

SUMMARY: Sampling

1 Concept of Population and Sample

1. A **population** is the complete set of items being investigated.
2. A **sample** is a subset of a population.
3. A **random sample of size n** is a sample which is taken in such a way that every sample of size n of the population has an equal chance (equal probability) of being chosen. This also means that for a random sample of size n , every member of the population has an equal chance of being chosen. Otherwise, the sample is a **non-random sample**.

2 Estimation

	Actual population mean: μ	Actual population variance: σ^2
	Sample mean: $\bar{x}$	Sample variance (no notation)
	Unbiased estimate for population mean: $\bar{x}$	Unbiased estimate for population variance: s^2
		$s^2 = \frac{n}{n-1} \times \text{sample variance}$
If Σx and Σx^2 are given:	$\bar{x} = \frac{\Sigma x}{n}$	$s^2 = \frac{1}{n-1} \left[\Sigma x^2 - \frac{(\Sigma x)^2}{n} \right]$ (in MF27)
If $\Sigma(x-c)$ and $\Sigma(x-c)^2$ are given:	$\bar{x} = \frac{\Sigma(x-c)}{n} + c$	$s^2 = \frac{1}{n-1} \left[\Sigma(x-c)^2 - \frac{(\Sigma(x-c))^2}{n} \right]$
If $\Sigma(x-\bar{x})^2$ is given:		$s^2 = \frac{n}{n-1} \left(\frac{\Sigma(x-\bar{x})^2}{n} \right)$ (in MF27)

3 Sampling Distribution

	Normal Population	Non-Normal/Unknown distribution Population
Known population variance σ^2	$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ and $X_1 + X_2 + \dots + X_n \sim N(n\mu, n\sigma^2)$ for any sample size n	$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \text{ approx.}$ and $X_1 + X_2 + \dots + X_n \sim N(n\mu, n\sigma^2)$ approx. by Central Limit Theorem where $n \geq 30$
Unknown population variance	$\bar{X} \sim N\left(\mu, \frac{s^2}{n}\right) \text{ approx.}$ where n is large* <i>*In our syllabus, if the population variance is unknown, the sample size n will be large.</i>	$\bar{X} \sim N\left(\mu, \frac{s^2}{n}\right) \text{ approx.}$ by Central Limit Theorem where $n \geq 30$