pdf of $Z = X^2$.

Sol: $Z = u(X) = X^2$

decreasing when $x < 0$

increasing $x > 0$

First, we find pdf of $\boxed{Y = |X|}$ and then find pdf of $Z = Y^2 \,(= X^2)$

we have seen before

$$g(y) = 2\underbrace{N(y; 0, 1)}$$

$$= \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} \quad \text{when } y > 0$$

and $g(y) = 0$ otherwise

Then the task is to find pdf of $Z = Y^2$, $y > 0$ and for which $g(y) \neq 0$

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

$$\therefore\ h(z) = \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}z} \left|\frac{1}{2} z^{-1/2}\right|$$

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

for $z > 0$.

and $h(z) = 0$ otherwise

Observe that this is chi-square distribution with $r = 1$

(b) Transformation technique for ~~more~~ two random variables.

$$X_1,\ X_2$$

$$Y = u(X_1, X_2)$$

$$(Y_1, Y_2) = u(X_1, X_2)$$

If the relation between $\underbrace{y \text{ and } x_1}$ (x_2 is const.)

or $\underbrace{y \text{ and } x_2}$ (x_1 is const).

Then

$$g(y, x_2) = \underbrace{f(x_1, x_2)}\left|\frac{\partial x_1}{\partial y}\right|$$

$$g(x_1, y) = \underbrace{f(x_1, x_2)}\left|\frac{\partial x_2}{\partial y}\right|$$

Exp. The pdf of X_1 & X_2

$$f(x_1, x_2) = \begin{cases} e^{-(x_1+x_2)}, & x_1>0, x_2>0 \\ 0 & \text{otherwise} \end{cases}$$

Find the pdf of $Y = \dfrac{X_1}{X_1+X_2}$

Soln: Suppose x_1 is constant

Using the transformation technique, we can find the joint density of X_1 & Y.

$$y = \frac{x_1}{x_1+x_2} \Rightarrow x_2 = x_1 \cdot \frac{1-y}{y}$$

$$\Rightarrow \boxed{\frac{\partial x_2}{\partial y} = -\frac{x_1}{y^2}}$$

$$g(x_1, y) = e^{-x_1/y} \left| -\frac{x_1}{y^2} \right|$$

$$= \frac{x_1}{y^2} e^{-x_1/y}$$