PROBABILITY and STATISTICS

Pendulum – Deterministic Experiment (whose outcome is specific, can be determined)

Many Real life experiments – Random (Eg. Tossing of coin)

Sample Space → Discrete
→ Continuous (Non-intuitive to understand)

10/01/17

(1) Random experiment
(2) Sample Space → discrete, continuous, mixed??

R: Tossing a coin
$S/\Omega = \{H, T\}$
$S = [0, \infty)$

(3) Events
(Subset of sample space)

R: Tossing a coin until you get Heads.
(Random exp)
$S = \{0, 1, 2, \ldots\}$
$S = \{1, 2, 3, \ldots\}$
$S = \{H, TH, TTH, TTTH, \ldots\}$
E_5 = getting 5 or more tails
$= \{TTTTTH, \ldots\}$

(4) Probability (chances)
– assign probabilities to events

Assigning Probabilities
Relative frequency.

Ex.
Tossing a coin; $S = \{H, T\}$
$P(H) = ??$ → probability of occurence of head

Suppose a coin is tossed n number of times. Record the number of heads in these n trials. (#H)

Relative frequency $\boxed{f_H = \frac{\# H}{n}}$

$n=10$,

HTHHHTTTHT $\rightarrow$ $f_H = \frac{5}{10} = \frac{1}{2} = 0.5$

$n=10$.

TTHHTTH TTT $\rightarrow$ $f_H = \frac{3}{10} = 0.3$

$f_H \rightarrow P(H)$ when $n \rightarrow \infty$

(Write a computer program to visualize it)

Definition:

Let R be a random experiment, S be a sample space, $\mathcal{F}$ is collection of all possible events in S ($\mathcal{F}$ is called as Sigma algebra)

Then the probility function is a set function from $\mathcal{F}$ to $[0,1]$, $P: \mathcal{F} \rightarrow [0,1] / \mathbb{R}$ such that

Axioms of probability

(i) for $A \in \mathcal{F}$, $0 \leq P(A) \leq 1$

(ii) $P(S) = 1$

(iii) If events $A_1, A_2 \in \mathcal{F}$ are disjoint (mutually exclusive)

$$P(A_1 \cup A_2) = P(A_1) + P(A_2)$$

(iv) If events $A_1, A_2, \ldots \in \mathcal{F}$ are denumerable events which are mutually exclusive, then

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

Min. number of statements required to define probability. (Minimal set of statements)

Consequences:

(i) $P(\phi) = 0$

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

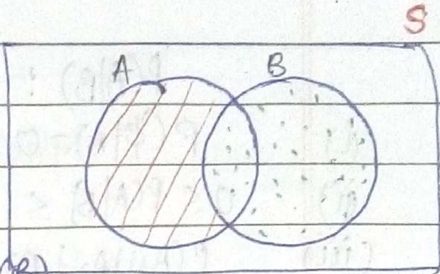
Hint: $A \cup A^c = S$, A & A^c are mutually exclusive

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

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

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

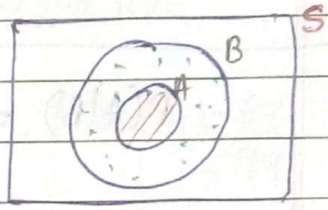
(iv) If $A \subseteq B$, then $P(A) \le P(B)$



* $A \cup B = B \cup (A \cap B^c)$ — disjoint

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

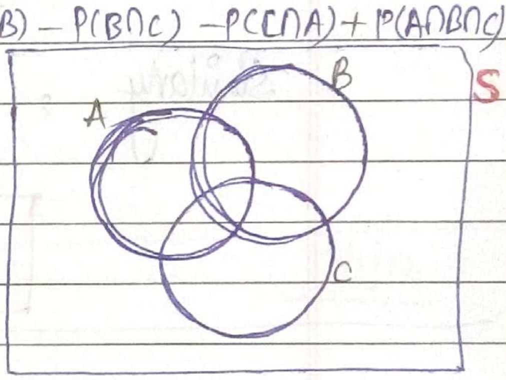
$\Rightarrow P(A \cup B) = P(B) + P(A \cap B^c) = P(B) + P(A) - P(A \cap B)$



* $P(B) = P(A) + \underbrace{P(B \cap A^c)}_{\ge 0}$

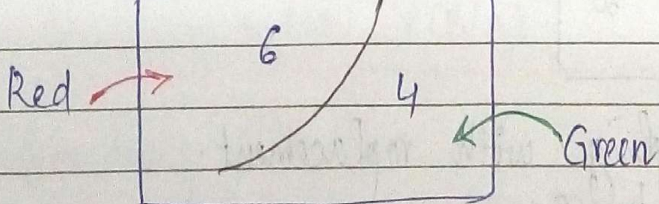
$\Rightarrow P(B) \ge P(A)$

(v) $P(\underbrace{A \cup B \cup C}_{x}) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C)$



16/01/17

CONDITIONAL PROBABILITY



$P(\text{Red Ball}) = 6/10$

$P(\text{Red Ball} \mid \text{Red ball was picked at first go})$ — added info. — altered sample space

$5/9$

$P(A|B) \rightarrow$ Probability of A given B

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}} \qquad P(B) \ne 0$$

Conditional probability is also a probability function

$P(A|B) : \mathcal{F} \to [0,1]$

(i) $P(\phi|B) = 0$

(ii) $0 \leq P(A|B) \leq 1$

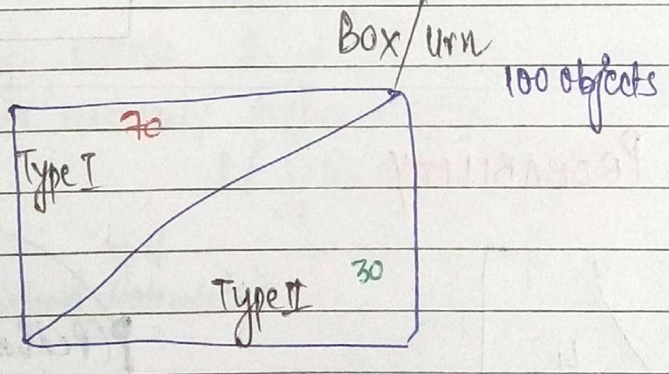
(iii) $P(A_1 \cup A_2 | B) = P(A_1|B) + P(A_2|B)$ given that A_1 and A_2 are mutually exclusive.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{given } P(B) \neq 0$$

$$\Rightarrow \boxed{P(A \cap B) = P(A|B) . P(B)}$$ Multiplication formula

Similarly, $P(B|A) = \dfrac{P(A \cap B)}{P(A)}$ given $P(A) \neq 0$

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



Experiments: (i) Sampling with replacement.
Sampling twice

A1: Type I object is selected in the first run. $\rightarrow 70/100$

A2: Type I object is selected in the second run $\rightarrow 70/100$

$$P(A_1) = P(A_2) = P(A_1 | A_2) = \frac{70}{100}$$

$\Downarrow$

A1 & A2 are independent of each other

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2) = \frac{49}{100}.$$

Expriments: (ii) Sampling without replacement
Sampling twice.

A_1 : Type 1 object is selected in the first run.
A_2 : Type I object is selected in the second run.

$P(A_1) = \frac{70}{100}$ $\quad P(A_2) = \frac{69}{99}$ or $\frac{70}{99}$, $P(A_2|A_1) = \frac{69}{99}$

$$P(A_2) = \frac{69}{99} \times \frac{70}{100} + \frac{70}{99} \times \frac{30}{100}$$

Independent events

A_1, A_2 are independent, then
$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2).$$

17/1/17

Conditional Probability.

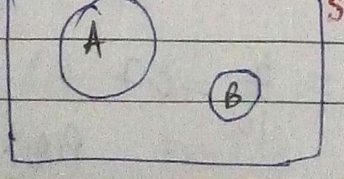
$P(A|B) = \frac{P(A \cap B)}{P(B)}$ when $P(B) \neq 0$

$$\boxed{\begin{aligned} P(A \cap B) &= P(B) \cdot P(A|B) && \text{when } P(B) \neq 0 \\ &= P(A) \cdot P(B|A) && \text{when } P(A) \neq 0 \end{aligned}}$$

Independence of 2 events:-

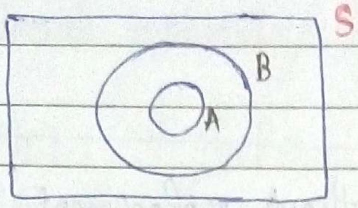
A & B are independent if and only if
$$P(A \cap B) = P(A) \cdot P(B).$$

Ex. A and B are mutually exclusive events.



$$P(A|B) = \frac{P(A \cap B)}{P(B)} = 0.$$

Ex: $A \subseteq B$.

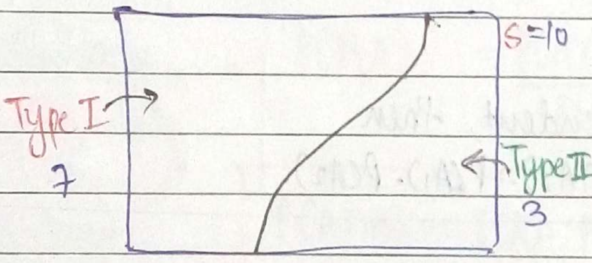


$P(B|A) = 1$

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

Example of independent events - came from the example of random exp. of sampling with replacement.

Example of events that are not independent (dependent) - sampling without replacement.



Case-I

Sampling without replacement.

A1: object picked is of type I at first run.

A2: object picked is of type I at second run.

$P(A_1) = \frac{7}{10}$ *, $P(A_2|A_1) = \frac{6}{9}$, $P(A_1 \cap A_2) = \frac{7}{10} \cdot \frac{6}{9} = \frac{42}{90} = 0.466$.

$P(A_2) = \frac{7}{10} \times \frac{6}{9} + \frac{3}{10} \times \frac{7}{9} = \frac{42+21}{90} = \frac{63}{90} = \frac{7}{10} = 0.7$.

$P(A_1) \cdot P(A_2) = 0.49. \neq P(A_1 \cap A_2)$

Case II

$S = 100$ Type I $= 70$ Type II $= 30$

$P(A_1) = \frac{70}{100}$ $P(A_2|A_1) = \frac{69}{99}$ $P(A_1 \cap A_2) = 0.7 \times \frac{69}{99} = 0.4878$.

$P(A_2) = \frac{70}{100} \times \frac{69}{99} + \frac{30}{100} \times \frac{70}{99} = \frac{7}{10} = 0.7$

$P(A_1) \cdot P(A_2) = 0.49$

PAGE:
DATE: / /

When $S=1000$, $T=700$, $\overline{T}=300$, $P(A_1 \cap A_2) \rightarrow 0.49 \equiv P(A_1) \cdot P(A_2)$

$\Rightarrow A_1$ and A_2 tends to become more independent.

Ex: A and B are independent events, then A & B^c are also independent.

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

$$= P(A)\left[1 - \cancel{P(B|A)}^{\,P(B)}\right]$$

$$= P(A)\left[1 - P(B)\right] = P(A) \cdot P(B^c)$$

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

$\Rightarrow$ A and B^c are independent.

* If A and B^c are independent $\Rightarrow$ A & B are also independent events.

Ex: A and B independent $\Rightarrow$ A^c and B^c are independent
A^c and B are independent.

Independence of Random Experiments.

Let R_1 and R_2 be two random experiments
Let A_1 and B_1 be any two events from R_1 and R_2 resp.

The experiments R_1 and R_2 are called independent if A_1 and B_1 are independent events for every choice of A_1 from R_1 and A_2 from R_2.

R_1: Tossing coin 1
R_2: Tossing coin 2.

$S_{R_1} = \{H, T\}$, $S_{R_2} = \{H, T\}$

$\mathcal{F}_{R_1} = \{\phi, \{H\}, \{T\}, \{H, T\}\}$ $\mathcal{F}_{R_2} = \{\phi, \{H\}, \{T\}, \{H, T\}\}$

Let $E^1_{R_1} = \phi$, $E^2_{R_1} = \{H\}$, $E^3_{R_1} = \{T\}$, $E^4_{R_1} = \{H, T\}$

$E^1_{R_2} = \phi$, $E^2_{R_2} = \{H\}$, $E^3_{R_2} = \{T\}$, $E^4_{R_2} = \{H, T\}$

To show: $P(E^i_{R_1} \cap E^j_{R_2}) = P(E^i_{R_1}) \cdot P(E^j_{R_2})$ for $i = 1,2,3,4$
$j = 1,2,3,4.$

$\Rightarrow R_1$ and R_2 are independent.

Total Probability

Independence:
Events $A_1, A_2, \ldots A_n$ are mutually independent if
$$P(A_{i_1} \cap A_{i_2} \cap \ldots A_{i_n}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdots P(A_{i_n})$$

$i_n = 2, 3, \ldots n.$

Partition Consider the random experiment R and the corresponding sample space S.
Then the collection of events $\{B_1, B_2, \ldots B_n\}$ is called a partition of S if:-

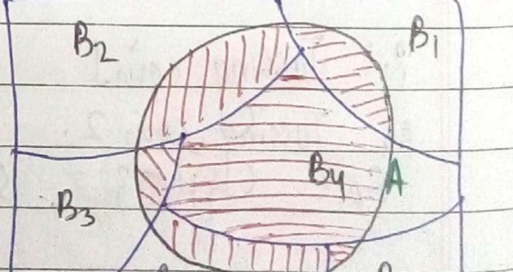
$\boxed{B_1, B_2, \ldots, B_n \text{ are mutually exclusive and non-empty}}$

and

$\boxed{B_1 \cup B_2 \cup \ldots \cup B_n = S.}$

Then for any event $A \subseteq S$

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots (A \cap B_n)$$



$\Rightarrow P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots$
$\ldots + P(A \cap B_n)$

$P(A) = P(B_1) \cdot P(A|B_1) + P(B_2) \cdot P(A|B_2) + \ldots$
$\ldots + P(B_n) P(A|B_n)$

$$\boxed{P(A) = \sum_{i=1}^{k} P(B_i) \cdot P(A|B_i)}$$ Total Probability law.

Orion

PAGE:
DATE: / /

Theorem: **Baye's Theorem**

If $\{B_1, B_2, \ldots, B_k\}$ form a partition of S and A is any arbitrary event of S, then for $i = 1, 2, \ldots, k$

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

Proof: $P(B_i|A) = \dfrac{P(B_i \cap A)}{P(A)}$

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

Example 1-10.

A: item is defective

B_i: item came from facility i.

$P(A|B_1) = 0.02$, $P(A|B_2) = 0.01$, $P(A|B_3) = 0.03$.

$P(B_1) = 0.15$, $P(B_2) = 0.80$, $P(B_3) = 0.05$

$$\sum_{i=1}^{3} P(B_i) \cdot P(A|B_i) = 0.15 \times 0.02 + 0.80 \times 0.01 + 0.05 \times 0.03$$
$$= 0.0030 + 0.0080 + 0.0015 = 0.0125.$$

$$P(B_1|A) = \frac{P(B_1) \cdot P(A|B_1)}{\sum_{i=1}^{3} P(B_i) \cdot P(A|B_i)} = \frac{0.0030}{0.0125} = \frac{6}{25}.$$

$$P(B_2|A) = \frac{P(B_2) \cdot P(A|B_2)}{\sum_{i=1}^{3} P(B_i) \cdot P(A|B_i)} = \frac{0.0080}{0.0125} = \frac{16}{25}.$$

$$P(B_3|A) = \frac{P(B_3) \cdot P(A|B_3)}{\sum_{i=1}^{3} P(B_i) \cdot P(A|B_i)} = \frac{0.0015}{0.0125} = \frac{3}{25}.$$

Ex. 1-39 A: Recruiting booklet is more than a month late.
B_i: Contract is held by i^{th} Printer.

$P(B_1) = 0.2$ $\quad$ $P(B_2) = 0.3$ $\quad$ $P(B_3) = 0.5$

$P(A|B_1) = 0.2$ $\quad$ $P(A|B_2) = 0.5$ $\quad$ $P(A|B_3) = 0.3$.

$$P(B_3|A) = \frac{P(B_3)\cdot P(A|B_3)}{\sum_{i=1}^{3} P(B_i)\cdot P(A|B_i)} = \frac{0.5\times 0.3}{0.2\times 0.2 + 0.3\times 0.5 + 0.5\times 0.3}$$

$$= \frac{0.15}{0.04+0.15+0.15}$$

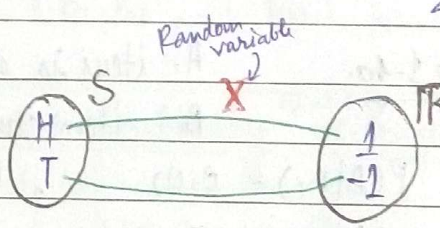
$$\boxed{P(B_3|A) = \frac{15}{34}.}$$

Random Variable.

23/1/17

R: random experiment

S: Sample space.



R: Tossing a coin

S = {Heads, Tails}

Random variable:

A function (measurable function) from S to ℝ

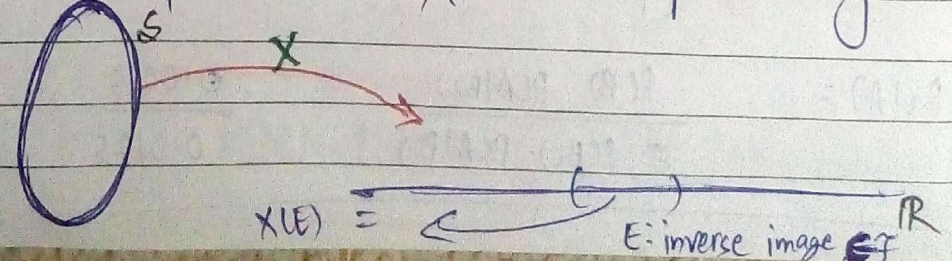
$X : S \to \mathbb{R}$

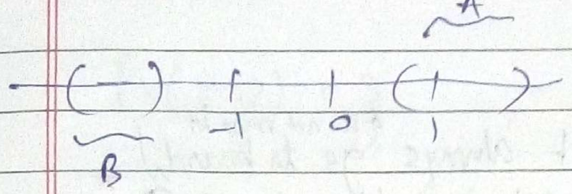
↳ co-domain (not range)

R

S

$\mathcal{F}$: Collection of all events (all subsets of S) to which probability can be assigned





$$X^{-1}(A) = \{H\} \subseteq S$$
$$\in \mathcal{F}$$
$$X^{-1}(B) = \{\} = \phi \subseteq S \in \mathcal{F}$$
$$X^{-1}(A) = \{w \in S \mid X(w) \in A\}$$

$$A = (-10, 10) \qquad X^{-1}(A) = S$$

Examples of random variables.

1) R = throwing a fair dice
$S = \{1, 2, 3, 4, 5, 6\}$.
$X : S \to \mathbb{R}$
$X(w) = w \quad \forall\, w \in S$

Trivial.

2) R = throwing 2 fair dies
$S = \{(1,1), (1,2), \ldots$
$\ldots (6,6)\}$

$X : S \to \mathbb{R}$
$X(w_1, w_2) = w_1 + w_2$

$X^{-1}(7) = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

3) R = Pick a real number.
$S = \mathbb{R}$
$X : S \to \mathbb{R}$
$X(a.b_1b_2b_3\ldots) = a.$ i.e $X(w) = \lfloor w \rfloor$
Range of $X = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$

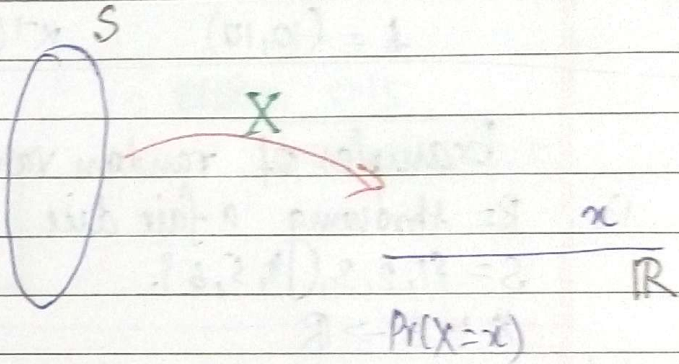
4) R = Throwing a dart
S = Board (Assuming dart always go to board & do not miss it)
X : Co-ordinates of every point on board. → (Random vector)

X: distance from the bull's eye of the point where dart is hit

---

Define

$Pr(X=x)$
$\parallel$
$Pr\{\omega \in S \mid X(\omega) = S\}$
$\parallel$
$Pr(X^{-1}(x))$



Random variables.

24/01/17

R: random exp. S: Sample space.
Then r.v. $X : S \to \mathbb{R}$ which is measurable
($X^{-1}(A)$ is an element of $\mathcal{F}$)

Ex: R: Throwing 3 fair dice.
$S = \{(1,1,1), (1,1,2), \ldots, (6,6,6)\}$ $\#S = 216 = 6^3$.
$X : S \to \mathbb{R}$
$X(\omega_1, \omega_2, \omega_3) = \dfrac{\omega_1 + \omega_2}{\omega_3}$

$R_X = \{\frac{1}{3}, \ldots, 12\}$

Ex: R: Tossing 4 fair coins.
$S = \{HHHH, HHHT, \ldots, TTTT\}$ $\#S = 2^4 = 16$
X = no. of tails
$R_X = \{0, 1, 2, 3, 4\}$

$P_X(X=2) = \frac{3}{8}$, $P_X(X=0) = \frac{1}{16}$, $P_X(X=4) = \frac{1}{16}$,

$P_X(X=3) = \frac{1}{4}$, $P_X(X=1) = \frac{1}{4}$.

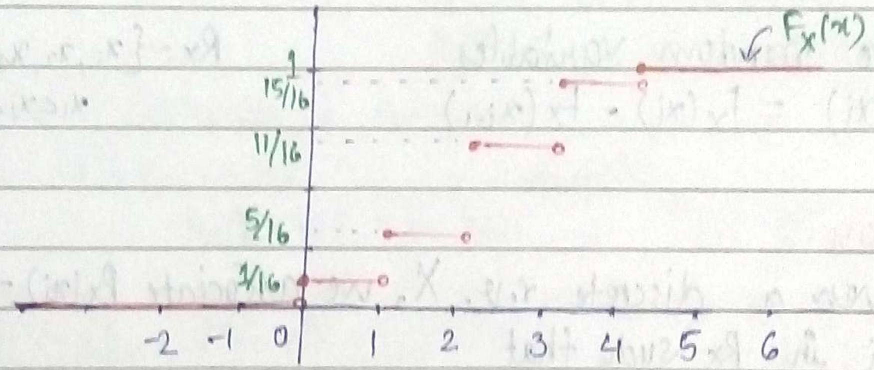
$F_X(x) = P_X(X \le x) \quad \forall x \in \mathbb{R}$ — Distribution function or Cumulative Distributive function

X → Random variable, x → Real No.

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

Discontinuous function (Jump discontinuity) (OR) at points in Range space of Random variable

* Stepwise discontinuous function



(both inclusive)

* Bounded function (bounded between 0 and 1).
* Non-decreasing
* $\lim_{x \to \infty} F_X(x) = 1$, $\lim_{x \to -\infty} F_X(x) = 0$.

* Right continuity: RHL = $F_X(x) \ne$ LHL.
  Right Hand Limit.

Distribution f^n:-

A function $F_X(x) : \mathbb{R} \to \mathbb{R}$ is called as distribution function iff

* $0 \le F_X(x) \le 1$
* $F_X(x)$ is right continuous. For any $\delta > 0$, $\lim_{\delta \to 0} F_X(x+\delta) = F_X(x)$
* $F_X(x)$ is a non-decreasing function. For $x_1 < x_2 \Rightarrow F_X(x_1) \le F_X(x_2)$
* $\lim_{x \to \infty} F_X(x) = 1$ and $\lim_{x \to -\infty} F_X(x) = 0$

Lebesgue Decomposition theorem

Every CDF (Cumulative distribution function) can be decomposed as continuous part and non-continuous part.

$$F_X(x) = \underbrace{G(x)}_{\text{Cont.}} + \underbrace{H(x)}_{\text{discont. f}^n}$$

$G(x)=0$, X discrete
$H(x)=0$, X continuous
$\left.\begin{array}{l} G(x)\neq 0 \\ H(x)\neq 0 \end{array}\right\}$, X is mixed

---

Discrete random variables.

$P_X(X=x_i) = F_X(x_i) - F_X(x_{i-1})$ $\qquad R_X = \{x_1, x_2, x_3, \ldots\}$
$\qquad x_1 < x_2 < x_3 \ldots$

Definition.

Given a discrete r.v. X, we associate $P_X(x_i) = P_X(X=x_i)$ to each x_i in R_X such that

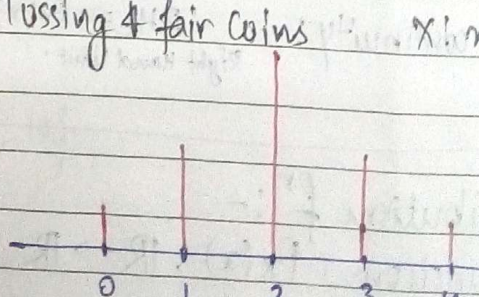
1. $P_X(x_i) \geq 0$ $\qquad \forall x_i \in R_X$
2. $\sum_i P_X(x_i) = 1$ $\qquad \sum_i P_X(x_i) = 1$

Such an assignment is called as probability mass function (pmf).

---

x_i	$P_X(x_i)$
0	1/16
1	1/4
2	3/8
3	1/4
4	1/16

R: Tossing 4 fair coins $\quad X$: no. of tails.



Sec. 1.4 - Counting principles - Combinatorics. (Read yourself).

Bernouli R.V.

X: discrete

$R_X = \{0, 1\}$ for some

$p_X(0) = 1-p$, $p_X(1) = p$ $0 \le p \le 1$.

x_i	$p(x_i)$
0	$1-p$
1	p

(1) ## Binomial R.V.

R: n trials with each trial having only two outcomes

success with probability p

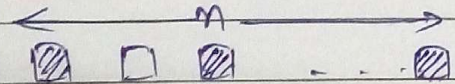
Failure with probability $1-p = q$.

X: no. of successes in n trials.

$R_X = \{0, 1, 2, \ldots n\}$. discrete.

$p_X(X = x_i) = p_X(x_i) = {}^nC_{x_i}\, p^{x_i} (1-p)^{n-x_i} \equiv \binom{n}{x_i} p^{x_i} q^{n-x_i}$

PMF → $\boxed{\begin{aligned} P(X=x) &= \binom{n}{x} p^x (1-p)^{n-x}, & x &= 0, 1, 2, \ldots n \\ &= 0, & &\text{otherwise} \end{aligned}}$



x_i places can be chosen from n places in ${}^nC_{x_i}$ ways.

$X \overset{\text{follows}}{\sim}$ Binomial (n, p)

n & p are parameters of this distribution.

n = no. of trials.

p = probability of success.

Orion

PAGE:
DATE: / /

Ex: X: Number of trials 'before' the first success in (Bernoulli trial setting)

$R_X = \{0, 1, 2, \dots\}$ → Infinite Range Space.

$p_X(x_i) =$

$p_X(0) = p$ S

$p_X(1) = (1-p)p$ FS

$p_X(2) = (1-p)^2 p$ FFS

$p_X(x_i) = (1-p)^{x_i} p$

⇒ PMF → $\boxed{\begin{aligned} p_X(x) &= \underset{\uparrow}{q}^x p, \quad x = 0, 1, 2, \dots \\ &= 0, \quad \text{otherwise} \end{aligned}}$

$$\sum_{x_i} p_X(x_i) = \sum_{x=0}^{\infty} q^x p = p \frac{1}{(1-q)} = 1$$

X ~ negative Binomial (p) (we are counting no. of trials before 1st success)

No. of trials before k Successes
↳ Negative Binomial Geometric Distribution.

Ex: X: Number of trials before first k successes.

$\underbrace{HHH..H}_{k}$

$\underbrace{HH...H}_{k-1}TH$

HW (Submit it) do for k=5.

Moments

$$\mu = E(X) = \sum_{x_i} x_i\, p_X(x_i)$$

Weights

$p_X(x_i) \to \frac{1}{n}$

usual concept of mean

$$R_X = \{x_1, x_2, \ldots x_n\}$$

and $p_X(x_i) = \frac{1}{n} \quad \forall x_i$

30/1/17

Expectation of a random variable.

R: Random expt.

S: Sample space

$X: S \to \mathbb{R}$ (measurable f^n)

$F_X : \mathbb{R} \to [0,1]$ CDF — discrete / mixed / continuous

P_X : pmf for discrete r.v. (probability mass function)

Expectation : (mean, average)

X: discrete random variable with pmf given by

$$p_X(x_i) = \text{Prob}(X = x_i)$$

$$\boxed{E(X) = \sum_{x_i} x_i\, p_X(x_i)}$$

Ex. Discrete uniform distribution.

X: discrete random variable

$R_X = \{x_1, x_2, \ldots, x_n\}$

$p_X(x_i) = \frac{1}{n}$ for $i = 1, 2, \ldots n$

$= 0$ otherwise.

$$\Rightarrow \boxed{E(X) = \frac{\sum_{i=1}^{n} x_i}{n}}$$

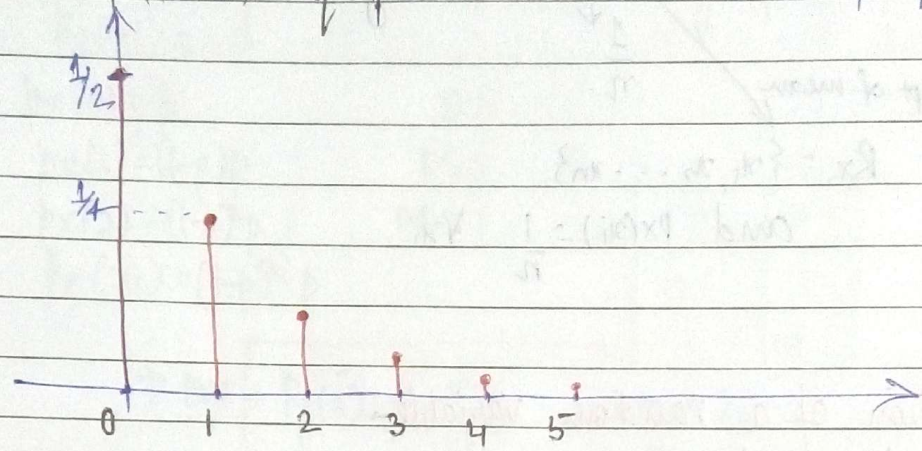
Orion

PAGE:
DATE: / /

Ex. Negative Binomial

$R_X = \{0, 1, 2, \ldots\}$.

$p(X = x_i) = q^{x_i} p$. $\qquad p = q = 1/2$.



$E(X) = \sum_{x_i} x_i \, p_X(x_i) = 0.p + 1\,pq + 2\,pq^2 + 3\,pq^3 + \ldots$

$q\,E(X) = 1\,pq^2 + 2pq^3 + \ldots$

$E(X) - qE(X) = pq + pq^2 + pq^3 + \ldots$

$(1-q)\,E(X) = pq \frac{1}{(1-q)} = q$

$E(X) = \frac{q}{p} = 1.$

Ex. Binomial distⁿ

$R_X = \{0, 1, 2, \ldots n\}$

$P_X(x_i) = \binom{n}{x_i} p^{x_i} (1-p)^{n-x_i} \qquad x_i = 0, 1, 2, \ldots n$

$= 0 \qquad$ otherwise

$E(X) = \sum_{x_i=0}^{n} x_i \, p_X(x_i) = 0.\,{}^nC_0\, p^0 q^n + 1.\,{}^nC_1\, p^1 q^{n-1} + 2.\,{}^nC_2\, p^2 q^{n-2} + \ldots$
$\ldots + (n-1)\,{}^nC_{n-1}\, p^{n-1} q^1 + n\,{}^nC_n\, p^n q^0$

$\left\{ k\,{}^nC_k = n\,{}^{n-1}C_{k-1} \right\}$

$= \sum k\,{}^nC_k\, p^k q^{n-k} = \sum n\,{}^{n-1}C_{k-1}\, p^k q^{n-k} = np \underbrace{\boxed{\sum {}^{n-1}C_{k-1}\, p^{k-1} q^{n-k}}}_{=1}$

$= np.$

Computer Ex. : Simulate Binomial distribution with $n=50$ $p=\frac{1}{2}=q$ (Toss Coin).

~~Take n=100, p:~~

Orion

PAGE:
DATE: / /

Variance

Variance measures variability in values taken by the random variable around the mean. $(E(X))$

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

Define $E(X) = \mu$

$$\boxed{Var(X) = \sum_{x_i} (x_i-\mu)^2 \, p(X=x_i)}$$

Find variance for: * Bernoulli pq

* Binomial npq

* Negative Binomial

Bernoulli :

x_i	$p_X(x_i)$
0	q
1	p

$$E(X) = 0.q + 1.p = p.$$

$$Var(X) = (0-p)^2.q + (1-p)^2.p = p^2q + q^2p = pq(p+q) = \boxed{pq = p(1-p)}.$$

Binomial : $E(X) = np. = \mu.$

$$Var(X) = (0-np)^2\, {}^nC_0 p^0 q^n + (1-np)^2\, {}^nC_1 p^1 q^{n-1} + \dots + (n-np)^2\, {}^nC_n p^n q^0$$

$$= \sum (\mu-k)^2\, {}^nC_k p^k q^{n-k} = \sum (\mu^2+k^2-2\mu k)\, {}^nC_k p^k q^{n-k}$$

$$= \mu^2 \underbrace{\sum {}^nC_k p^k q^{n-k}} + \underbrace{\sum k(k-1)\, {}^nC_k p^k q^{n-k}} + \sum (1-2\mu) \underbrace{k\, {}^nC_k p^k q^{n-k}}$$

$$\sum p^2 n(n-1) \underbrace{{}^{n-2}C_{k-2}\, p^{k-2} q^{n-k}} \qquad \sum p(n-1) \underbrace{{}^{n-1}C_{k-1}\, p^{k-1} q^{n-k}}$$

$$= \boxed{npq}$$

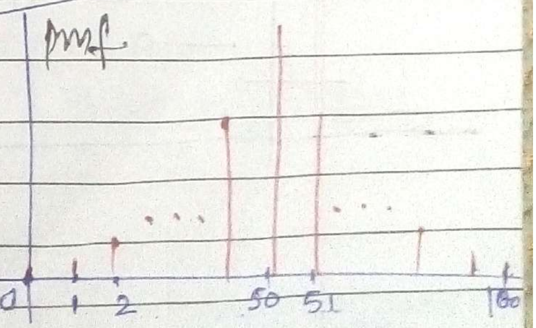
Simulate

Take $n=100$, $p=\frac{1}{2}=q$

X: no. of heads in 100 trials.

$R_X = \{0,1,2,\dots\dots 100\}$.

Plot pmf X.



R: Toss coin 100 times and count no. of heads (x_i) } Repeat this 1000 times

Find average.

ANUBHAV JAIN NOTES

Orion
PAGE:
DATE: / /

$x_1, x_2, \ldots, x_{1000} \rightarrow$ No. of heads.

Find $\dfrac{x_1 + x_2 + \ldots + x_{1000}}{1000} \overset{\text{Compare}}{\longleftrightarrow} np.$

Count no. of x_i's from $x_1, x_2, \ldots, x_{1000}$ which lie in the interval $(50 - 3\times5, 50 + 3\times5)$
$(35, 65)$.

31/01/17

Lebesgue Decomposition Theorem

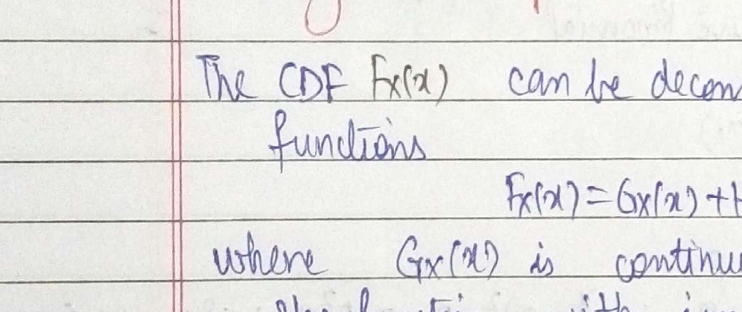
The CDF $F_X(x)$ can be decomposed as sum of two component functions

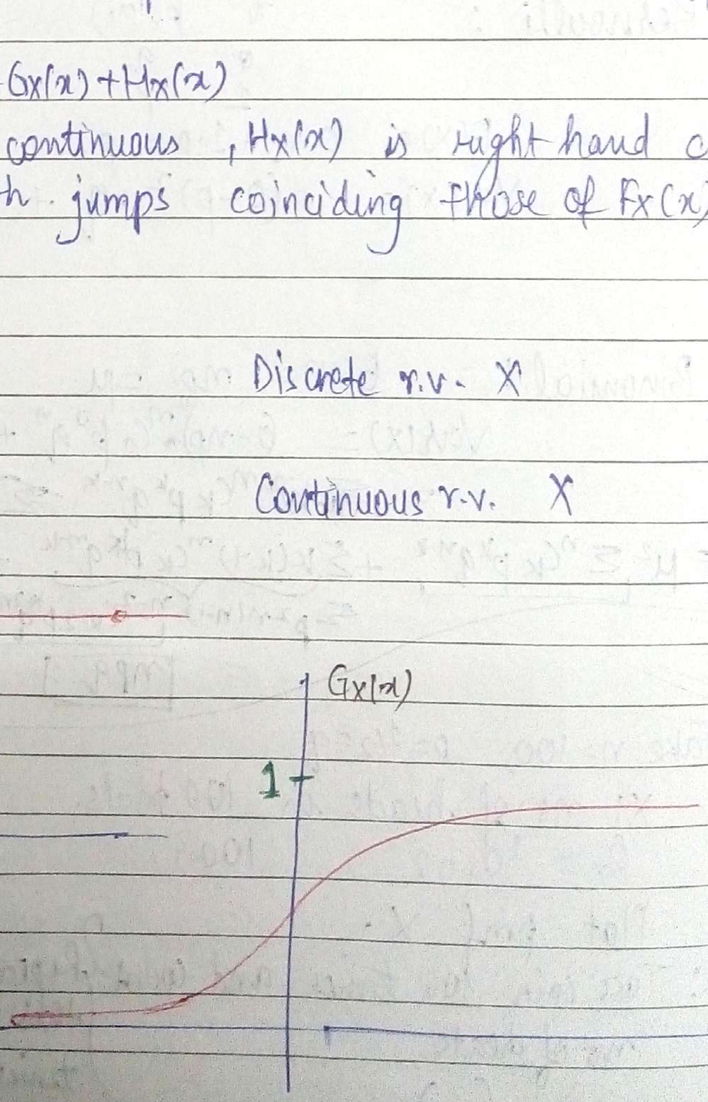
$$F_X(x) = G_X(x) + H_X(x)$$

where $G_X(x)$ is continuous, $H_X(x)$ is right hand continuous step function with jumps coinciding those of $F_X(x)$ and $H_X(-\infty) = 0$.

$G_X(x) = 0$ Discrete r.v. X

$H_X(x) = 0$ Continuous r.v. X





$F_X(x) = G_X(x) + H_X(x)$

If $H_X(x) = 0$, then x is a continuous random variable.

Then $F_X(x) = G_X(x)$ is continuous function.

(Further, $F_X(x)$ can be differentiated almost everywhere)

$$f_X(x) = \frac{d}{dx} F_X(x) \longleftarrow \text{Piecewise continuous } f^n.$$

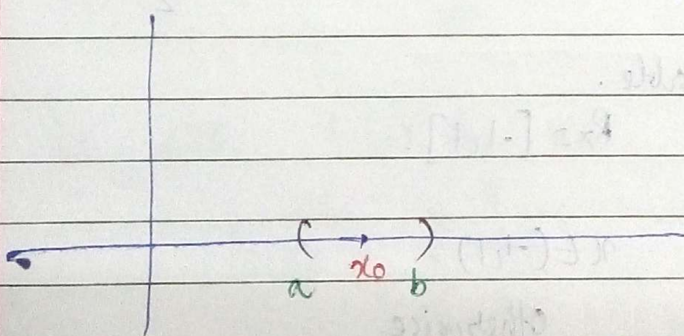
$$\boxed{P_X(x) = \lim_{\delta \to 0} \left[F_X(x+\delta) - F_X(x) \right] = 0}$$

Probability of singleton point is always zero, in case of continuous random variable.

$$f_X(x) = \frac{d}{dx} F_X(x) \qquad \text{probability density function}$$

$\downarrow$

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx = \int_{-\infty}^{x} f_X(t)\,dt$$



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

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

$P_X(x_0) = 0$

$$P(a \le X \le b) = F_X(b) - F_X(a) = \int_{-\infty}^{b} f_X(x)\,dx - \int_{-\infty}^{a} f_X(x)\,dx.$$

Probability density function

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

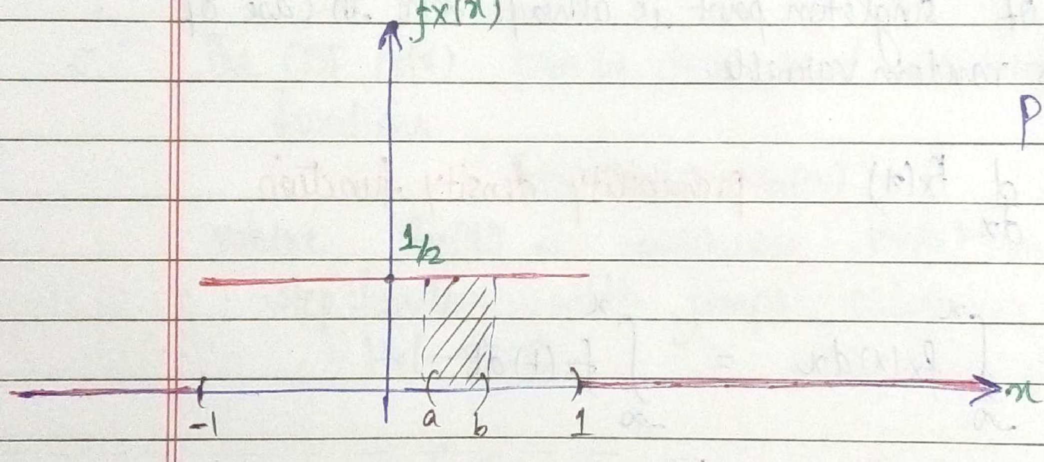
ii) $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$.

iii) $f_X(x)$ is a piecewise continuous function.

iv) $f_X(x) = 0$ for $x \notin R_X$

$$P\{e \in S \mid a \leq X(e) \leq b\} = P(a \leq X \leq b) = \int_a^b f_X(x)\,dx.$$

At point x_0, $f_X(x_0)$ does NOT represent probability at $x = x_0$.



$$P(a < X < b) = \int_a^b f_X(x)\,dx = \int_a^b \frac{1}{2}\,dx = \frac{1}{2}(b-a)$$

X - cts. random variable.

$R_X = [-1, 1]$.

$$f_X(x) = \begin{cases} 1/2 & x \in (-1, 1) \\ 0 & \text{otherwise} \end{cases}$$

For any $\boxed{\text{continuous}}$ random variable X,

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

Ex. Let X be a continuous random variable with pdf

$$f_X(x) = \frac{2x}{9} \quad 0 < x < 3$$
$$= 0 \quad \text{otherwise}$$

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^3 \frac{2x}{9}\,dx = \frac{2}{9}\frac{x^2}{2}\Big|_0^3 = \frac{x^2}{9}\Big|_0^3 = 1.$$

Write down CDF of X.

For any real no. $a \in \mathbb{R}$

$$F_X(a) = \int_{-\infty}^{a} f_X(x)\,dx$$

For $a \le 0$

$$F_X(a) = \int_{-\infty}^{a} 0\,dx = 0$$

For $0 < a < 3$

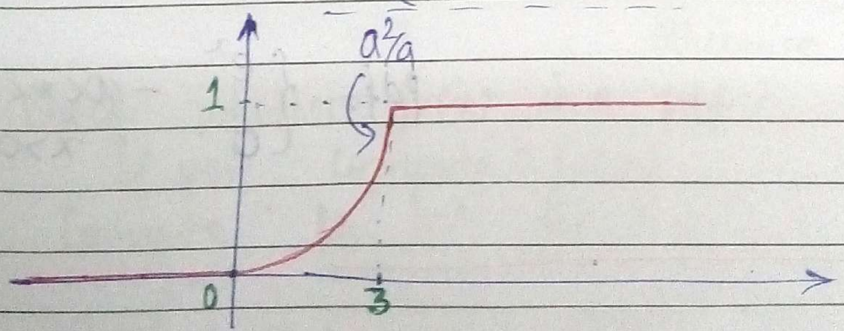
$$F_X(a) = \int_{-\infty}^{a} \frac{2x}{9}\,dx = \frac{x^2}{9}\Big|_0^a = \frac{a^2}{9}$$

For $a > 3$

$$F_X(a) = \int_{-\infty}^{a} f_X(x)\,dx = 1.$$

The CDF $F_X(a)$ for X

$$F_X(a) = \begin{cases} 0 & -\infty < a \le 0 \\ a^2/9 & 0 < a \le 3 \\ 1 & a > 3. \end{cases}$$



Moments for continuous random variables.

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

$$Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\, dx$$

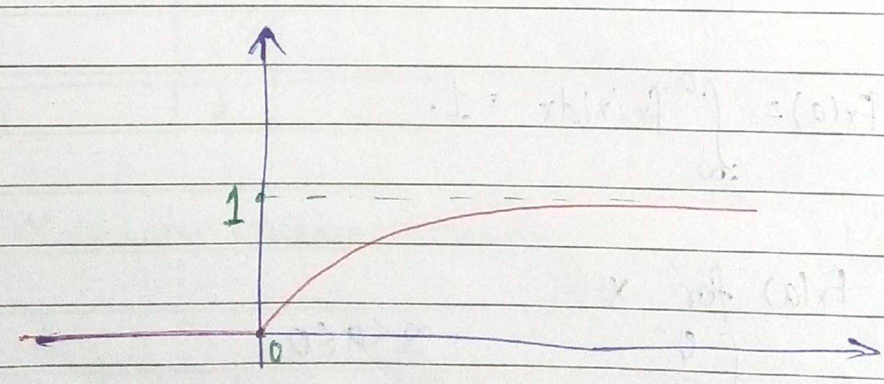
Ex// Check whether the following functions are CDFs. If yes, compute pdfs.

(a) $F_X(x) = 1-e^{-x}$ $\quad 0<x<\infty$ $\quad$ ✗ NOT CDF {f" not defined at $-\infty$ }.

(b) $F_X(x) = e^x$ $\quad -\infty < x \le 0$

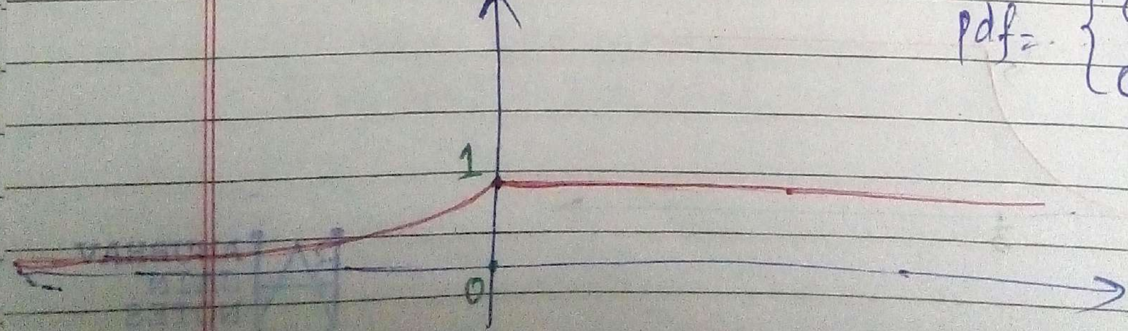
$= 1$ $\quad x>0$

$$F_X(x) = \begin{cases} 1-e^{-x} & 0<x<\infty \\ 0 & -\infty < x \le 0 \end{cases}$$ ✓ CDF



(b)

$$pdf = \begin{cases} e^x & -\infty<x<0 \\ 0 & x>0 \end{cases}$$



Ex: Consider the following function

$$P_X(x) = \frac{e^{-20}\, 20^x}{x!} \qquad x = 0, 1, 2, \ldots$$

$$= 0 \qquad \text{otherwise}$$

Is P_X a pmf?

$P_X(x) \geq 0 \quad \forall x = 0, 1, 2, \ldots$

Check $\sum_{x=0}^{\infty} P_X(x) = \sum_{x=0}^{\infty} \frac{e^{-20} \cdot 20^x}{x!}$

$$= e^{-20} \sum_{x=0}^{\infty} \frac{20^x}{x!} = e^{-20}\left[e^{20}\right] = 1.$$

Mean $= \sum_{x=0}^{\infty} x\, p_X(x) = \sum_{x=0}^{\infty} \frac{x \cdot e^{-20}\, 20^x}{x!}$

$$= e^{-20} \sum_{x=0}^{\infty} \frac{x \cdot 20^x}{x!} = e^{-20}\left(0 + \sum_{x=1}^{\infty} 20 \cdot \frac{20^{x-1}}{(x-1)!}\right)$$

$$= 20 e^{-20}\left[e^{20}\right] = \boxed{20}$$

Ex: Consider the following function

$$f_X(x) = 2e^{-2x} \qquad x \geq 0$$

$$= 0 \qquad \text{otherwise}$$

(i) Check whether $f_X(x)$ is a pdf?

(ii) If yes, compute $F_X(x)$.

(iii) Compute μ, σ^2 for X.

(i) $$\int_{-\infty}^{\infty} 2e^{-2x}\,dx = \left.\frac{2 \cdot e^{-2x}}{-2}\right]_0^{\infty} = 1$$

(ii) For $x<0$, $f_X(x)=0$

For $x>0$, $\int_0^x f_X(t)\,dt = \int_0^x 2e^{-2t}\,dt = 2\left[\frac{e^{-2t}}{-2}\right]_0^x$

$= 1-e^{-2x}$

Ex:

$$f_X(x) = \begin{cases} kx & 0\le x\le 1 \\ 2-x & 1\le x<2 \\ 0 & \text{otherwise} \end{cases}$$

If $f_X(x)$ is a pdf, find k and hence find first 2 moments of X.

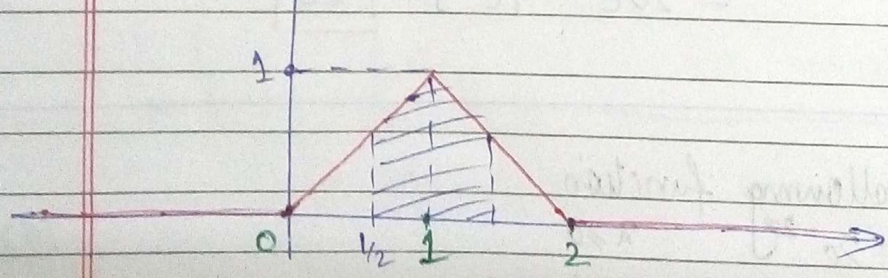
$$\int_0^2 f_X(x)\,dx = 1 \Rightarrow \int_0^1 kx\,dx + \int_1^2 (2-x)\,dx = 1.$$

$$\left[\frac{kx^2}{2}\right]_0^1 + \left[2x-\frac{x^2}{2}\right]_1^2 = 1$$

$$\left(\frac{k}{2}-0\right) + \left[2(2-1) - \frac{(4-1)}{2}\right] = 1$$

$$\frac{k}{2} + \frac{1}{2} = 1$$

$$\Rightarrow \boxed{k=1}$$



$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 x^2\,dx + \int_1^2 x(2-x)\,dx$$

$$= \left[\frac{x^3}{3}\right]_0^1 + \left[x^2\right]_1^2 - \left[\frac{x^3}{3}\right]_1^2$$

$$= \frac{1}{3} + 3 - \frac{7}{3} = 1$$

PAGE:
DATE: / /

$$Var(X) = \int_{-\infty}^{\infty}(x-\mu)^2 f_X(x)\,dx$$

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

$$= \int_0^1 x\cdot(x-1)^2\,dx + \int_1^2 (x-1)^2(2-x)\,dx$$

$$= \int_0^1 x(x^2-2x+1)\,dx + \int_1^2 (x^2-2x+1)(2-x)\,dx$$

$$= \left[\frac{x^4}{4} - \frac{2x^3}{3} + \frac{x^2}{2}\right]_0^1 + \int_1^2 (2x^2-4x+2-x^3+2x^2-x)\,dx$$

$$= \left(\frac{1}{4} - \frac{2}{3} + \frac{1}{2}\right) + \left[\frac{4x^3}{3} - \frac{x^4}{4} - \frac{5x^2}{2} + 2x\right]_1^2$$

$$= \frac{1}{12} + \frac{4(7)}{3} - \frac{15}{4} - \frac{5(3)}{2} + 2$$

$$= \frac{1+112-45-90+24}{12} = \frac{2}{12} = \boxed{\frac{1}{6}}$$

$$P\left(\tfrac{1}{2} \le X \le \tfrac{3}{2}\right) = \int_{1/2}^{3/2} f_X(x)\,dx$$

$$= 1 - \underbrace{2\times\left(\tfrac{1}{2}\times\tfrac{1}{2}\times\tfrac{1}{2}\right)}_{\text{Area of 2 unshaded } \Delta\text{'s}} = 1 - \frac{1}{4} = \frac{3}{4}$$

Ex:

$$f_X(x) = \frac{2}{9}x \qquad 0 < x < 3$$

$$= 0 \qquad \text{otherwise}$$

pdf of X.

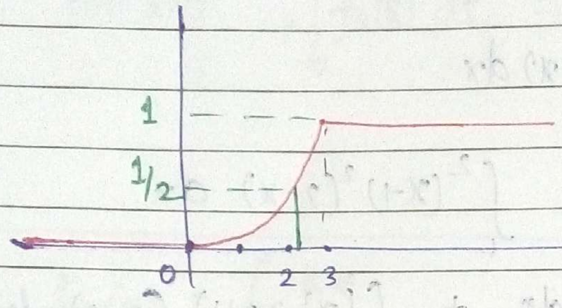
Find a number m such that

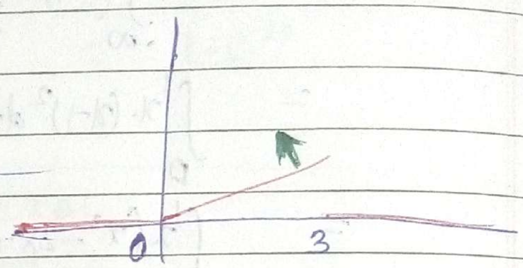
$$\underbrace{Prob(X \le m)}_{F_X(m)} = Prob(X \ge m) = 1/2$$

(Graph drawn 2 Pages back)

$$* \quad \frac{a^2}{9} = \frac{1}{2} \Rightarrow a = \sqrt{\frac{9}{2}} = \frac{3}{\sqrt{2}} = m$$

This number m is called as median of this $dist^n$ or random variable.





$$E(X) = \mu = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_0^3 x \cdot \frac{2}{9} x \, dx = \left[\frac{2}{9} \frac{x^3}{3}\right]_0^3$$

$$= \boxed{2} \neq m = \frac{3}{\sqrt{2}}$$

Mean is almost same as median.

Theorem: Chebyshev's inequality.

Let X be a random variable (continuous or discrete) and let K be any positive number, then

$$P_X\left(|X-\mu| \geq K\sigma\right) \leq \frac{1}{K^2}.$$

where

μ: Mean of X and σ: Standard deviation of X.

Moment Generating Function.

X: random variable < discrete / continuous

X - cdf
- pdf / pmf
- moment generating function

Moments (raw) of r.v. X

$$\mu_r = \sum_x x^r p_X(X=x) \quad \text{for discrete r.v.}$$

$$= \int_{-\infty}^{\infty} x^r f_X(x)\,dx \quad \text{for continuous r.v.}$$

Moment generating function (mgf)

$$M_X(t) = E(e^{tx}) \quad \text{for a given random variable X}$$

$$= \sum_x e^{tx} p_X(X=x) \quad \text{for discrete r.v.}$$

$$= \int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx \quad \text{for continuous r.v.}$$

IMPORTANCE OF MGF

$$e^{tx} = 1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \cdots$$

$$E(e^{tx}) = E\left(1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \cdots\right) \overset{*}{=} E(1) + E(tx) + E\left(\frac{t^2x^2}{2!}\right) + \cdots$$

$$\Rightarrow M_X(t) = E(1) + E(tx) + E\left(\frac{t^2x^2}{2!}\right) + \cdots$$

Linearity of Expectation

$$\boxed{M_X(t) = 1 + t\,E(x) + \frac{t^2}{2!}E(x^2) + \cdots + \frac{t^r}{r!}E(x^r) + \cdots}$$

expectation is linear
t → Not random ⇒ can be pulled out.

* : Conditions Apply.

PAGE:
DATE: / /

* $$\left.\frac{d}{dt} M_X(t)\right|_{t=0} = E(X)$$

$\rightarrow \frac{d}{dt} M_X(t) = 0 + E(X) + \dots + \frac{t^{r-1}}{(r-1)!} E(x^r) + \dots$

* $$\left.\frac{d^2}{dt^2} M_X(t)\right|_{t=0} = E(X^2)$$

$\rightarrow \frac{d^2}{dt^2} M_X(t) = E(x^2) + \dots + \frac{t^{r-2}}{(r-2)!} E(x^r) + \dots$

$$\left.\frac{d^r}{dt^r} M_X(t)\right|_{t=0} = E(X^r)$$

Ex.

$X \sim Bin(n, p)$

$R_X = \{0, 1, 2, \dots n\}$

p: Prob. of success of Bernoulli

$q = 1-p$.

$$p_X(X=x) = \binom{n}{x} p^x q^{n-x} \qquad x = 0, 1, 2, \dots n$$
$$= 0 \qquad \text{otherwise}$$

pmf of X.

$$M_X(t) = E(e^{tx}) = \sum_{x=0}^{n} e^{tx} p_X(X=x)$$

$$= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x q^{n-x}$$

$$\boxed{M_X(t) = (pe^t + q)^n}$$ $\leftarrow$ M g f for Binomial r.v.

$$\mu = E(X) = \left.\frac{d}{dt} M_X(t)\right|_{t=0} = \left. n(pe^t + q)^{n-1} (pe^t)\right|_{t=0}$$

$$= n(1)^{n-1} pe^0 = \boxed{np.}$$

$$E(x^2) = \left.\frac{d^2}{dt^2} M_X(t)\right|_{t=0} = \left.\frac{d}{dt}\left(npe^t (pe^t + q)^{n-1}\right)\right|_{t=0}$$

$$= np\left\{e^t (pe^t+q)^{n-1} + e^t (n-1) pe^t (pe^t+q)^{n-2}\right\}\Big|_{t=0} = np\{1 + (n-1)p\}$$

$$= \boxed{n^2p^2 + npq}$$

PAGE:
DATE: / /

$$\begin{aligned} Var(X) &= E(X - E(X))^2 \\ &= E(X-\mu)^2 \quad \text{where } \mu = E(X) \\ &= E(X^2 - 2X\mu + \mu^2) \\ &= E(X^2) - 2\mu E(X) + \mu^2 \quad \text{.. laws of linearity of Expectation} \\ &= E(X^2) - 2\mu^2 + \mu^2 \\ &= E(X^2) - \mu^2 \end{aligned}$$

$$\boxed{Var(X) = E(X^2) - [E(X)]^2} \quad = (n^2p^2 + npq) - (np)^2 = npq$$

Given a random variable, we know

- CDF of X — continuous / discrete
- pdf/ pmf of X (parameters of the distribution)
- mean and variance, median
- mgf

Geometric distribution

X: no. of tails before first heads in a coin toss exp.

$R_X = \{0, 1, 2, \dots\}$ prob (Heads) = p

$$p_X(X=x) = q^x p \quad x \in R_X$$
$$= 0 \quad \text{otherwise}$$

Sample space associated with the random experiment

$S = \{H, TH, TTH, \dots\}$

Compute $P(X=10) = q^{10}p$.

$P(X=5) = q^5 p$.

$P(X=10 \mid X=5) = 0$

ANUBHAV JAIN NOTES

Prepare table for all properties studied.

$$P(X \geq 10 \mid X \geq 5) = \frac{P(X \geq 10)}{P(X \geq 5)} = \frac{1 - P(X < 10)}{1 - P(X < 5)}$$

CDF

$$= \frac{1 - F_X(9)}{1 - F_X(4)}$$

$F_X(x)$ for geometric distribution:-

$$F_X(x) = \sum_{n=0}^{x} q^n p = \frac{\cancel{p}(1-q^{x+1})}{\cancel{(1-q)}} \quad , \text{ for } x = 0,1,2 \ldots$$

$$\boxed{F_X(x) = 1 - q^{x+1}}$$

$$\Rightarrow P(X \geq 10 \mid X \geq 5) = \frac{1 - (1 - q^{10})}{1 - (1 - q^5)} = \boxed{q^5} = P(X \geq 5)$$

For s and t positive integers,

$$P(X \geq s+t \mid X \geq t) = P(X \geq s)$$

for geometric distribution

Memoryless property.

H, TH, TTH, ~~TT~~TH, ~~TT~~TTH, ...

(H, TH, TTH, ...)

Special Discrete Distributions

1. Bernoulli

$X \sim \text{Bernoulli}(p)$ $\quad\quad 0<p<1$

$R_X = \{0,1\}$

Sample space associated with the random experiment

$S = \{F, S\}$ $\quad\quad M_X(t) = E(e^{tX}) = q + pe^t$

Name	Parameters	pmf	Mean	Variance	MGF
Bernoulli	$0<p<1$	$R_X = \{0,1\}$ $P_X(0) = q = 1-p$ $P_X(1) = p$ $= 0$, otherwise	p	pq	$q + pe^t$
Binomial	$n = 1, 2, \ldots$ $0<p<1$	$R_X = \{0,1,2,\ldots n\}$ $P_X(x) = \binom{n}{x} p^x q^{n-x}$ $x = 0, 1, \ldots n$ $= 0$, otherwise	np	npq	$(q+pe^t)^n$
Geometric Distribution	$0<p<1$	$R_X = \{1, 2, \ldots\}$ $P_X(X=x) = q^{x-1} p$, $x \in R_X$ $= 0$, otherwise	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{pe^t}{1-qe^t}$
Negative Binomial / Pascal	$r = 1, 2, \ldots$ $0<p<1$	$R_X = \{r, r+1, \ldots\}$ $P_X(x) = \binom{x-1}{r-1} p^r q^{x-r}$, $x \in R_X$ $= 0$, otherwise	$\frac{r}{p}$	$\frac{rq}{p^2}$	$\left(\frac{pe^t}{1-qe^t}\right)^r$
Hypergeometric	N, D, n	$R_X = \{0, 1, 2, \ldots \min(n, D)\}$ $P_X(X=x) = \frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}}$, $x \in R_X$ $= 0$ otherwise	$n \cdot \left[\frac{D}{N}\right]$	$n \cdot \left[\frac{D}{N}\right]\left[1-\frac{D}{N}\right]\left[\frac{N-n}{N-1}\right]$	

Orion

PAGE:
DATE: / /

2. Binomial

$X \sim \text{Binomial}(n,p)$

Random expt: n independent Bernoulli trials.

$$S = \{\underbrace{SSS\ldots SS}_{n}, \underbrace{SSS\ldots SF}_{n-1}, \ldots, \underbrace{FFF\ldots F}_{n}\}$$

X_i: Bernoulli r.v. associated with trial i.

$X = \sum_{i=1}^{n} X_i$ = counting no. of S's in n trials.

(Binomial $\underbrace{\text{distribution}}_{\text{r.v.}}$ as sum of n independent Bernoulli $\underbrace{\text{distribution}}_{\text{r.v.}}$)

$R_X = \{0, 1, 2, \ldots n\}$

$$P_X(X=x) = \binom{n}{x} p^x q^{n-x} \qquad x = 0, 1, 2, \ldots n$$

$$= 0 \qquad \text{Otherwise.}$$

$$E(X) = \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

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

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

$E(X) = E(X_1 + X_2 + \ldots + X_n)$

$= E(X_1) + E(X_2) + \ldots + E(X_n)$

$= p + p + \ldots + p = np.$

$E(X) = np$

$Var(X) = Var(X_1 + X_2 + \ldots + X_n)$

$= Var(X_1) + Var(X_2) + \ldots + Var(X_n)$ ⟵ Only when $X_1, X_2, \ldots X_n$ are independent

$= pq + pq + \ldots + pq$

$Var(X) = npq$

$$MGF(X) = M_X(t) = M_{\Sigma x_i}(t) = E\left(e^{(\Sigma x_i)t}\right)$$

$$= E\left(e^{x_1 t} \cdot e^{x_2 t} \cdot e^{x_3 t} \cdots e^{x_n t}\right)$$

$$= E(e^{x_1 t})\, E(e^{x_2 t}) \cdots E(e^{x_n t})$$

$$MGF(X) = (q + pe^t)^n$$

Application of Binomial distribution.

$$X \sim \text{Binomial}(n, p)$$

Define a new random variable

$$\hat{p} = \frac{X}{n}$$

$$E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{1}{n} E(X) = \frac{1}{n} np.$$

$$E(\hat{p}) = p$$

$$Var(\hat{p}) = Var\left(\frac{X}{n}\right) = \frac{1}{n^2} Var(X) = \frac{1}{n^2} npq$$

$$Var(\hat{p}) = pq/n$$

In gen.

$$Var(ax) = \int_{-\infty}^{\infty} (ax - \underbrace{a\mu}_{E(ax)})^2 f_x(x)\, dx = a^2 \int_{-\infty}^{\infty} (x-\mu)^2 f_x(x)\, dx$$

$$= a^2\, Var(x)$$

$$Prob(\hat{p} \le p_0) \quad \text{for } 0 < p_0 < 1$$

$$= Prob\left(\frac{X}{n} \le p_0\right) = Prob(X \le np_0) = F_X(np_0)$$

$$\boxed{Prob(\hat{p} \le p_0) = \sum_{x=0}^{\lfloor np_0 \rfloor} \binom{n}{x} p^x q^{n-x}}$$

3. Geometric distn

$0<p<1$. p = Prob. of Success

$S = \{S, FS, FFS, \ldots\}$.

X: no. of trials to obtain 1st success.

$R_X = \{1, 2, 3, \ldots\}$.

$$P_X(X=x) = q^{x-1}p \quad x=1,2,\ldots$$
$$= 0, \quad \text{otherwise}$$

$E(X) = \mu = \frac{1}{p}$ $\quad Var(X) = \frac{q}{p^2}$

$M_X(t) = \frac{pe^t}{1-qe^t}$

$E(X) = \sum_{x_i} x_i p_X(x_i) = 1 \cdot q^0 p + 2 \cdot qp + 3 \cdot q^2 p + \ldots$

$qE(X) = \qquad 1 \cdot qp + 2q^2p + \ldots$

$(1-q)E(X) = p + qp + q^2p + \ldots = \frac{p}{1-q} = 1$

$\Rightarrow E(X) = \frac{1}{1-q} = \frac{1}{p}$.

4. Negative Binomial / Pascal distribution

Expt: Perform indep. Bernoulli trials until one gets r successes

$S = \{\underbrace{SS\ldots S}_{r}, \underbrace{FSS\ldots S}_{r}, SFS\ldots S, SSFS\ldots S, \ldots SS\,SFS, \ldots\}$

X: no. of trials to get first r successes.

$R_X = \{r, r+1, r+2, \ldots\}$.

$$\boxed{P_X(X=x) = \binom{x-1}{r-1} p^r q^{x-r} \quad x=r, r+1, \ldots \qquad = 0 \quad \text{otherwise}}$$

Require x trials to obtain r success $\equiv (X=x)$

$$\underbrace{\_\ \_\ \_\ \_\ \cdots\ \_\ \underline{S}}$$

there are x places.

$\Rightarrow x-1$ places which are to be filled with $r-1$ successes.
This can be achieved in $\binom{x-1}{r-1}$ ways.

---

5. Hypergeometric distribution

Finite population of N items
Defective items $= D$ $(D \le N)$
Draw random sample of size n from this population (without replacement) X: no. of defective items in the sample of size n

$R_X = 0, 1, 2, \ldots, \min(n, D)$

$$P_X(X=x) = \frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}} \qquad x = 0, 1, 2, \ldots \min(n, D)$$

$$= 0 \qquad \text{otherwise.}$$

$$E(X) = n\left[\frac{D}{N}\right]$$

$$Var(X) = n\left[\frac{D}{n}\right]\left[1-\frac{D}{n}\right]\left[\frac{N-n}{N-1}\right]$$

Discrete uniform

X = discrete uniform

$R_X = \{x_1, x_2, \ldots, x_n\}$

$$P_X(X=x_i) = \frac{1}{n}, \quad \text{for } x = x_1, x_2, \ldots x_n$$

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

Orion
PAGE:
DATE: / /

Q 8.8

Suppose a random sample of size 200 is taken from a process that is 0.07 fraction defective. What is the probability that $\hat{p}$ will exceed the true fraction defective by one std. deviation? By two std. deviations? By three standard deviations?

$$\hat{p} = \frac{X}{n}$$

$$P(\hat{p} - 0.07 > \sigma) = P\left(\frac{X}{n} - 0.07 > \sigma\right)$$

$$= P(X > 0.07n + n\sigma) = P\left(X > 0.07n + n\sqrt{\frac{pq}{n}}\right)$$

$$= P(X > \sqrt{npq} + 0.07n) = 1 - F_X(\sqrt{npq} + 0.07n)$$

$$= \boxed{1 - F_X(25.41)}$$

Q 5.14

A submarine's probability of sinking an enemy ship with any one firing of its torpedoes is 0.8. If the firings are independent, determine the probability of a sinking within the first two firings. Within the first three.

$0.8 + (0.2)(0.8) = \boxed{0.96}$

Q 5.16

A potential customer enters an automobile dealership every hr. The prob. of a salesperson concluding a transaction is 0.10. She is determined to keep working until she has sold three cars. What is the prob. that she will have to work exactly 8 hours? More than 8 hours?

X: no. of hours taken to sell 3^{rd} cars. $p = 0.1$ $q = 0.9$

Exactly 8 hours $P(X=8) = {}^{7}C_2\, p^3 q^5 = 21\,(0.1)^3 (0.9)^5$

$= \boxed{0.124}$

More than 8 hours $P(X>8) = \sum_{x=9}^{\infty} \binom{x-1}{r-1} p^r q^{x-r}$

$$= 1 - P(X \le 8) = 1 - \sum_{x=3}^{8} \binom{x-1}{r-1} p^r q^{x-r}$$

$$= 1 - F(8)$$

$$= 1 - \left\{ {}^{2}C_2\, p^3 + {}^{3}C_2\, p^3 q + {}^{4}C_2\, p^3 q^2 + {}^{5}C_2\, p^3 q^3 + {}^{6}C_2\, p^3 q^4 + {}^{7}C_2\, p^3 q^5 \right\}$$

$$= 1 - \left\{ p^3 + 3p^3 q + 6p^3 q^2 + 10p^3 q^3 + 15p^3 q^4 + 21p^3 q^5 \right\}$$

$$= 1 - p^3 \left\{ 1 + 3q + 6q^2 + 10q^3 + 15q^4 + 21q^5 \right\} = 1 - 0.038 = \boxed{0.9619}$$