Chapter 1

Background material

1.1 Distribution Summary

1. **Binomial**(n, p); $\quad f(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0, 1, \ldots, n$

range

where $\binom{n}{x} = \frac{n!}{(x!)(n-x)!}$.

Consider n independent repetitions of an experiment each of which has only two possible outcomes, say (S, F) where

$$P\{S\} = p \text{ is constant, i.e. the same for each experiment}$$

Let $X = \#S$'s in n trials

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

Example: Let X be the number of heads in n tosses of a fair coin.

Note: We often use the Binomial Distribution when we **Sample with Replacement.**

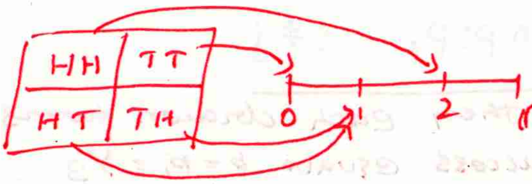
$E(X) = np \quad , \quad Var(X) = np(1-p)$

Let, $n = 2$. What is the Sample space of the experiment i.e, the domain of X?

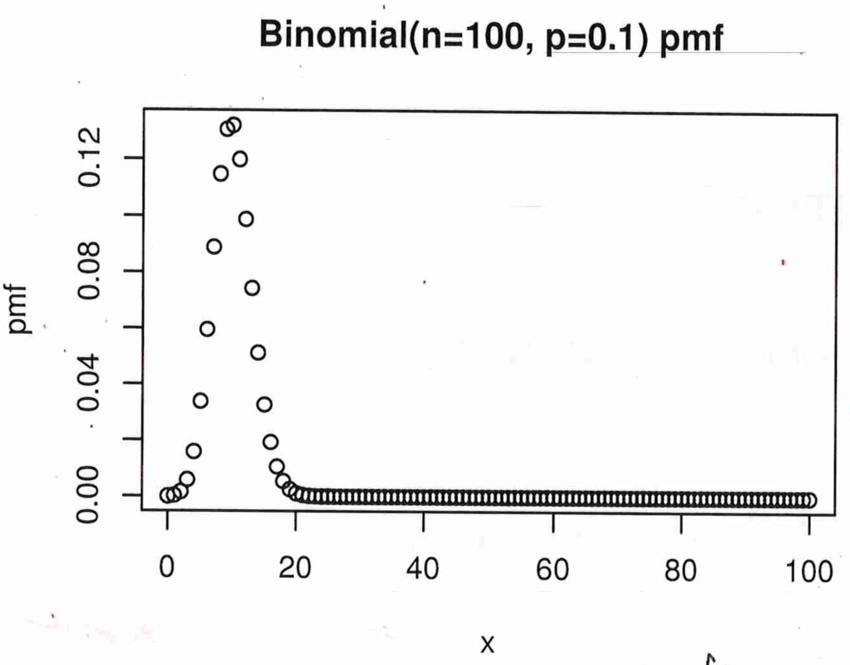
Since, X is the # heads in 2 tosses

$X = \{0, 1, 2\}$

$X \sim Bin(2, p = 0.5)$



$\Omega = \{HH, HT, TH, TT\}$



x is discrete

(pmf)

Figure 1.1: Binomial probability mass function, n=100, p=0.1

2. **Multinomial**$(n, p_1, \ldots, p_k)$; $\quad f(x_1, \ldots, x_k) = \binom{n}{x_1 \ldots x_k} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$

   where $\binom{n}{x_1 \ldots x_k} = \frac{n!}{(x_1!)(x_2)!\cdots(x_k)!}$ and

   $x_i = 0, 1, \ldots, n,$ such that $x_1 + \cdots + x_k = n$ and $p_1 + \cdots + p_k = 1.$

   Consider n independent repetitions of an experiment for which each outcome can be classified in exactly one of k mutually exclusive ways, $A_1, A_2, \ldots, A_k$.

   Let   When $k=2$; this is Binomial(n, p_1)

   $$p_i = P\{\text{an outcome of one trial is of class } A_i\}$$
   $$X_i = \#\text{ outcomes that are of class } i \text{ out of } n \text{ repetitions}$$

   Then $(X_1, X_2, \ldots, X_k) \sim$ Multinomial $(n, p_1, \ldots, p_k)$

   X_i's have marginal Binomial Distributions

   $$E(X_i) = np_i, \quad Var(X_i) = np_i(1-p_i)$$
   $$Cov(X_i, X_j) = -np_ip_j \quad i \neq j$$

We draw marbles from a bag W.R.

$\boxed{\begin{matrix} R & B & R \\ B & B & G \end{matrix}}$

Let, $X_1 = \#$ red, $X_2 = \#$ Blue, $X_3 = \#$ Green marbles

$(X_1, X_2, X_3) \sim$ Multi$(3, 1/3, 1/2, 1/6)$

Now, let's focus on only on red marbles; then each drawn turns into a yes/no trial, with a probability of success equals $p = p_1 = 1/3$ and failure being $1-p = 1-p_1 = 2/3 = 1/2 + 1/6 = p_2 + p_3$.

$\therefore X_1 = $ Bin$(3, 1/3)$ and Similarly, $X_2 \sim$ Bin$(2, 1/3)$, $X_3 \sim$ Bin$(3, 1/6)$

Note: $\sum_{i=1}^{k} p_i = 1 \quad \sum_{i=1}^{k} X_i = n$

Example: Toss a fair die $n = 100$ times and let $(X_1, X_2, \ldots, X_6)$ be the observed frequencies of the numbers $1, 2, 3, 4, 5, 6$ from the tosses of the die. Since the die is fair, then $p_i = 1/6$ for $i = 1, \ldots, 6$.

3. **Negative Binomial**(r, p); $\quad f(x) = \binom{x+r-1}{r-1} p^r (1-p)^x$, $x = 0, 1, \ldots$

Consider independent repetitions of an experiment each of which, has exactly two possible outcomes, say (S, F).

Let $P(S) = p$ constant, i.e. the same for each experiment

Let $X = \#$ F's before the r^{th} S

Then $X \sim \text{NegBin}(r, p)$ X is discrete

Example: Continue flipping a fair coin and stop when you observe the first head. $X =$ the number of tails before the first head has a Negative Binomial distribution with $r = 1$.

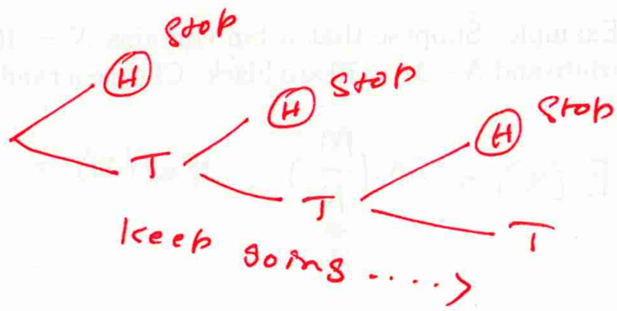
$$E(X) = \frac{r(1-p)}{p}, \quad Var(X) = \frac{r(1-p)}{p^2}$$

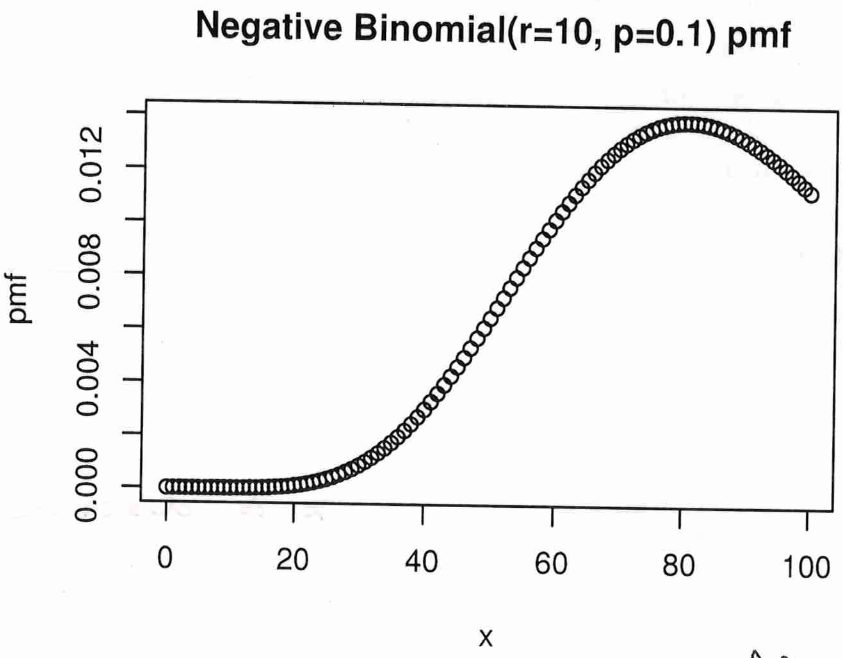
What is the sample space here?

$$\Omega = \{H, TH, TTH, TTTH, \ldots\}$$

How many possible outcomes?

Countably infinitely many outcomes.





(cpmf)

Figure 1.2: Negative Binomial probability mass function, r=10, p=0.1

4. **Geometric**(p); is the same as Negative Binomial $(r = 1, p)$ $f(x) = p(1-p)^x ; x = 0, 1, \ldots$
or, $f(x) = p(1-p)^{x-1} ; x = 1, 2, \ldots$

5. **Hypergeometric(N, M, n)**; $f(x) = \binom{M}{x}\binom{N-M}{n-x}/\binom{N}{n}$ where $\max(0, n - N + M) \le x \le \min(n, M)$.

   Consider a finite population of size N. Let each object in the population be characterized as either a S or F, where there are $M \le N$ S's in the population. Draw a random sample of size n from the population without replacement.

   Let $X = \#\ S$'s in the sample of size n.

   Then $X \sim$ Hypergeometric(N, M, n).

   Example: Suppose that a bin contains $N = 100$ balls, of which $M = 30$ are white and $N - M = 70$ are black. Choose a random sample of $n = 10$ balls from

$$E(X) = n\Big(\underbrace{\tfrac{M}{N}}_{p}\Big), \quad Var(X) = \frac{N-n}{N-1}\, n\Big(\underbrace{\tfrac{M}{N}}_{p}\Big)\Big[1 - \underbrace{\tfrac{M}{N}}_{p}\Big]$$

Explained at the end of this pdf (last page).

the bin without replacement. X = the number of white balls in the sample has a Hypergeometric($N = 100, M = 30, n = 10$) distribution.

Example: A shipping container contains $N = 10,000$ iPhone 7's of which $M = 30$ are defective and the remainder are not defective. Choose a random sample of $n = 100$ iPhone 7's from the shipping container without replacement. Then X = the number of defectives in the sample has a Hypergeometric($N = 10,000, M = 30, n = 100$) distribution.

In this example, $n/N = 100/10,000 = 0.01 \leq 0.05$. Then X = the number of defectives in the sample is approximately distributed as Binomial($n = 100, p = 30/10,000 = 0.003$).

Let, $Y \sim HYG(N, M, n)$

If, $\frac{n}{N} \leq 0.05$

then,

$Y \sim Bin(n, M/N)$

or, $Y \sim Bin(n, p)$

where,

$p = M/N$.

If $\frac{n}{N} \leq .05$, can use Binomial to approximate

when your sample size is less than or equal to 5% of your population, the change in probabilities at each trial from not replacing is negligible

=> can treat the trials as independent.

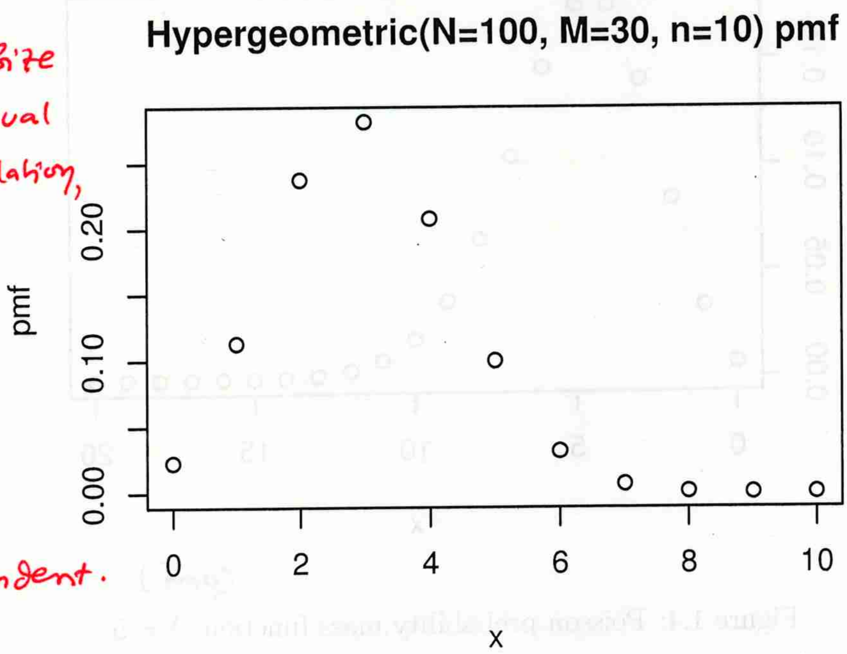


Figure 1.3: Hypergeometric probability mass function, N=100, M=30, n=10

6. **Poisson** (λ); $f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$ $x = 0, 1, \ldots$

Discrete

Models the number of occurrences of random events in space and time, where the average rate, λ per unit time (or area, or volume) is constant.

Recall 1-
pois(np) $\approx$ Bin(n, p)
when $n \geq 20$ and
$np \leq 5$
large n,
small p)

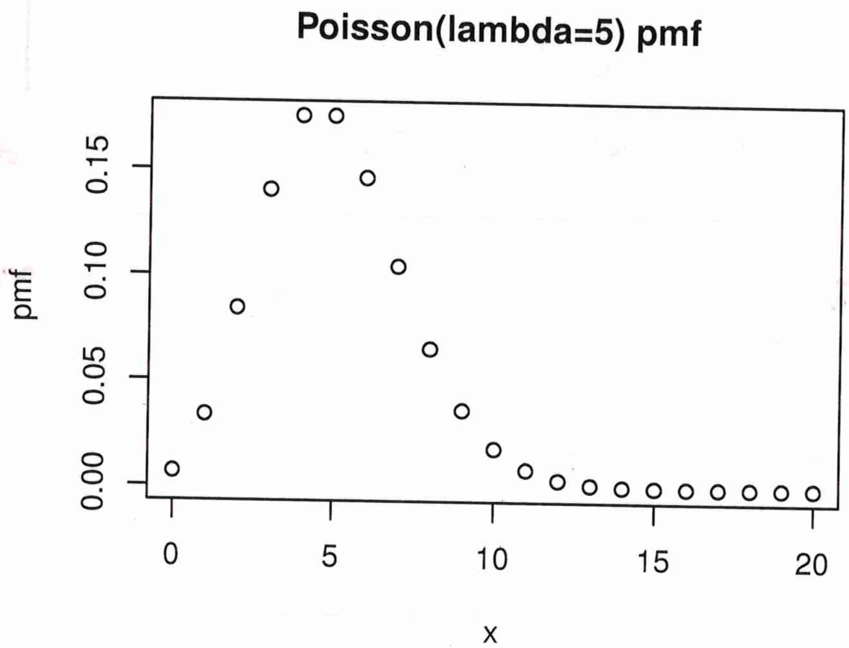
$$\text{Let } X = \#\text{ events in } t \text{ units of time}$$
$$\text{Then } X \sim \text{Poisson } (\lambda t)$$

Example: Let X = the number of customers arriving at a bank in a given one hour time interval.

What is the sample space?

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

$\Omega = \{$Nobody, One person, ... $\}$



(pmf)

Figure 1.4: Poisson probability mass function, $\lambda = 5$

7. **Exponential** (mean θ); $f(x) = \frac{1}{\theta}\, e^{-x/\theta}$ $\quad x > 0 \quad \theta > 0$ Continuous!

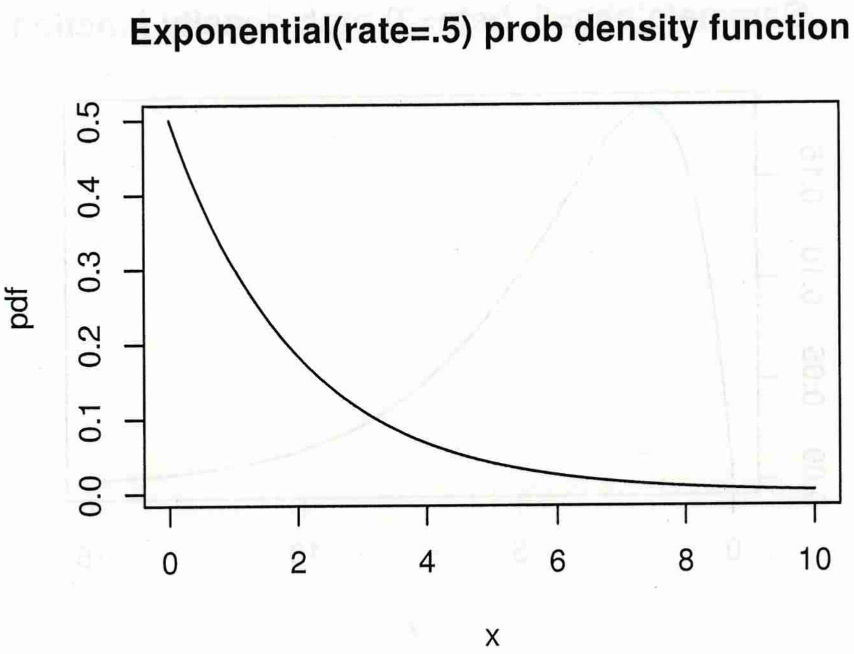
Models lifetimes where there is no deterioration with age - or - waiting times between successive random events in a Poisson process. We also parameterize the exponential distribution using the rate parameter, $\lambda = 1/\theta$.

$E(X) = \theta$, $Var(X) = \theta^2$

Memoryless property

$P(X > a+b \mid X > a) = P(X > b)$,
$a, b > 0$.

$\boxed{\text{If, } X \sim \text{pois}(\lambda) \text{ is counting the \# of customers entering in an hour, then } Y \sim \text{Exp(mean } 1/\lambda) \text{ denotes the time between consecutive entrances.}}$



mean

Figure 1.5: Exponential density function, $\theta = 1/.5 = 2$

λ = rate = .5

8. **Gamma**(α, β); $\quad f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha}\, x^{\alpha-1}\, \exp(-x/\beta),\ x > 0,\ \alpha, \beta > 0.$

$E(X) = \alpha\beta$, $Var(X) = \alpha\beta^2$

flexible model used to model lifetimes.

When $\alpha = 1$, what does this simplify to?

Gamma$(1, \beta) \Rightarrow f(x) = \frac{1}{\Gamma_1\, \beta^1}\, x^{1-1} \exp(-x/\beta)$

$(\Gamma_1 = 0! = 1)$ $\qquad = \frac{1}{\beta} e^{-x/\beta}$

$\therefore$ This becomes, Exp(mean β)

or, Exp(rate $1/\beta$)

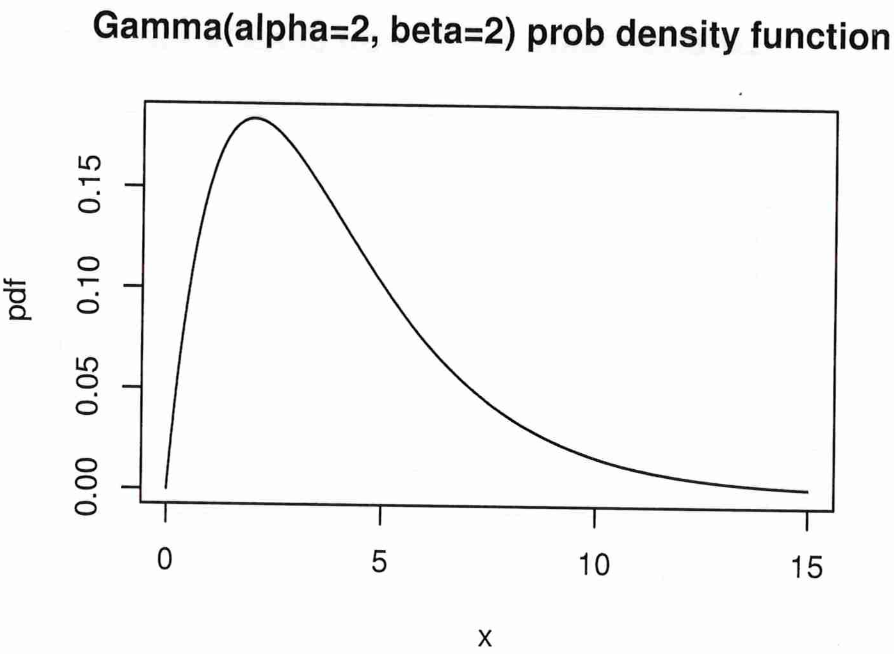


Figure 1.6: Gamma density function, $\alpha = 2, \beta = 2$

location parameter

scale parameter.

9. **Normal** (μ, σ^2); $\quad f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$, $x, \mu \in \Re$, $\sigma^2 > 0$.

   Many measurements are approximately normal.

   mean (μ) ("mu")

   If $X \sim N(\mu, \sigma^2)$,

   Variance (σ^2) ("sigma")

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1).$$

(Standard Normal Distribution)

$E(X) = \mu$, $Var(X) = \sigma^2$, $sd = \sigma$ = standard deviation

We assume $\sigma^2 > 0$, because $\sigma^2 = 0$ is trivial

↪ leads to degenerated dist$^{\underline{n}}$.

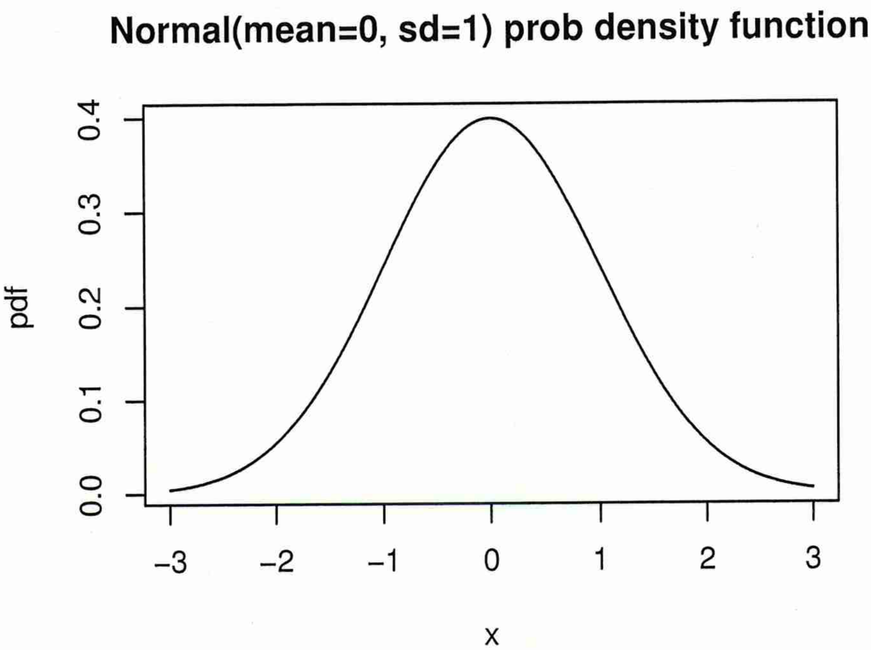


Figure 1.7: Normal density function, $\mu = 0, \sigma = 1$

① If, $X_1 \sim N(\mu_1, \sigma_1^2)$
$X_2 \sim N(\mu_2, \sigma_2^2)$

Notation : $X_1 \perp X_2$

Note: If $X_1, \ldots, X_n$ are independent with $X_i \sim N(\mu_i, \sigma_i^2)$ and $a_1, \ldots, a_n$ are constants,

$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

then $\sum_{i=1}^{n} a_i X_i \sim N(\sum a_i \mu_i, \sum a_i^2 \sigma_i^2)$

★ Linear Combinations of Normal R.V's are Normal.

$X_1 + 2X_2 \sim N(\mu_1 + 2\mu_2, \sigma_1^2 + 4\sigma^2)$

Central Limit Theorem (CLT)

Let $S_n = \sum_{i=1}^{n} X_i$ be the sum of n independent random variables each with mean μ, variance σ^2. Then

$$\boxed{\frac{S_n - n\mu}{\sigma\sqrt{n}}} \approx N(0,1) \quad \text{for large } n,$$

where $\approx$ means approximately distributed as.

Alternatively, the CLT is expressed as

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \approx N(0,1)$$

finite variance
$\sigma^2 < \infty$.

From Jan 9 lecture [Maybe helpful for your Lecture Assignment-1

Let, X be a R.V (random variable)

1) If, X is discrete, the distribution function is called probability mass function (pmf)

$$f(x) = \mathbb{P}(X = x) \ ; \ \forall x \in \mathcal{X}$$

for all x belonging to $\mathcal{X}$.

$f(x)$ has these constraints

ⓐ $f(x) \geq 0 \ ; \ \forall x \in \mathcal{X}$

ⓑ $\sum_{x \in \mathcal{X}} f(x) = 1.$

2) If, X is Continous, then it can take on values in an interval {subset of $\mathbb{R}$} and we call the distribution function 'probability density function' (pdf)

Consider; $F(x) = \mathbb{P}(X \leq x) \ ; \ \forall x \in \mathcal{X}$

The pdf is defined as : $f(x) = \frac{d}{dx} F(x)$

$$\Longleftrightarrow$$

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

Fundamental Theorem of Calculus

$f(x)$ has the following Constraints :

ⓐ $f(x) \geq 0 \ ; \ \forall x \in \mathcal{X}$

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

ⓒ $\mathbb{P}(X = x) = 0 \ ; \ \forall x \in \mathcal{X}$

What does it mean for $X = x$?

↳ Both $X \geq x$ and $X \leq x$ are true

so, $\mathbb{P}(X = x) = \mathbb{P}(x \leq X \leq x)$

$= F(x) - F(x) = 0$

Solution of the treat question from Jan 10

* Consider a bin of $N = 100$ marbles, with $M = 30$ white marbles and $N - M = 70$ black marbles. Suppose, we sample $n = 40$ marbles from the bin without replacement. How many possible outcomes make up the sample space of this experiment? How about with $n = 10$ marbles?

Sol :→ For, $\boxed{n = 40}$

2 types of marbles → black (× 70) B
→ white (× 30) W

A string of length 40 consisting of B's and W's or, may be no W's, only B's, because # white $< n$ $(n = 40)$

Let, r be the # white marbles drawn

$\therefore \quad 0 \leq r \leq 30$

For each r, there are $\binom{n}{r}$ ways to arrange the white and black marbles,

Therefore, $|\Omega| = \sum_{r=0}^{30} \binom{n}{r} = \sum_{r=0}^{30} \binom{40}{r} = \sum_{r=0}^{30} \frac{40!}{r!(40-r)!}$

i.e, Case 1; $r = 0$, $\frac{40!}{0!(40-0)!} = 1$

Case 2; $r = 1$, $\frac{40!}{1!\,39!} = 40$

$\vdots \qquad \vdots \qquad \vdots$

Case 31; $r = 30$, $\frac{40!}{30!\,10!}$

Add up all the Cases

$= \sum_{r=0}^{30} \frac{40!}{r!(40-r)!}$ (Total Outcomes)

For, $\boxed{n = 10}$ should be $(r \leq n)$

Similarly, $\sum_{r=0}^{10} \frac{10!}{r!(10-r)!} = 1024 = 2^{10}$ (Total Outcomes)

In this case, we can simply think, both the # white and Black marbles are $> n$, so, there will be 2^n outcomes.

Let see, how this problem is related to an hyper Geometric distribution,

What is the probability of getting 6 white balls
a) in the sample of size 10 ?
b) in the sample of size 40 ?

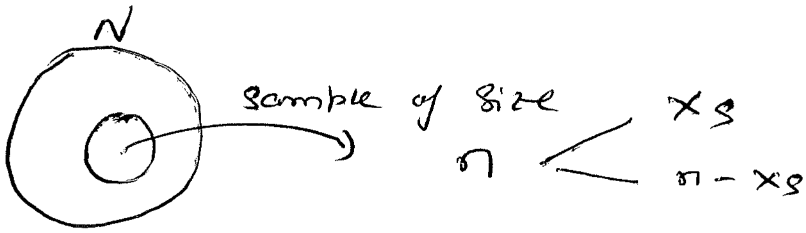
↳ Here, $N = 100$, $M = 30$ (white balls)

a) $n = 10$; b) $n = 40$

$$P(X=6) = \frac{\binom{30}{6}\binom{70}{4}}{\binom{100}{10}} \quad ; \quad P(X=6) = \frac{\binom{30}{6}\binom{70}{36}}{\binom{100}{40}}$$

Hyper Geometric Distribution

$$N \begin{cases} M \\ N-M \end{cases}$$



⊙ "N" contains "M" number of Success "S"
then, there are "N−M" number of Failure "F"
we are taking a sample of size "n" from "N"
and, we are interested in how many Success "S" are there in the sample of size "n"

∴ Let X be a r.v, which defines the # S in the sample of size n.

$$\therefore P(X=x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$