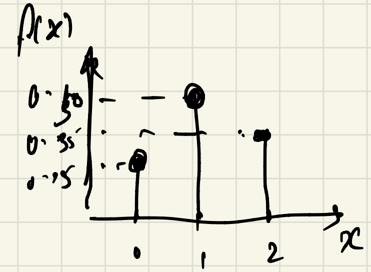
(1) PMF :

for discrete random variable X with possible outcomes $x_1, x_2, \dots x_n$, a PMF (Prob. Mass funct^n)

s.t. (i) $f(x_i) \geq 0$

(ii) $\sum_i f(x_i) = 1$

(iii) $f(x_i) = P(X = x_i)$



$P(X=0) = 0.15$
$P(X=1) = 0.35$
$P(X=2) = 0.50$

(2) CDF :

for discrete R.V X, $F(x)$ is,

$$F(x) = \sum_{x_i \leq x} f(x_i)$$

satisfies the properties;

(i) $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$

(ii) $0 \leq F(x) \leq 1$

(iii) If $x \leq y$, then $F(x) \leq F(y)$

(3) **Mean** : Central value of probability distribution

μ, $E[x]$, $\sum_x x f(x)$

(4) **Variance:** Dispersion or variability in the distribution.

$$\sigma^2 = V(X) = E[x-\mu]^2$$

$$= \sum_x (x-\mu)^2 f(x)$$

$$= \sum_x x^2 f(x) - \mu^2$$

$$= E[x^2] - (E[x])^2$$

$$V(X) = \sum_x (x-\mu)^2 f(x)$$

$$= \sum_x (x^2 + \mu^2 - 2x\mu) f(x)$$

$$= \sum_x x^2 f(x) + \mu^2 \underbrace{\sum_x f(x)}_{=1} - 2\mu \underbrace{\sum_x x f(x)}_{=E[x]}$$

$$= \sum_x x^2 f(x) + \mu^2 - 2\mu \times \mu$$

$$= \sum_x x^2 f(x) + \mu^2 - 2\mu^2$$

$$= \sum_x x^2 f(x) - \mu^2$$

$$= E[x^2] - (E[x])^2$$

(5) Expected value of a funct^n of discrete R.V.

Consider $h(x)$ a function of R.V. X:

$$E[h(X)] = \sum_x h(x) f(x)$$

Interpretation: The E.V. of h is the wt. average of function evaluated at values of the random variable.

Special cases:

$$E[h(X)] \neq h(E[X])$$

Consider $h(x) = aX + b$

$$E[ax + b] = aE(X) + b \quad — (1)$$

$$Var(ax + b) = a^2 V(x) \quad — (2)$$

(1)

$$E[aX + b] = \sum_x h(x) \cdot f(x)$$

$$= \sum_x (ax + b) \cdot f(x)$$

$$= \sum_x ax \cdot f(x) + b \underbrace{\sum_x f(x)}_{=1}$$

$$= a \underbrace{\sum_x x f(x)}_{= E[X]} + b$$

$$= a E[X] + b$$

⑦ $Var(h(x)) = \sum_{x} (h(x) - E[h(x)])^2 \cdot f(x)$

$$= \sum_{x} (ax + \cancel{b} - aE(x) \cancel{+b})^2 \cdot f(x)$$

$$= \sum_{x} (ax - aE(x))^2 \cdot f(x)$$

$$= \sum_{x} (a[x - E(x)])^2 \cdot f(x)$$

$$= a^2 \underbrace{\sum_{x} (x - E(x))^2 \cdot f(x)}_{= Var(x)}$$

$$= a^2 \cdot Var(X)$$

⑥ Uniform Distribution :

A random variable has n values in the range each with equal probability

$$f(x_i) = 1/n$$

Assume interval (a, b) for the example:

Now $E[X] = \sum_{a}^{b} x_i \cdot f(x_i) \Rightarrow f(x_i) = \frac{1}{b-a+1}$

Now $\sum_a^b k = \frac{b(b+1)}{2} - \frac{a(a+1)}{2}$

$= \underline{b(b+1) - a(a+1)}$

$\therefore E[x]$ or $\mu = \sum_a^b x_i \cdot f(x_i)$

$= \frac{1}{(b-a+1)} \sum_a^b x_i$ ~ Arithmetic series whose sum is $n/2$ (First Term + last term)

$= \frac{1}{\cancel{(b-a+1)}} \cancel{(b-a+1)} \frac{(a+b)}{2}$

$\mu = \frac{a+b}{2}$ ∎

$Var(x) = \frac{(b-a+1)^2 - 1}{12}$ when $(b-a+1) = n$

Assume any n, then $E[x] = \sum x_i f(x_i)$

$= \frac{1}{n} \sum x_i = \frac{\cancel{n}(n+1)}{2\cancel{n}}$

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

$Var(x) = E[x^2] - (E[x])^2$

$= \sum x^2 f(x) = 1/n$

$= 1/n \sum_n x^2$

$= 1/\cancel{n} \cdot \cancel{n} \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2$

$= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$

$= \frac{2n^2 + 2n + n + 1}{6} - \frac{n^2 + 2n + 1}{4}$

$= \frac{4n^2 + \cancel{6n} + 2}{12} - \frac{3n^2 + \cancel{6n} + 3}{12}$

$= \frac{n^2 - 1}{12}$ $\boxed{\text{Holds for any } n}$

7 Binomial distribution:

A random experiment consists of n Bernoulli trials s.t.

(i) Trials are independent
(ii) Each trial results in only 2 outcomes, success and failure
(iii) The prob. of success is p and is constant.

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

Assume you run 3 independent trials: where you may only have success / fail outcomes:

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

You can have 0, 1, 2, 3 successes in any of the given order (sequence)

1 = success 0 = failure

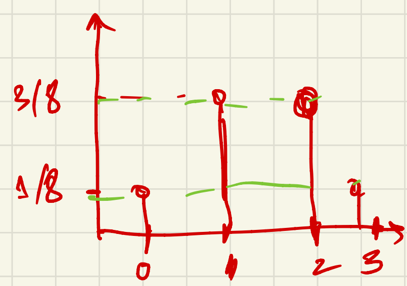
Now $x = 0, 1, 2, 3$

0 = 000 = 1 cnt

1 = 010, 001, 100 = 3 cnt

2 = 011, 110, 101 = 3 cnt

3 = 111 = 1 cnt



$$\sum f(x) = 1 = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = \frac{8}{8}$$

Note: $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$

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

$$= \sum_{k=1}^{n} k \binom{n}{k} p^k q^{n-k}$$

$k\binom{n}{k} = n\binom{n-1}{k-1}$ factors of binomial coefficient.

$$= \sum_{k=1}^{n} n \binom{n-1}{k-1} p^k q^{n-k}$$

→ Take out np

$$= \sum_{k=1}^{n} (np) \binom{n-1}{k-1} p^{k-1} q^{n-k}$$

$n-k = (n-1)-(k-1)$

$$= np \underbrace{\sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{(n-1)-(k-1)}}_{\text{Binomial expansion of } (\underbrace{p+q}_{=1})^{n-1}}$$

$$\boxed{\mu = np}$$

Alternatively consider that each Bernoulli trial is independent and can be treated as a sequence of R.V.s.

$$X_i = X_1, X_2 \dots X_n$$

$$E[X_i] = 1(p) + 0(1-p) = p$$

$$V[X_i] = (1-p)^2 \cdot p + (0-p)^2 \cdot (1-p) = p(1-p)$$

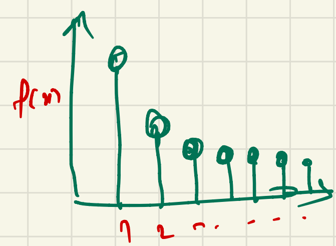
$$\boxed{\therefore\ E[X] = np \qquad Var(X) = np(1-p)}$$

8 Geometric distribution:

In a series of Bernoulli trials a r.v. X equals the # of trials before the first success is a geometric random variable with parameter $0 < p < 1$

$$f(x) = (1-p)^{x-1} p \quad \text{for } x = 1, 2 \dots \infty$$

The probability/height of the line at x is $(1-p)$ times height at $x-1$



Lack of memory:

The trials are independent, the count of # of trials until next success can be started at any trial w/o changing the probability distribution for the R.V.

$$E(X) = \mu = 1/p \qquad \sigma^2 = Var(X) = \frac{1-p}{p^2}$$

Alternatively, we could be looking at the number of failures before hitting 1st success:

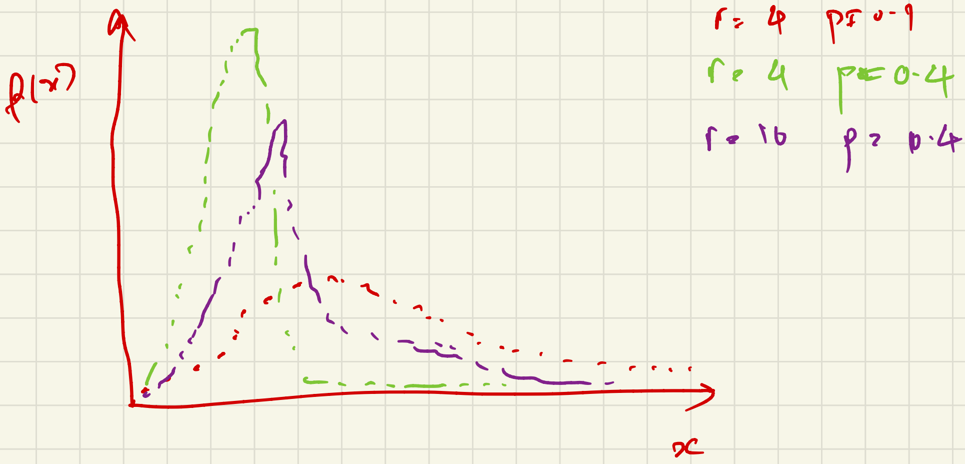
$$E(X) = \mu = \frac{1-p}{p} \qquad \sigma^2 = Var(X) = \frac{1-p}{p^2}$$

(9) Negative binomial :-

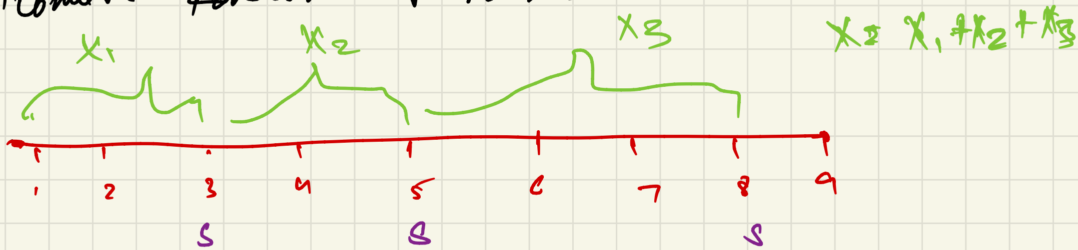
A generalization of geometric distribution in which R.V. is the # of Bernoulli trials required to obtain r successes is negative binomial.

$$f(x) = \binom{x-r}{r-1} p^r q^{x-r}$$

Here, we assume that there are r-1 successes in the first x-1 trials and r^{th} success is obtained on the x^{th} trial.



N.B can also be represented as a sum of Geometric Random Variable.



Alternatively X = Negative Binomial with p

it can also be represented as $Y = X_1 + X_2 + X_3$

where X_1, X_2, X_3 are geometric with param p

$E[X] = \mu = r/p$ $\sigma^2 = Var(X) = r(1-p)/p^2$

(10) Hypergeometric distribution:

A set of N objects
- K are classified as success
- $N-K$ are " " failures

A sample of n objects is selected at random from N objects where $n \le N$ and $K \le N$.
w/o replacement

The R.V. X that equals # of successes is a hypergeometric random variable and

$$f(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

$x = \max\{0, n+K-N\}$ to $\min\{K, n\}$

$E[X] = \mu = np$ $\qquad$ $Var(X) = \sigma^2 = np(1-p)\underbrace{\left(\frac{N-n}{N-1}\right)}$

where $p = K/N$

N = total # objects

K = # of successes

n = # of objects sampled

Finite population correction factor

Sampling is done w/o replacement from an infinite set in binomial trials and thus the corresponding hyper geometric R.V. without replacement would have mean and variance equal to B.R.V.

Without replacement changes the probability of success and thus the variance is lower since we are drawing from a finite set.

(ii) Poisson distribution:

Events occur randomly in a given interval.

The R.V. of interest is the count of events that occur within that interval.

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

$E[X] = \mu = \lambda T \qquad Var(X) = \sigma^2 = \lambda T$

Imagine partitioning the interval into n sub intervals $\Delta t = T/n$.

Furthermore, assume that probability of the flaw occurring is the same across all subintervals. Our variable can now be treated as a Binomial Random Variable.

$E[X] = \lambda T = np \qquad p = \frac{\lambda T}{n}$

$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$ with small subintervals

$$\binom{n}{x}\left(\frac{\lambda T}{n}\right)^x \Rightarrow \frac{\lambda T^x}{x!}\left(1 - \frac{\lambda T}{n}\right)^{-x} \Rightarrow 1 \quad \left(1 - \frac{\lambda T}{n}\right)^n \rightarrow e^{-\lambda T}$$

$$\lim_{n \to \infty} P(X = x) = \frac{e^{-\lambda T} (\lambda T)^x}{x!} \; ; \quad x = 0, 1, 2, \ldots$$