Population   Sample

Mean
Variance / Standard deviation

Parameter: These are the entities that summarize data for an entire population.

Statistic: These are the entities that summarize data for a sample.

Note 1: Values of statistics will vary from one sample to another.

Note 2: The difference in the values of a statistic is called sampling fluctuations.

Note 3: We can construct a distribution of a statistic, which is called the sampling distribution of that statistic.

	Parameter	Statistic
Mean	μ	$\bar{X}$, $\bar{x}$
Variance	σ^2	s^2
Standard deviation	σ	s

Q1 GOI wants to estimate average height of mens aged 20 years or older. From a random sample of 45 men's, research obtaing a sample mean height of 63.9 inches. List the parameter and statistic for this study.

Parameter: AH of mens aged 20 years or older (Unknown)

Statistic: AH of 63.9 inches from the sample of 45 mens.

Q2 Determine # Kms the average person in India drive a car in one day.

P: # Km the average person in India drive a car in one day.

$\bar{x}$

S: 25.6 Km a person in India drive a car in one day

(I) Sample Size (n) ≡ n=10

(II) No of Samples (m) ≡ m=6

Samp# (10)	Average Km
1	25.6
2	50.2
3	15.1
4	43.9
5	36.8
6	60.2

$\bar{x}$

Sample Size:

a) Fixed:

$\bar{x} \equiv \bar{X}$

$f(x) \equiv PDF \equiv$ Distributions

$\int_{R_X} f(x)\,dx = 1, \; f(x) \geq 0$

Sampling distribution of Mean

$n = 10$

# Sam	Avg e/cm	
1.	25.6	x_1
2	50.2	x_2
3	15.1	x_3
4	43.9	
5	36.8	
6	60.2	x_6

(b) SS is not fixed:

$n=10 \Rightarrow X_1 \rightarrow E(X_1)$

$n=20 \Rightarrow X_2 \rightarrow E(X_2)$

$n=30 \Rightarrow X_3 \rightarrow E(X_3)$

$\vdots$

ε_1, ε_2, $\vdots$ Standar error of the mean

$\Downarrow$

It measures the variability in the sampling distribution

$\Downarrow$

Standar deviation of the mean

Sample Mean:

Let $X_i : i = 1, 2, 3, \ldots, N$ form a population each having mean μ and variance σ^2 ($E(X_i) = \mu$ $Var(X_i) = \sigma^2$ $\forall i$)

Construct a R.Sample of this population with $n << N$ observation, let they be $X_1, X_2, \ldots, X_n$. Then, the sample mean is given by

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \quad (RV's)$$

$$Mean(\bar{X}) = E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right) = \frac{1}{n}\left(\sum_{i=1}^{n} E(X_i)\right)$$

$$= \frac{n\mu}{n} = \mu$$

$$\Rightarrow \boxed{E(\bar{X}) = \mu}$$

$$Variance(\bar{X}) = Var\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right)$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2} n \cdot \sigma^2$$

$$= \sigma^2/n$$

S.tD $= \sqrt{var(\bar{X})} = \dfrac{\sigma}{\sqrt{n}}$

The sample variance:

Let $X_1, X_2, - X_n$ be a random sample where each RV is having mean (μ) and variance (σ^2). Let $\bar{X}$ be the sample mean. Then, the sample variance can be defined as

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \Rightarrow RV's$$

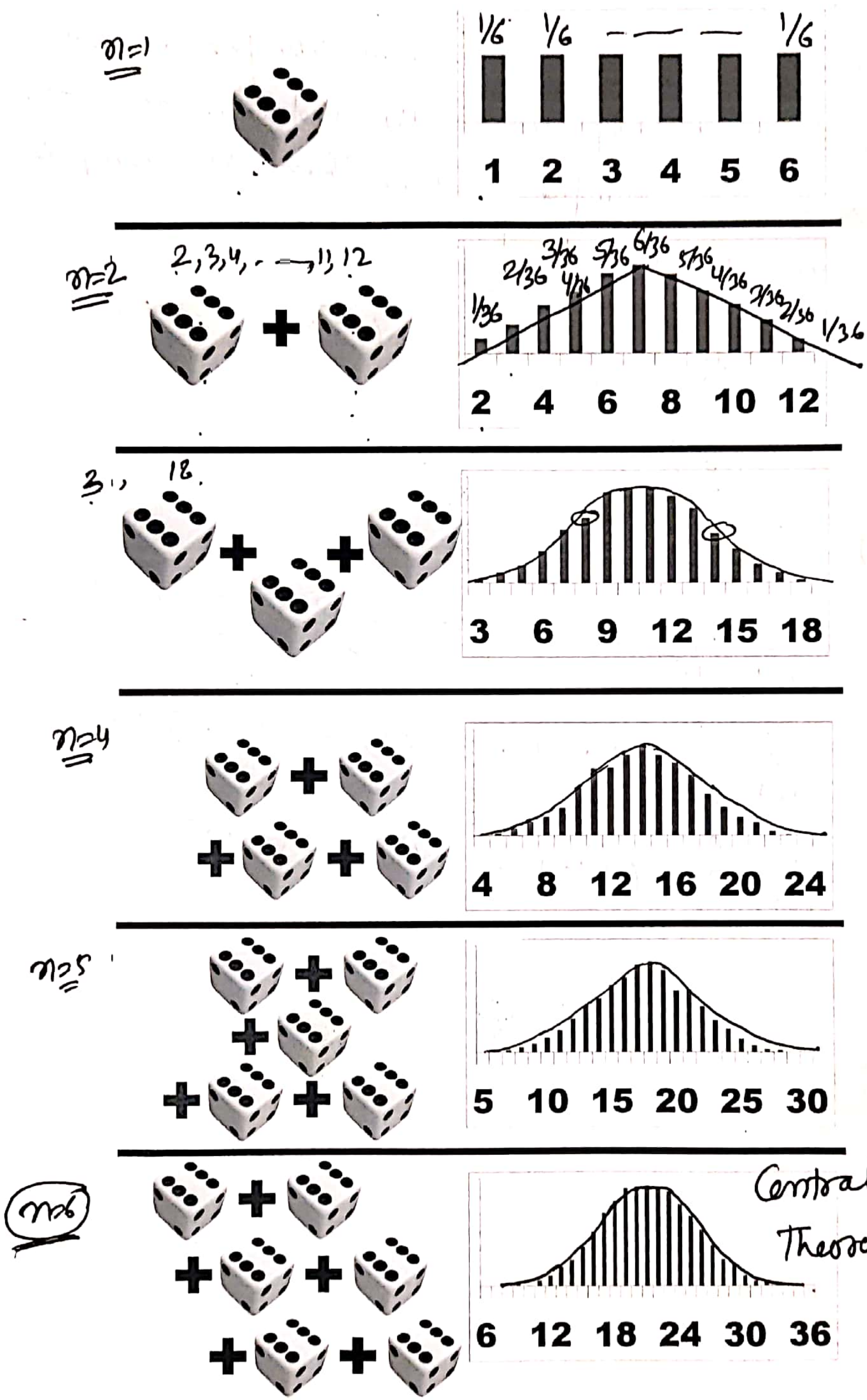
$$S^2$$

$$E(S^2) = E\left[\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2\right]$$

$\Rightarrow$ $E(S^2) = \sigma^2$ (HW)

* $Var(S^2) = ??$ (HW)

Rolling of dice: Sum of digits on upfront face



Central Limit Theorem

Central Limit Theorem :

Let $x_1, x_2, - x_n$ be a sequence of independent and identical distributed RV's (~~iid~~ iid RV's) each having mean (μ) and variance (σ^2). Then for large n, the distribution of

$$\bar{X} = \sum X_i \sim N(n\mu, n\sigma^2)$$

$$\Rightarrow Z = \frac{\bar{X} - n\mu}{\sigma\sqrt{n}} \sim N(0,1)$$

Q An IC has 25000 policy holders. If the yearly claim of a policy holder is a RV with mean 320 and std 540. Find the probability that the total yearly claim exceeds 8.3 million rupees.

Assume X_i : RV of yearly claim of i^{th} P.H $\quad E(X_i) = 320$

$Var(X_i) = 540^2$

X : RV of yearly claim

$X = \sum_{i=1}^{n} X_i \quad$ Using CLT $\quad X \sim N(n \times 320, n \times 540^2)$

$\sim N(25000 \times 320, 25000 \times 540^2)$

$\sim N(\underset{\mu}{\underline{8 \times 10^6}}, \underset{\sigma^2}{\underline{(8.538 \times 10^4)^2}})$

$P(X > 8.3 \times 10^6)$

$= P\left(\frac{X-\mu}{\sigma} > \frac{8.3 \times 10^6 - \mu}{\sigma}\right)$

$\approx P(Z > 3.51)$

$= 0.00023$