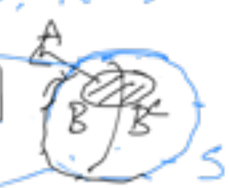
Probability and Statistics, Lec 5.

Prob Consider a construction work on campus. 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.

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



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

$$\overset{+3}{=} P(A \cap B) + P(A \cap B^c)$$

Conditional

$$= P(B)\,P(A \mid B) + P(B^c)\,P(A \mid B^c)$$

$$= (0.60 \times 0.35) + [(1 - 0.60) \times 0.85]$$

$$= **$$

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

$A = (A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_5)$

Rule of total probability:

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

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

Partition of a set: Let S be a set. Then $A_1, \dots, A_k$ 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 for Agency 3 need oill change.

What is the probability that a rented car given to IIT Kgp needs an oil change?

Ans. Let A be the even 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.

$P(B_1) = 0.6 \quad P(B_2) = 0.3, \quad P(B_3) = 0.1$

$P(A \mid B_1) = 0.09, \quad P(A \mid B_2) = 0.2, \quad P(A \mid B_3) = 0.06.$

$$P(A) = P(B_1)\,P(A \mid B_1) + P(B_2)\,P(A \mid B_2) + P(B_3)\,P(A \mid B_3)$$

$= \cancel{\ldots}$

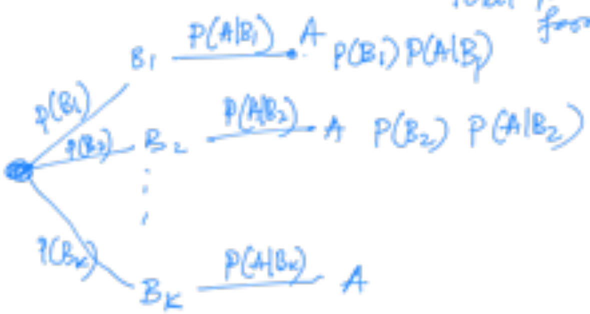
Bayes' Theorem.

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

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

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

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



Pro. For the Agency problem: what ... rented car

Prob. ... is the probability that the ... which needs an oil change came from Agency 1?

Random variables:

Objective: Associate real numbers with possible outcomes and formulate the desired events through this association.

Explain the events through the numbers which correspond to the outcomes / sample points.

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

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

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

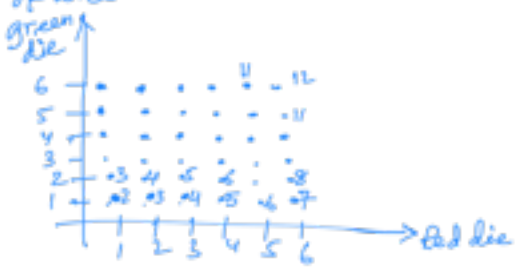
$X(H) = 2$ | $X(H) = 0$

$X(T) = 1$ | $X(T) = 1$

Exp. roll a pair of dice

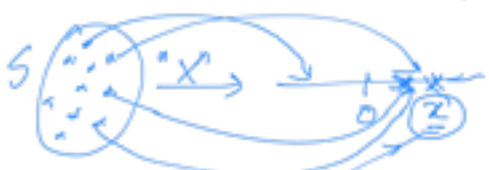
$X : S \to \mathbb{R}$

$X(s_r, s_g) = s_r + s_g$



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



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

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

$\boxed{X = 5}$ is the event given by $\{ s \in S \mid \underline{X(s) = 5} \}$

$X = a \in \mathbb{R}$ 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

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

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

$X = a$ corresponds to the event Φ if $a \neq 100$, S if $a = 100$

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

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

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

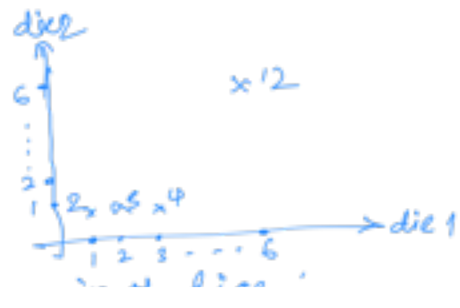
Test-2 on Feb 23 (12:10 - 1 pm)
45 mins
Topic: Random variables and special distributions

Tests will be - "Fill in the blanks type"

Random Variables

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

The sample space corresponding to the



corresponding example of rolling a pair of dice

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

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

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

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

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

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

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

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

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

"$\underbrace{a \le X \le b}$" $\equiv$ "$2 \le X \le 5$"

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

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

Sample space → discrete → discrete random variable

Sample space → continuous → continuous random variable.

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

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

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

$= f(x) \leftarrow$ p.m.f.

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

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

x	$p(X=x)$
2	1/36
3	2/36
4	3/36
⋮	⋮
12	1/36

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

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

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

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

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

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

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

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

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

	Prob.	x
$s =$ HHHH	1/16	4
HHHT	4/16	3
		⋮

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

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

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

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

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

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

$f(x) = 0$, $x \in \mathbb{R} \setminus \{0,1,2,3,4\}$

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

$S = \{ s_1, s_2, \ldots \}$, $X : S \to \mathbb{R}$, $X(s_i) = x_i$, $i \geq 1$

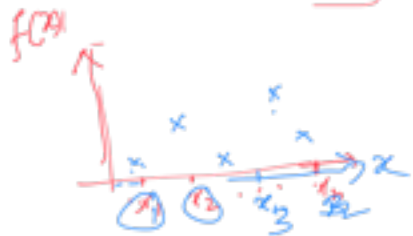
$$\text{p.m.f.} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \end{bmatrix}$$

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

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

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

H.W. why?



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

Does $f(x)$ represent a p.m.f ?

Lec-7.

Random Variables (Contd.)

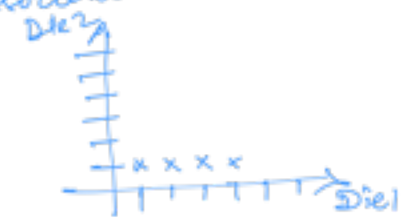
(Exam next Tuesday, Feb 02, 2021)

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

— PS Ln, $n = 1, 2, 3, 4, 5, 6, \ldots$

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

↳ purpose is to associate numbers to sample points



explaining events as subsets of IR through X

$$\underbrace{P(X = \boxed{x})}_{\text{event}} = P\Big(\underbrace{\{ s \in S \mid X(s) = x \}}\Big)$$

$f(x)$ = pmf of X

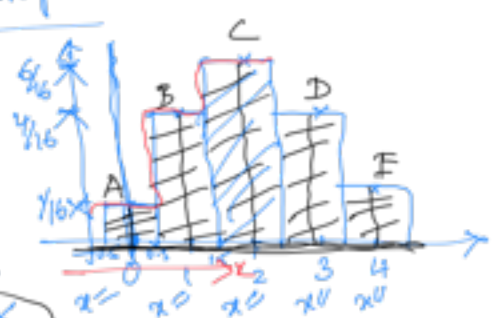
$f : \mathbb{R} \to \mathbb{R}$

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

Probability Histogram.

Exp. tossing of a unbiased coin 4 times

X = # of heads.

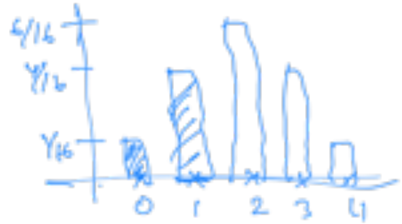


$P(X = x) = \boxed{f(x)}$

$f(0)$ = area of rectangle A

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

Bar-chart:



Cumulative distribution of a random variable 'X'

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

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

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

$= \sum f(t)$

(c.d.f) $\boxed{t \le x}$ $\downarrow$ proof.

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

Exp. $f(0) = \frac{1}{16}$, $f(1) = 4/16$
$f(2) = 6/16$, $f(3) = 4/16$
$f(4) = 1/16$

$F(0) = f(0) = 1/16$
$F(1) = f(0) + f(1) = \frac{5}{16}$
$F(4) = 1$

$F(2) = f(0) + f(1) + f(2) = 11/16$
$F(3) = 15/16$
$\vdots$

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

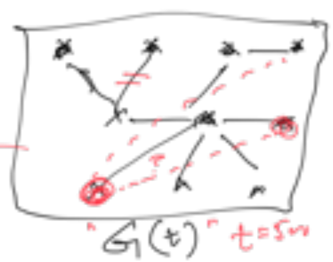
Properties of c.d.f.

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?)

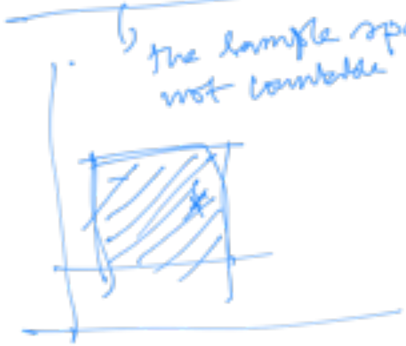
$X \to f(x)$ we ??
$X \to F(x)$ we ??

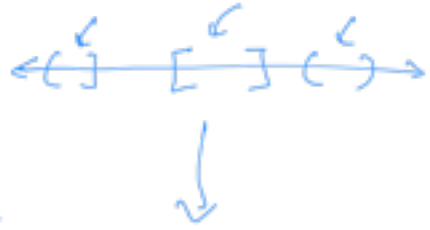


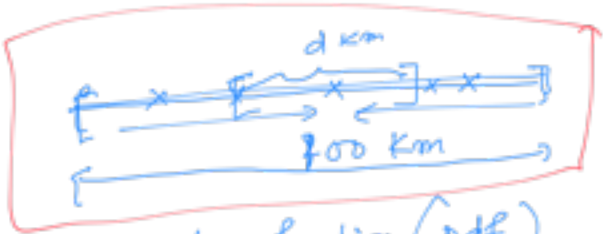
X $\leftarrow$ $f(x)$, $F(x)$

Continuous random variable

The sample space is not countable



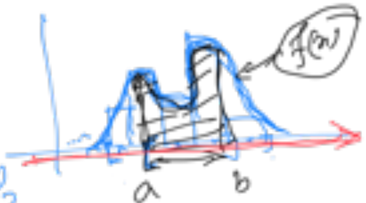




Probability density function (pdf)

Goal: Define probability $a < b$ for events

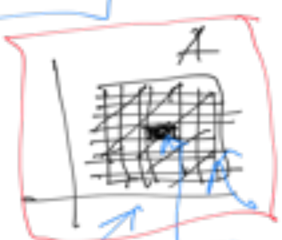
$P(a \le X \le b) = P(\{s \in S \mid a \le X(s) \le b\})$



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

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

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



probability of an interval.

Area of a point zero

it implies that

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

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

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

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

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

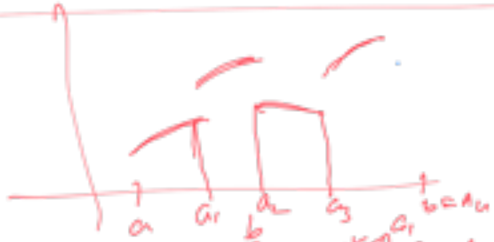
$\int_a^b \boxed{f(x)}\,dx$

Ave. No.

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

① $f(x) \geq 0$

② $\int_{-\infty}^{\infty} f(x) = 1$



$\int_a^b f(x) = \int_a^{a_1} f(x)\,dx + \int_{a_1}^{a_2} \cdot$

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

Find the values 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

Pf. $f(x) \geq 0$

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

$f(x) \to k$

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

X

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

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

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

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

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

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

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

$\therefore$ implies $k = 3$.

→ If ...

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

$= \times \times \times$

Cumulative density / distribution fn

$F(x) = P(''X \le x'')$

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

$-\infty < x < \infty$

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

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

Q. If we know $F(x)$ what can we about probability of certain 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. $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_0^x 3e^{-3t}\,dt$

$= -e^{-3t}\Big|_0^x$

$= 1 - e^{-3\cdot 1}$.

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

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

Random Experiment (Sample space, set of events)

↓

Random variable → discrete (sum) → continuous (integration)

↓

pmf/pdf

cdf

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

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

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

Reminder: Think of random variable as a measuring something

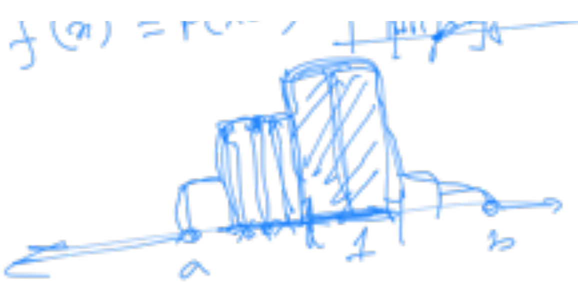
$$\int_a^b f(x)\,dx = P(a \le X \le b)$$

$a = b$ $\quad P(c \le x \le c)$

$= P(X = c)$

$= \int_c^c f(x)\,dx = 0.$

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



Lec-8

Random Variables

discrete
pmf (Histogram)
cdf

Continuous
pdf
cdf

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

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

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

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

Expected value of a random variable.

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

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

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

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

$$\text{Avg}(x_1, x_2, \ldots, x_n) = \frac{x_1 + x_2 + \cdots + x_n}{n}$$
$$= \frac{1}{n}x_1 + \frac{1}{n}x_2 + \cdots + \frac{1}{n}x_n$$
$$= P(X=x_1)x_1 + P(X=x_2)x_2 + \cdots + P(X=x_n)x_n$$

"Y" = "X" + 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

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

Ex. 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:
$$E(e^{3X/4}) = \int_0^{\infty} e^{3x/4} e^{-x} dx$$
$$= \int_0^{\infty} e^{-x/4} dx$$

$$= *$$

Formula of expectation of 'affine' functions

Obs. (1) $\boxed{E(aX+b) = aE(X) + b}$ 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) $E(aX) = aE(X)$

(3) $E(b) = b$

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

$X(s) = b$

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:

moment about the mean of X

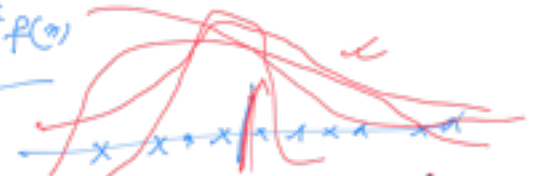
The k-th [illegible] the mean [illegible] of a r.v. X is

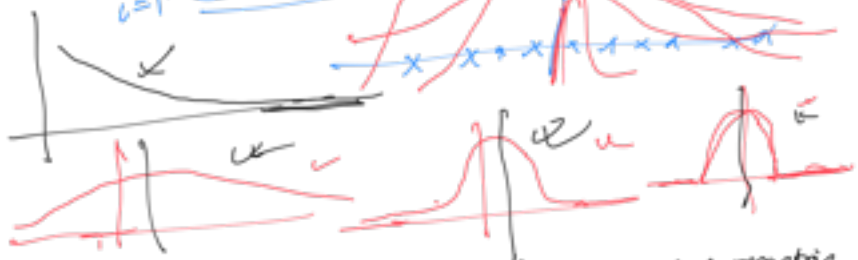
$$\mu_k = E\left((X-\mu_1')^k\right)$$

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

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

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





Measure of skewness / lack of symmetry

how much skew the pdf is μ_3 / σ^3

μ_4 / σ^4

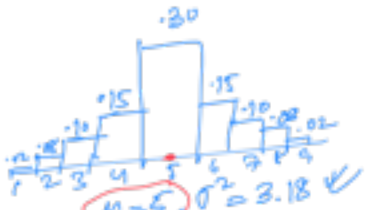
Q. How to measure the spread of a pdf

Def. $\mu_2 \rightarrow$ 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

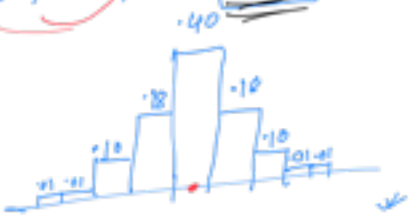
The positive square root of the variance is called standard [illegible] of the

deviation of corresponding random variable.

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



$\mu = 5$, $\sigma^2 = 3.18$



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

$\mu = 5$, $\sigma^2 = \boxed{0.88}$

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~~ ~~val~~ large value of σ^2 there is a greater probability of getting a value that is not close to the mean.

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

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

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

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

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

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

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

$E(aX) = a E(X)$

Obs. If X has variance σ^2
then $Var(aX+b) = a^2\sigma^2$

(H.W.)

If $b=0$ $\quad Var(aX) = a^2\sigma^2$

If $a=1$ $\quad \sigma^2$

$Var(aX+b) = Var(X+b)$?

$= \sigma^2$

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

---

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

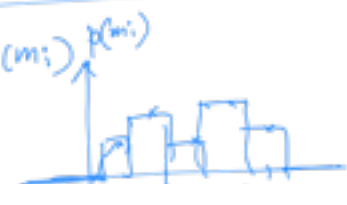
$\mu_k = E(\boxed{(X-\mu)^k})$, $k=1,2,\ldots$

Exp. Class Test - ?.

$m_1, m_2, \ldots, m_k$

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

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

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

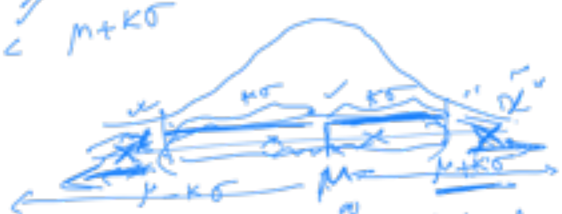
(2) $\text{Variance}(aX+b) = \boxed{a^2\sigma^2}$

Chebyshev's Theorem:

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

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

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



Pf.

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

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

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 $x \geq \mu + k\sigma$ or $x \leq \mu - k\sigma$

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

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

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

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

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

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

"Moment" generating function (mgf)

Let X be a random variable.

Then the mgf of X is defined as

$$M_X(t) = E(e^{tX})$$

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

$P(X=x)$

$$e^{tx} = 1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \cdots + \frac{t^Kx^K}{K!} + \cdots$$

If X is discrete

$$M_X(t) = \sum_x \left(1 + tx + \cdots + \frac{t^Kx^K}{K!} + \cdots\right) f(x)$$

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

$$\cdots + t\,E(X) + \frac{t^2}{2!}E(X^2) + \cdots$$

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

$$= 1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \dots + \frac{t^k}{k!}\mu_k' + \dots$$

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

Exp. Let $f(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$ be the pdf of a random variable X. Find mgf of X?

Sol: $$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} e^{-x}\,dx$$
$$= \int_0^{\infty} e^{-x(1-t)}\,dx$$
$$= \boxed{\frac{1}{1-t}} \text{ for } t < 1$$

Properties of mgf:

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

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

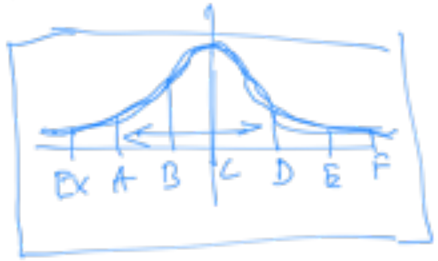
③ $M_{\frac{X+a}{b}}(t) = E\left(e^{\frac{X+a}{b}t}\right) = e^{\frac{a}{b}t} M_X\left(\frac{t}{b}\right).$

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

C_4 — 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 $P(X=x) = f(x) = \frac{1}{k}$, $x = x_1, \dots, x_k$

all distinct. $x_i \neq x_j$ if $i \neq j$

② Bernoulli distribution.

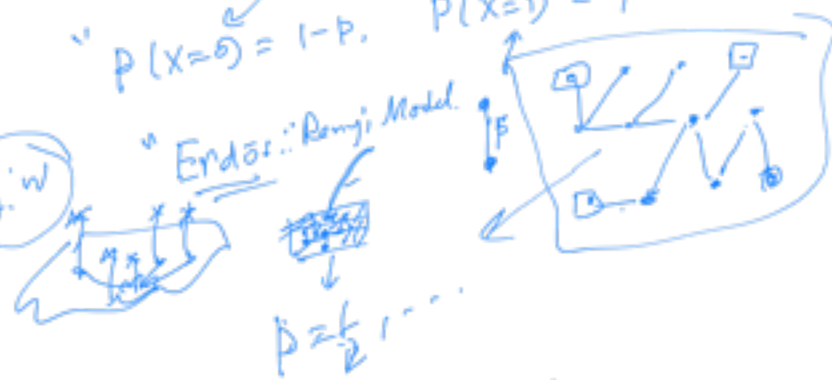
A random variable X is said to have a Bernoulli distribution if the pmf is

$$f(x) = \underbrace{p^x (1-p)^{1-x}}_{0 \le p \le 1}, \quad \boxed{x = 0, 1}$$

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

"Erdős–Renyi Model

H.W



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

$\hookrightarrow$ Bernoulli ~~trials~~

(3) The binomial distribution

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

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

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

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

" X ~~to~~ # of successes in 'n' independent trials"

$\boxed{0 \mid 0\ 1\ 0\ 0\ 0\ 1}$

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

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

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

8th Bernoulli trial and 4th success

$\boxed{0\ 1\ 0\ 1\ 0\ 1\ 0\ [1]\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1}$

'1' denotes success

failure

: ϑ denote ...

④ "Negative" Binomial distribution.

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

$$P(X=x) = f(x) = \underbrace{\binom{x-1}{k-1} p^k (1-p)^{x-k}, \quad x = k, k+1, k+2, \ldots}_{\text{the parameters } k, p.}$$

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

Explain

$X \equiv$ the # of trial on which kth success occurs.

binomial.

This is called "negative"

look at the expansion of

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

successive terms of this expansion are the probabilities.

Lec-10

Special discrete probability distribution functions

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

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

Bernoulli distribution: (Bernoulli trial)

p = probability of success

$B(p) = f(x) = p^x (1-p)^{1-x}, \; x \in \{0, 1\}.$

Binomial distribution:

(independent Bernoulli trials)

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

Negative binomial distribution:

(Pascal distribution) $\overline{B}(k, \cancel{x,} p) = f(x)$

$\binom{x-1}{k-1} p^k (1-p)^{x-k}$

$$= \binom{x-1}{k-1} p^k (1-p)^{x-k},$$

$k \to$ # of successes

$x \to$ index for the trial

$x \in \{k, k+1, k+2, \dots\}$

Geometric distribution

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

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

$\boxed{0\,0\,0\,1\,0\,0\,1 \dots 0\,1 \dots}$

Hypergeometric distribution.

Assumpt: the trials are Not independent.

100 times

$\boxed{0\,1\,0\,0\,1\,0\,0\,1 \dots\dots 0\,1 \dots 0}$

$\{0,1\}^{100}$ $\boxed{2^{100}}$

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 observation:

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

$$P(A \cap B) = P(A)P(B) = \quad B = \binom{N-M}{n-x} = \text{ways of choosing } n-x \text{ of the } N-M \text{ failures.}$$

obs. 1

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

$\boxed{n=3,\ x=2}$

$\boxed{1} \to$ success

$0 \to$ failure

# of succes = 2

of ~~choices~~ ways choosing n elements out of N elements $\binom{N}{n}$

Lec-11

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

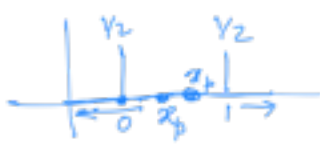
Percentile for random variable

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

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

Which means: The 100pth 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 $1-p$

$P(X \le x_p) = p$



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

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

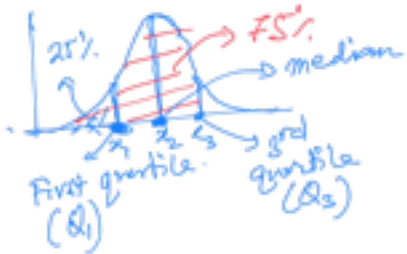
$$= \int_{-\infty}^{x_p} f(x)\,dx = \int_{-\infty}^{x_p} e^{x-2}\,dx$$

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

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

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

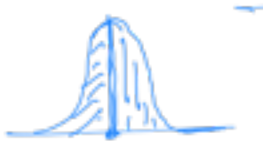
Q_1: is a number such that 25% of the observations are less than it



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

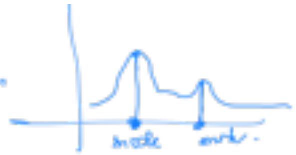
Defn. The 50th percentile of any distribution is called the 'median'.

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 23

Random variables, 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 prices 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.

... identically ...

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

$X_1, X_2, X_3, \ldots, X_n, \ldots$

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 next value depends on past values through the current value

Bernoulli Process

A sequence of Bernoulli trials.

$X_1, X_2, X_3, \ldots$

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

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

X_i's are i.i.ds

X_is are "Memoryless"

... random variables associated Bernoulli Process

Consider two ... with Bernoulli ...

Y_K — the time of K-th succes

"T_K" — Kth interarrival time which represents the number of trials follow $(K-1)$st succes untill Kth succ.

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

Previously we observed:

$T_1 \rightarrow$ the time untill the first succes is geometrically distributed

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

⋮

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

interarrival times : these are independent and geometrically distributed

Suppose K-th arrival Time.

" $Y_K = \boxed{T_1} + \boxed{T_2} + \cdots + T_K$ " ~~[illegible]~~

which is a sum of independent identically distributed random variables.

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

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

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

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' $\rightarrow$ moderate value.

going from discrete to continuous

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

Let $np = \lambda$

Poisson distribution.

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

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

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

Lec-12

Stochastic Processes

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

$\mapsto$ Bernoulli process

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

Exp. traffic accidents in a city
success — atleast one traffic accident in a minute

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

when we increase the # of observations 'n' and 'p' $\to 0$

$\boxed{\lambda = np} \longrightarrow$ 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. $n \to \infty$

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

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

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

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

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

$n = \dots$

~~for~~ ~~Bernoulli~~

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

Given, n, p and k

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

Rule of thumb:

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

if $n \geq 100$, $p \leq 0.01$, and $\lambda = np$.

Then Binomial closely approximates the Poisson.

MGF of Poisson.

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

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

	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

... exactly

Let $p(k;\tau) = P\left(\text{there are } k \text{ arrivals during an interval of length } \tau\right)$

'λ' $\longrightarrow$ arrival rate or intensity

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

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

(b) (Independence) The number of arrivals in any interval is independent of the history of arrivals outside this interval

(c) (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, March~~
Tomorrow 23rd 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.

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

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

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

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

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

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

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

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

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

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

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

$$\Gamma(\alpha) = (\alpha-1)\,\Gamma(\alpha-1)$$
$$= (\alpha-1)(\alpha-2)\,\Gamma(\alpha-2)$$
$$\vdots$$
$$= \boxed{(\alpha-1)!}$$ when α is a positive integer

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

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

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

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

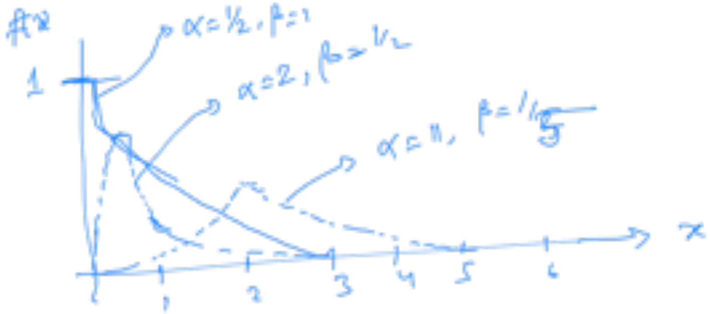
$= 1$

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

Gamma distribution: 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 $\alpha>0$, $\beta>0$

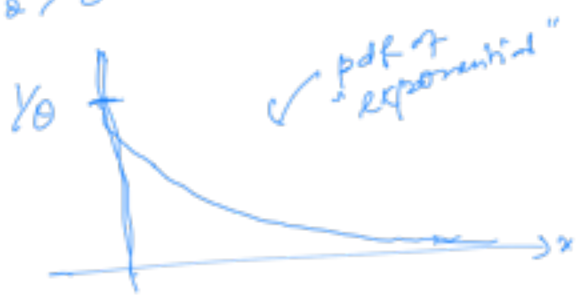


③ Setting $\alpha=1$, "$\beta=\theta$"
Then the Gamma is called the exponential distribution:

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 $\boxed{t}$ ($= p(x,t)$ ~~for~~ during the discussion of Poisson probability)

① the prob. of a

success

Lec-13

Gamma distribution.

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

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

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

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

chi-square. $\alpha = \frac{\nu}{2}, \quad \beta = 2$

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

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

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

$$= \frac{\beta^r}{\Gamma(\alpha)} \Gamma(r+\alpha) \longrightarrow (*)$$

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

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

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

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

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

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

For exponential dist.

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

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

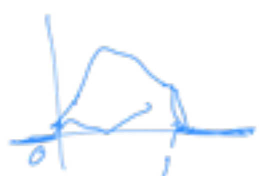
Mgf. of gamma

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

Beta distribution. pdf is

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

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



why Beta.

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

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

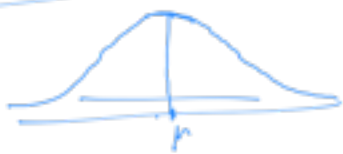
$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx = \boxed{\frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)}}$$

$$\Rightarrow \underbrace{\int_0 x \ldots}_{\substack{\text{Beta function.} \\ B(\alpha,\beta)}}$$

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

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

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.

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

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.

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

Pf.

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

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

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

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

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

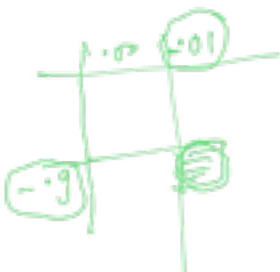
Prob. $X \sim N(\boxed{3}, \boxed{16})$ then what is $P(4 \le X \le 8)$?

Sol^n: $P(4 \le X \le 8)$

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

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

$P(Z \le -.91$



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

$$= 0.8944 - 0.5987$$

$$= \boxed{0.2957}$$

Thm. $X \sim N(\mu, \sigma^2)$ then

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

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

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

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

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

pdf of standard normal.

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

$$= \boxed{\frac{1}{\sqrt{2\pi w}} e^{-\frac{1}{2}w}}$$

Lognormal distribution.

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

— distribution of stars in the universe

— dist. of size of incomes

(Cobb - Douglas distribution)

pdf.

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

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

$\ldots x^{\beta-1} e^{-\alpha x^\beta}$, $x>0$

$$f(x) = \begin{cases} k\, x\ e & , \ldots \\ 0 & \text{otherwise} \end{cases}$$

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

$\beta = 1$ then it becomes exponential

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

Cauchy distribution.

$$\frac{\beta/\pi}{\ldots} \qquad -\infty < x < \infty$$