• Discrete Random Variables

- The no. of free throws an NBA player makes in his next 20 attempts.

  e.g.– Possible Values : 0, 1, 2, . . . , 20

- The no. of rolls of a die needed to roll a 3 for the first time.

  e.g.– Possible Value : 1, 2, 3, . . .

- The Profit on a \$1.50 bet on Black in roulette

  e.g.–Possible Value : −1.50, 1.50

Ex: Approximately 60% of the full time newborn babies develop jaundice. Suppose we randomly sample 2 full-time newborn babies and let X represent the no. ~~that~~ develop jaundice. What is the probability distribution of X?

	JJ	JN	NJ	NN
Value of X	2	1	1	0
Probability =	0.6×0.6	0.6×0.4	0.4×0.6	0.4×0.6

Probability distribution of X

X	0	1	2
P(x)	0.16	0.48	0.36

$P(X=0) = 0.16$

$P(X = x) \equiv P(x)$

Random Variable

Value of Random Variable

All discrete Prob. distributions must satisfy

① $0 \le p(x) \le 1$ for all x

② $\sum_{x=0}^{n} p(x) = 1$

x	0	1	2
$p(x)$	0.16	0.48	0.36

$p = 0.6$

$p(x) = {}^{2}C_{x} (0.6)^{x} (1-0.6)^{2-x}$

Prob. Mass function for $x = 0, 1, 2$

⊛ <u>Expectation and Variance of Discrete Random Variable</u>

Probability distribution of D.R.V. X

X	0	1	2
$p(x)$	0.16	0.48	0.36

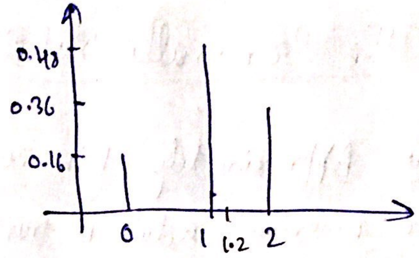
$E(X) = \mu = \sum_{x=0}^{n} x \cdot p(x)$ expected value (expectation)

$E[g(x)] = \sum_{x=0}^{n} g(x) \cdot p(x)$ expectation of a function $g(x)$

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

Ques Suppose you bought a novelty coin that has a prob. of 0.6 of coming up heads when flipped. Let X represent the no. of heads when this coin is tossed twice

x	0	1	2
$p(x)$	0.16	0.48	0.36



What is $E(X)$

$$E(x) = \sum x \cdot p(x)$$

$$= 0 \times 0.16 + 1 \times 0.48 + 2 \times 0.36$$

$$= 1.2$$

$$\boxed{\mu = E(x) = 1.2}$$

$$E(x^2) = \sum x^2 \cdot p(x)$$

$$= 0^2 \times 0.16 + 1^2 \times 0.48 + 2^2 \times 0.36$$

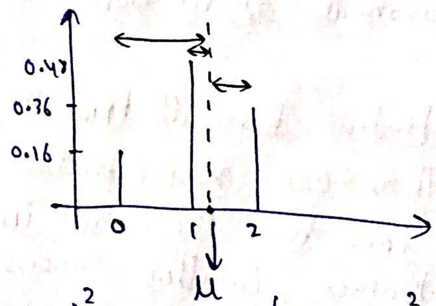
$$= 1.92$$

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

The expectation of the distance from the Mean

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

$$\sigma^2 = \sum (x-\mu)^2 \cdot p(x)$$



$$\sigma^2 = (0-1.2)^2 \times 0.16 + (1-1.2)^2 \times 0.48 + (2-1.2)^2 \times 0.36$$

$$\boxed{\sigma^2 = 0.48}$$

$$\sigma = \sqrt{0.48} \quad (S.D)$$

Que Approximately 1 in 200 American adult are lawyer. One American adult is randomly selected. what is the distribution of the no. of lawyers?

Solⁿ Bernoulli with $p = \frac{1}{200}$

$$P(X=x) = p^x (1-p)^{1-x}$$

$$P(X=1) = \frac{1}{200} \qquad P(X=0) = \frac{199}{200}$$

- **Binomial Distribution**

The no. of success in n independent Bernoulli trials has a binomial distribution.

⊛ Independent Bernoulli trial means

- There are n independent trials.
- Each trial can result in one of the two possible outcomes, labelled success and failure.

$$P(\text{success}) = p \qquad P(\text{Failure}) = 1-p$$

— Success and failure are mutually exclusive and exhaustive.

⊛ X represent the no. of success in n trials.

X has a binomial distribution

$$P(X=x) = {}^nC_x \, p^x (1-p)^{n-x}$$

$$x = 0, 1, \ldots n$$

$\mu = E(X) = n \cdot p$

$\sigma^2 = n \cdot p \cdot (1-p)$

Ques A balanced, six-sided die is rolled 3 times. What is the prob. a 5 comes up exactly twice?

Success = 5 Rolling.

Failure = Rolling anything but a 5

X represent the no. of fives in 3 rolls

X is a binomial distribution with $n = 3$ and $p = 1/6$

$P(X=x) = {}^{n}C_{x}\, p^{x} (1-p)^{n-x}$

$P(X=2) = {}^{3}C_{2} \left(\frac{1}{6}\right)^{2} \left(1-\frac{1}{6}\right)^{3-2}$

$= 0.0694$

Ques According to statistics Canada life tables, the prob. a randomly selected 90 year-old Canadian male survives for at least another year is approximately 0.82.

If twenty 90-year old Canadian males are randomly selected, what is the prob. exactly 18 survive for at least another year?

Success = The Man survive for at least one year

Failure = The Man dies within one year.

Let X represent the no. of men that survive for at least one year.

$n = 20 \quad p = 0.82$

$P(X=18) = {}^{20}C_{18} (0.82)^{18} (1-0.82)^{20-18} = 0.173$

Scanned with CamScanner

$\mu = np = 20 \times (0.82) = 16.4$

$\sigma^2 = n \cdot p \cdot (1-p) = 20 \times (0.82)(1-0.82)$

$\sigma = 2.952$

$\sigma = \sqrt{2.952}$

What is the probability that at least 18 survive for a year?

$P(x \geq 18) = P(x=18) + P(x=19) + P(x=20)$

$= {}^{20}C_{18} (0.86)^{18} (1-0.82)^{2} + {}^{20}C_{19} (0.86)^{19} (1-0.86)$

$+ {}^{20}C_{20} (0.86)^{20} (1-0.86)^{0}$

$= 0.173 + 0.083 + 0.019$

$= 0.275$

- <u>Geometric Distribution</u>

The geometric distribution is the distribution of the no. of trials needed to get the first success in repeated Bernoulli trials.

- For the first success to occur on the x^{th} trial;
  1) The first $x-1$ trials must be failures $(1-p)^{x-1}$
  2) The x^{th} Trial must be a success p

$P(X=x) = (1-p)^{x-1} p$

$$\mu = 1/p$$

$$\sigma^2 = \frac{1-p}{p^2}$$

Ques In a large population of adults, 30% have recieved CPR training. If adults from this population are randomly selected. what is the probability that the 6th person sampled is the first that has recieved CPR training?

$$P(X=x) = (1-p)^{x-1} p$$

$$P(X=6) = (1-0.3)^5 \, 0.3$$

$$P(X=6) = 0.0504$$

$$\mu = 1/p = \frac{1}{0.3} = 3.3$$

$$\sigma^2 = \frac{1-p}{p^2} = \frac{1-0.3}{0.3^2} = 7.7$$

$$\sigma = \sqrt{7.7}$$

Ques What is the probability that the first person trained in CPR occurs on or before the 3rd person sample

$$P(X \leq 3) = P(X=1) + P(X=2) + P(X=3)$$

$$= (1-0.3)^0 \, 0.3 + (1-0.3)^1 \, 0.3 + (1-0.3)^2 \, 0.3$$

$$= 0.3 + 0.7 \times 0.3 + (0.7)^2 \times 0.3$$

$$= 0.657$$

$P(X > 3) = 0.7^3$

$P(X \leq 3) = 1 - 0.7^3$

$= 0.657$

* CDF for geometric distribution

$$F(x) = P(X \leq x) = 1 - (1-p)^x \quad \text{for } x = 1, 2, 3, \ldots$$

- **Poisson Distribution**

Counting the no. of occurrences of an event in a given unit of time, distance, area or volume.

for e.g ① The no. of car accidents in a day

② The no. of dandelions in a square meter plot of land.

- Event are occure independently
- The prob. that an event occures in a given length of time does not change through time.

(Event are occuring randomly and independently)

X, the no. of events in a fixed unit of time, has a poisson distribution.

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad e \approx 2.71828$$

for $x = 0, 1, 2, \ldots$

$\sigma^2 = \lambda$

$\mu = \lambda$

Ques One nanogram of Plutonium-239 will have an average of 2.3 radioactive decays per second, and the no. of decays will follow a Poisson distribution. What is the prob. that in a 2 second period there are exactly 3 radioactive decays?

Solⁿ Let x represent the no. of decays in a 2 second period.

$$\lambda = 2.3 \times 2$$
$$= 4.6$$

$$P(x=3) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{(4.6)^3 e^{4.6}}{3!} = 0.163$$

$$\mu = \lambda = 4.6$$
$$\sigma^2 = 4.6$$
$$\sigma = \sqrt{4.6}$$

⊗ What is the prob. there are no more than 3 radioactive decays? $(\lambda = 4.6)$

$$P(x \le 3) = P(x=0) + P(x=1) + P(x=2) + P(x=3)$$

$$= \frac{(4.6)^0 e^{-4.6}}{0!} + \frac{(4.6)^1 e^{-4.6}}{1!} + \frac{(4.6)^2 e^{-4.6}}{2!} + \frac{(4.6)^3 e^{-4.6}}{3!}$$

$$= 0.010 + 0.046 + 0.106 + 0.163$$
$$= 0.326$$

- **The Relation b/w the Binomial and Poisson Distribution**

The binomial distribution tends toward the Poisson distribution as $n \to \infty$, $p \to 0$ and np stays Constant.

The Poisson distribution with $\lambda = np$ closely approximates the binomial distribution if n is large and p is small.

- **Negative Binomial Distribution**

negative Binomial distribution is the distribution of the number of trials needed to get the r^{th} Success.

Differen b/w Binomial and negative binomial

The binomial distribution is the distribution of the no. of successes (X) in a fixed no. (n) of independent Bernoulli trials.

The negative binomial distribution is the distribution of the no. of trials (X) needed to get a fixed no. of successed (r).

For the r^{th} success to occur on the x^{th} trial

- The first $x-1$ trials must result in $r-1$ success.

$$^{x-1}C_{r-1}\, p^{r-1} (1-p)^{(x-1)-(r-1)}$$

- The x^{th} trial must be a success which has a prob. of p.

The prob. the r^{th} success occurs on the x^{th} trial is

$$P(X=x) = p \times {}^{x-1}C_{r-1}\, p^{r-1} (1-p)^{(x-1)-(r-1)}$$

$$P(X=x) = {}^{x-1}C_{r-1}\, p^{r} (1-p)^{x-r}$$

for $x = r, r+1, \ldots \infty$

$$\mu = r/p$$

$$\sigma^2 = \frac{r(1-p)}{p^2}$$

Ques A person conducting telephone surveys must get 3 more completed surveys before their job is finished. One each randomly dialed no., there is a 9% chance of ~~read~~ reaching an adult who will complete the survey. What is the prob. the 3^{rd} completed survey occurs on the 10^{th} call?

$P(X=10)$

$p = 0.09$

$r = 3$

$$P(X=x) = {}^{x-1}C_{r-1}\, p^{r} (1-p)^{x-r}$$

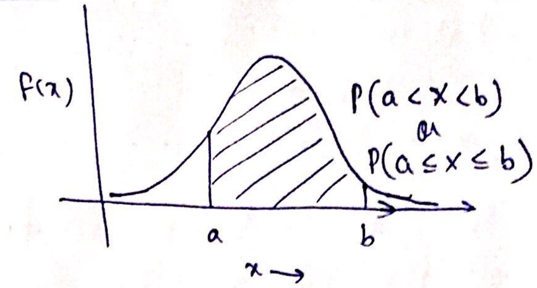
$$P(X=10) = {}^{10-1}C_{3-1}\, (0.09)^3 (1-0.09)^{10-3}$$

$$= 0.01356$$

$$\mu = r/p = \frac{3}{0.09} =$$

- Continuous Random Variables

Area under the entire curve $= 1$



Que Suppose for a random variable X:

$$f(x) = cx^3 \quad \text{for } 2 \le x \le 4$$

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

What value of c makes this a legitimate probability distribution?

Sol^n

- Two condition must satisfy for legitimate distribution
  ① Never take on negative value
  ② Area under the entire curve must equal 1.

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

So,

$$\int_{-\infty}^{2} f(x)\,dx + \int_{2}^{4} f(x)\,dx + \int_{4}^{\infty} f(x)\,dx = 1$$

$$0 + \int_{2}^{4} cx^3\,dx + 0 = 1 \quad \text{——— (1)}$$

$$c\int_{2}^{4} x^3\,dx = c\left[\frac{x^4}{4}\right]_2^4 = c\left[\frac{4^4}{4} - \frac{2^4}{4}\right]$$

$= 60C$

Put it in ①

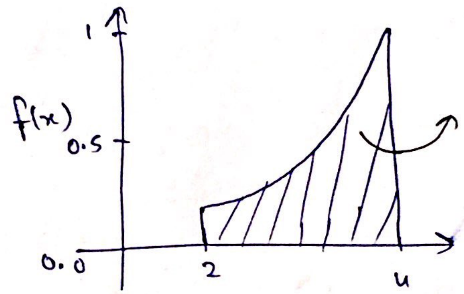
$60C = 1$

$\Rightarrow \boxed{C = 1/60}$

this is the Constant which make above Probability distribution a Legitimate probability distribution.

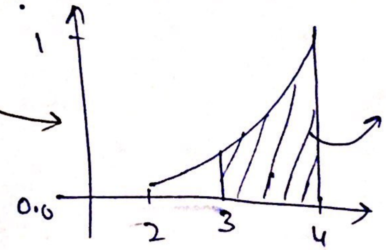
$$f(x) = \frac{1}{60}x^3 \quad \text{for } 2 \le x \le 4$$
$$0 \quad \text{otherwise}$$

Hence, the Curve is



⊛ What is $P(x>3)$?

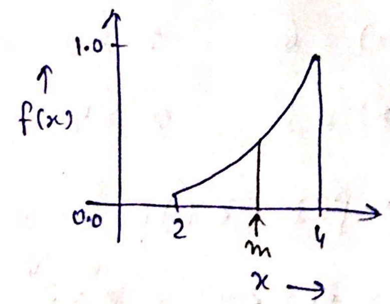
$f(x) = \frac{1}{60}x^3$ for $2 \le x \le 4$



$$= \int_3^4 \frac{1}{60} x^3 dx$$
$$= \frac{1}{60}\left[\frac{x^4}{4}\right]_3^4$$
$$= \frac{1}{60}\left[\frac{4^4}{4} - \frac{3^4}{4}\right] = \frac{175}{240} \approx 0.729$$

What is the median of this distribution?

= Value of the variable that splits the area into halfs



Area to the Left of m is half

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

$$0 + \int_{2}^{m} \frac{1}{60} x^3\,dx = \frac{1}{2}$$

$$\frac{1}{60}\left[\frac{x^4}{4}\right]_2^m = \frac{1}{2}$$

$$\frac{1}{60}\left[\frac{m^4}{4} - \frac{2^4}{4}\right] = \frac{1}{2}$$

$$\frac{m^4 - 2^4}{240} = \frac{1}{2}$$

$$\Rightarrow m^4 = 136$$

$$\boxed{m = \sqrt[4]{136}} \cong 3.415$$

What is the cumulative distribution function $F(x) = P(X \leq x)$?

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

$$= 0 + \int_{2}^{x} \frac{1}{60} t^3\,dt \;*$$

$$= \frac{1}{60}\left[\frac{t^4}{4}\right]_2^x$$

$$= \frac{1}{60}\left[\frac{x^4}{4} - \frac{2^4}{4}\right]$$

$$F(x) = \frac{x^4 - 2^4}{240} \quad \text{for } 2 \le x \le 4$$

Hence, $F(x) = 0 \quad x < 2$

$F(x) = 1 \quad x > 4$

⊛ What is $P(X \le 2.7)$?

$$F(x) = P(X \le x) = \frac{x^4 - 2^4}{240}$$

$$F(2.7) = \frac{(2.7)^4 - 16}{240}$$

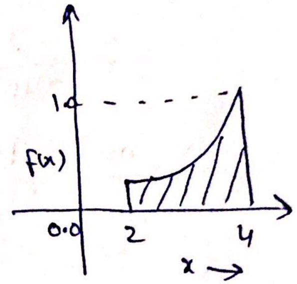
$$\boxed{F(2.7) \cong 0.155}$$

Driving Mean and Variance

Suppose for a R.V X:

$$f(x) = \frac{1}{60}x^3 \quad \text{for } 2 \le x \le 4$$

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



$$f(x) = \frac{1}{60}x^3$$

⊛ Finding the Mean and Variance of this distribution

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

$$= \int_{-\infty}^{2} x f(x)\,dx + \int_{2}^{4} x f(x)\,dx + \int_{4}^{\infty} x f(x)\,dx$$

$$= 0 + \int_2^4 x \cdot \frac{1}{60} x^3 \, dx + 0$$

$$= \frac{1}{60} \int_2^4 x^4 \, dx$$

$$= \frac{1}{60} \left[\frac{x^5}{5} \right]_2^4 = \frac{1}{60} \left[\frac{4^5}{5} - \frac{2^5}{5} \right]$$

$$E(x) = \frac{\cancel{992}}{\cancel{300}} \; \frac{248}{75} \cong 3.3$$

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

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

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

$$= \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{60} x^3 \, dx$$

$$= \frac{1}{60} \int_{-\infty}^{\infty} x^5 \, dx$$

$$= 0 + \frac{1}{60} \int_2^4 x^5 \, dx + 0$$

$$= \frac{1}{60} \left[\frac{x^6}{6} \right]_2^4 = \frac{1}{60} \left[\frac{4^6}{6} - \frac{2^6}{6} \right]$$

$$= \frac{4032}{360}$$

$$E(x^2) = 56/5 \cong$$

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

$$= \frac{56}{5} - \left(\frac{248}{75}\right)^2$$

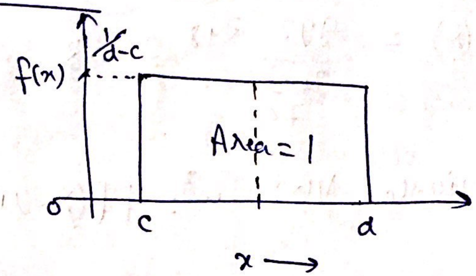
$$= \frac{1496}{5625} \cong 0.266$$

Continuous Uniform Distribution

Area = base × height

$= (d-c)\, f(x) = 1$

$\Rightarrow f(x) = \frac{1}{\cancel{dc}}\ \frac{1}{d-c}$



Q: The Pdf of the Uniform distribution

$$f(x) = \begin{cases} \frac{1}{d-c} & \text{for } c \le x \le d \\ 0 & \text{elsewhere} \end{cases}$$

Median $= \frac{c+d}{2}$ $\qquad \sigma^2 = \frac{1}{12}(d-c)^2$

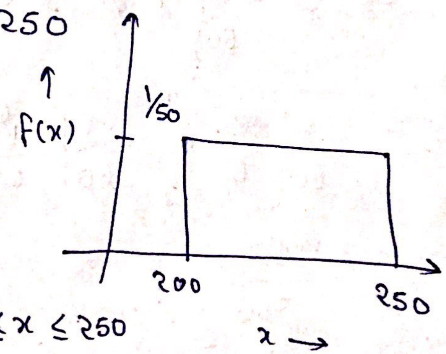
$\mu = \frac{c+d}{2}$

Q: Suppose X is a R.V that has a Uniform distribution with $c=200$ and $d=250$

$f(x) = \frac{1}{d-c}$

$= \frac{1}{250-200}$

$f(x) = \frac{1}{50}$ for $200 \le x \le 250$

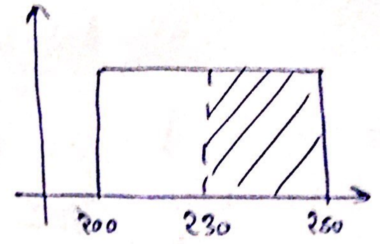


$= 0$ otherwise

(2) What is ~~P(X > 230)~~ $P(X > 230)$

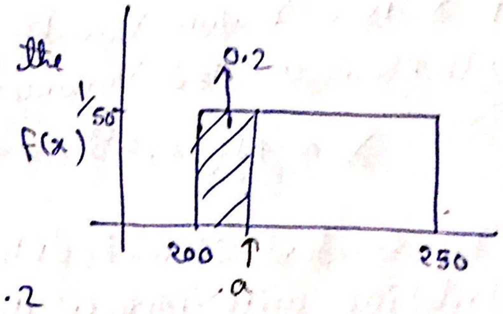
$P(X>230) = (250-230)\frac{1}{50}$

$\boxed{P(X>230) = 0.4}$



(3) What is the 20^{th} percentile of this distribution

find variable that have area to the left is 0.2



$b \times h = 0.2$

$(a-200) \times \frac{1}{50} = 0.2$

$\boxed{a = 210}$

- **Normal Distribution**

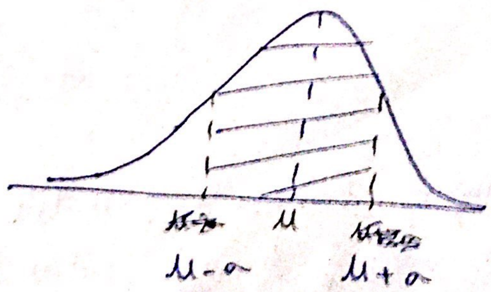
The probability density function (pdf) is:

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad \text{for } -\infty < x < \infty$$

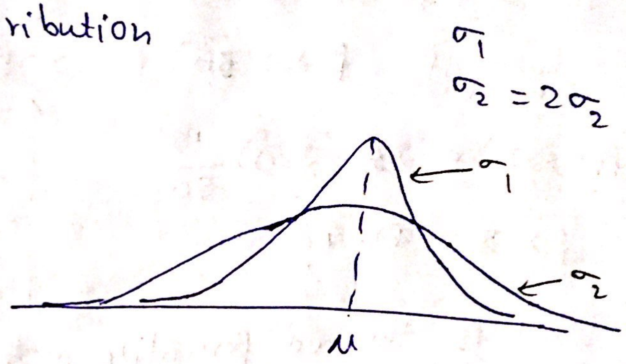
$\Rightarrow -\infty < \mu < \infty$

$\Rightarrow \sigma > 0$

$\Rightarrow \sigma^2 > 0$



* Two normal distribution



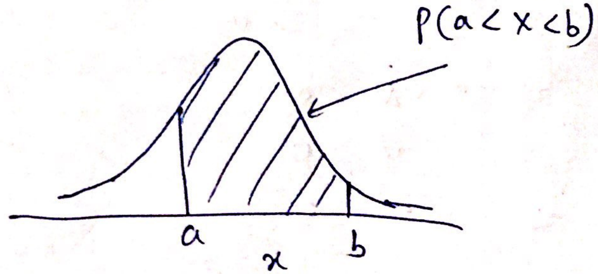
* If X is a random variable that has a normal distribution with mean μ and variance σ^2, we write this as

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

* The standard normal distribution is a normal distribution with mean 0 and variance 1

$$Z \sim N(\underset{\mu}{0}, \underset{\sigma^2}{1})$$

* Area under the standard normal curve



* Finding prob. and percentile requires integrating the probability density function (pdf)

- <u>Standardizing Normally distributed R.V.</u>

Suppose X is a normally distributed R.V with Mean (μ) and standard deviation (σ)

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

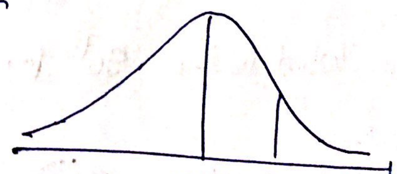
$$Z = \frac{X - \mu}{\sigma}$$ then Z has the SND

$$Z \sim N(0,1)$$

Suppose the height of adult American Male is approximately normally distributed with a Mean of 176.3 Cm and a standard deviation of 7.1 cm

Ans What is the Probability a randomly selected adult American male is taller than 180.0cm

$$P(X > 180.0)$$



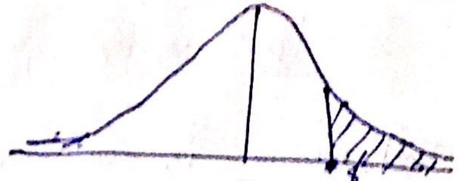
Substract μ and divide by σ both side,

$\left.\begin{array}{l}\sigma \text{ is (+)ive} \\ \mu \text{ is (+)ive}\end{array}\right\}$ so, Value of $\frac{X - \mu}{\sigma}$ is (+)ive

$$P\left(\frac{X-\mu}{\sigma} > \frac{180.0-\mu}{\sigma}\right)$$

$$Z = \frac{X-\mu}{\sigma} \Rightarrow P\left(Z > \frac{180.0 - 176.3}{7.1}\right)$$

$P(Z > 0.521)$



0

Area under the SN Curve to the Right of 0.521.

0.521

- What is the prob. a randomly selected adult American male has a height between 170.0 and 180.0 cm?

$P(170.0 < X < 180.0)$

$P\left(\frac{170.0-176.3}{7.1} < x < \frac{180.0-176.3}{7.1}\right)$ 170.0 176.3 180.0

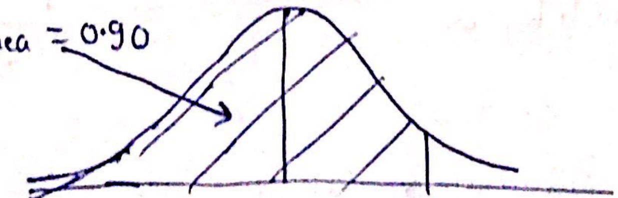
$P(-0.887 < Z < 0.521) \cong 0.511$

Area $= 0.511$

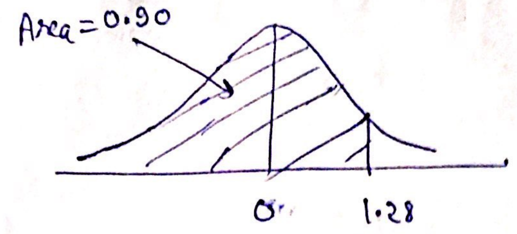
-0.887 0 0.521

- What is the 90th percentile of the height of American Males?

Area $= 0.90$



Using standard Normal table



$$Z = \frac{X - \mu}{\sigma}$$

$$X = \mu + \sigma Z$$

$$X = 176.3 + (7.1 \times 1.28)$$

90th percentile of the distribution $= 185.4$

Answer $= 185.4$ cm (~~height of American male~~ of the 90th percentile)