cumulative distribution Function

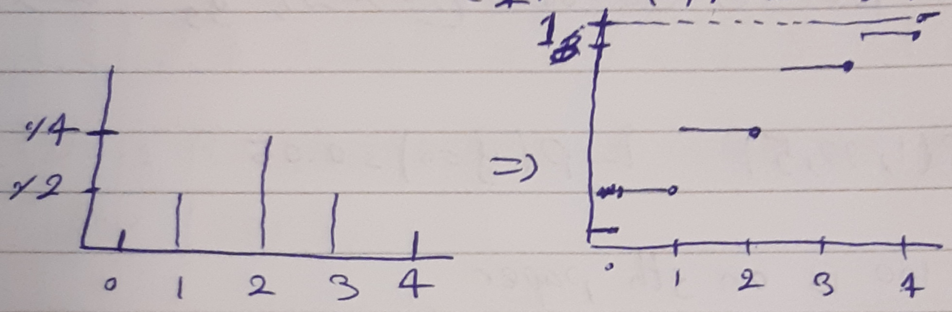
The cdf on function RV is a function

$$F_X(x) = P(\underbrace{X \le x}_{\text{event}})$$

Eg: $X \sim Bin(4, \frac{1}{2})$

$$F(1.5) = P(X \le 1.5) = P(X=0) + P(X=1)$$

$$= \left(\frac{1}{2}\right)^4 + \binom{4}{1}\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^3 = \frac{5}{6}$$

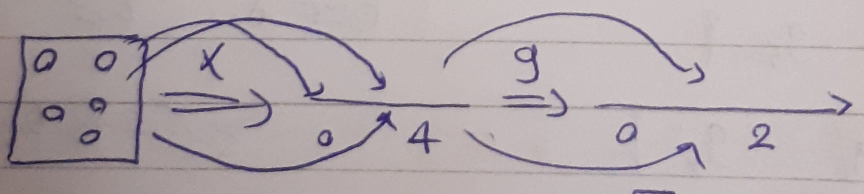


Function of Random variables

Def an experiment with sample space S

$g(x)$ is an RV that maps s to $g(x(s))$

Ex: e^x, x^2, $\sin(x)$ are RV



$g = \sqrt{x}$

if g is one-to-one ($y = g(x)$)

$$P(Y = g(x)) = P(g(X) = g(x)) = P(X = x)$$

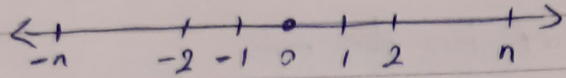
Randomwalk

A partike start at o

The particle move ~~one~~ step one unit at each steps

independent steps

the probability of moving left or right is the same.



Y: particle's position after n steps. PMF(Y) = ?

X: # rights $X \sim Bin(n, \frac{1}{2})$ take $0, 1, \dots n$

$Y = X \times 1 + (n - X) \times (-1) = 2X - n \to$ takes $-n, 2-n, \dots n$

$p(y = k) =$

$p(2x - n = k) = p\left(x = \frac{n+k}{2}\right) = \binom{n}{\frac{n+k}{2}} \left(\frac{1}{2}\right)^n$

position ی particle را مشخص می‌کند

probability of left and right are same

if g is not one-to-one:

$$x_1 \searrow y \nearrow x_2 \Rightarrow y = x^2 \quad x = 2, -2 \Rightarrow y = 4$$

pmf of $g(X) \Rightarrow p(g(X) = y) = \sum_{x \cdot g(x) = y} p(X = x)$

$p(y = 4) = p(X = 2) + p(X = -2)$

Function of two random variables

$X(s)$ $\quad g(x, y)$ is an RV

$Y(s)$ $\quad s \longrightarrow g(X(s), Y(s))$

Ex : mean of two die rolls

s	X	Y	max(X, Y)
1,2	1	2	2
⋮	1	6	6
1,6			
⋮			
6,6			

$P(\max(X, Y) = 1) = \frac{1}{36}$

$P(\max(X, Y) = 2) = \frac{3}{36}$

$$P(\max(X, Y) = 5) = \underbrace{P(X = 5, Y \le 4)}_{\frac{4}{36}} + \underbrace{P(X \le 4, Y = 5)}_{\frac{4}{36}}$$

$$+ \underbrace{P(X = 5, Y = 5)}_{\frac{1}{36}}$$

Independence of Random variable.

$X \longleftrightarrow Y$

Def : X and Y are said to be independent if

$P(X \le x, Y \le y) = P(X \le x)\, P(Y \le y)$

for $x, y \in R$

$P(X_1 \le x_1, \dots X_n \le x_n) = P(X_1 \le x_1) P(X_2 \le x_2) \dots$

$P(X_n \le x_n)$

why we don't consider all of the combination variable

if $x_2 \to \infty$ the $p(X_2 \leq x_2) = 1$ so we could eliminate x_2 from left part.

A roll of two fair dice : $X + Y \perp X - Y$?

آیا جمع و x-y مستقل است یا خیر

$p(X + Y = 12, X - Y = 1) \neq p(x+y=12)\, p(X - Y = 1)$

if $x = 6$, $y = 6$ $\quad 6-6 \neq 0 \quad \frac{1}{36} \times \frac{5}{36}$

Theorem : X and Y are independent : every function of x and every function of y are independent

Ex

if X and Y are independent $\Rightarrow X^2, Y^4$ are independent.

RVs → indepe

independent and identically distributed.

RVs → independent
→ have the same distribution

Example : Independent and Identical Distribution
X and Y two independent dice roll

Dependent and identically distributed

X : # head $\quad Bin(n, \frac{1}{2})$
Y : # Tail

story of Binomial distribution in algebraic form

$X \sim Bin(n,p)$ # success in n indepen Bernoli trial

$$X = X_1 + X_2 + \cdots X_n \quad X_j \sim Bern(p)$$
$$j = 1, \ldots n$$

Theorm if $X \sim Bin(n,p)$ and $Y \sim Bin(m,p)$
and X is independe of Y $(X \perp Y)$
$X + Y \sim Bin(n+m, p)$

proof: 1) LOTP

$$P(X+Y=k) = \sum_{j=0}^{k} P(X+Y=k \mid X=j)P(X=j)$$

$$P(X+Y=k \mid X=j) = P(Y=k-j \mid X=j) = P(Y=k-j)$$

$$\Rightarrow \sum P(Y=k-j)P(X=j) = \binom{n+m}{k} p^k q^{n+m-k}$$
$$\Downarrow$$
$$X+Y \sim Bin(n+m, p)$$

2) representation

$$X = X_1 + X_2 \cdots X_n, \quad Y = Y_1 + \cdots Y_m$$

$X + Y$ is sum of man

conditional independence

$X \perp Y \mid Z$ given a RV Z

$$P(X \le x, Y \le y \mid Z = z) = P(X \le x \mid Z = z) P(Y < y \mid Z = z)$$

for all $x, y \in R$, z in super Z

Fisher exact Test

A random sample of n women & m men tested for disease

	women	men	Total
Disease	x	$r-x$	r
not Disease	$n-x$	$m-(r-x)$	$n+m-r$
	n	m	

Test $P_1 = P_2$

$X \sim Bin(n, P_1)$ $\qquad$ $Y \sim Bin(m, P_2)$

\# of women disease $\qquad$ \# of me disease

Fisher exact test conditions on row & col sums (m, n, r)

$$P(X = x \mid X + Y = r) = \frac{P(X+Y=r \mid X=x) P(X=x)}{P(X+Y=r)}$$

$$= \frac{P(Y = r-x) P(X = x)}{P(X+Y=r)} \Rightarrow \frac{\binom{m}{r-x}\binom{n}{x}}{\binom{m+n}{r}} \Rightarrow H(n, m, r)$$