Chapter 9
Estimation

9.1 & 9.2 Introduction & Statistical Inference

Objectives:

1. **To introduce the statistical estimation of population parameters.**
2. **To introduce the two methods of statistical inference.**
3. **To define the two areas of statistical inference.**

Definition: The **statistical inference** consists of those methods of making inferences or generalizations about a population.

Note: There are two types of statistical inference:

1. **Classical Method**: depends on the information obtained from a sample drawn from the population.
2. **Bayesian Method**: utilizes prior subjective knowledge.

Note: Statistical inference can be divided into two areas:

1. **Estimation**: Approximation of population parameters.
2. **Testing hypotheses**: formulated on estimates of parameters.

9.3 Classical Methods of Estimation

Objectives:

1. **To define the unbiased estimator.**
2. **To define the most efficient estimator.**

Definition: The **estimator** $\hat{\Theta}$ is all the possible values $\hat{\theta}$ of a statistic $\hat{\Theta}$.

Definition: The **point estimate** of θ: is a single value $\hat{\theta}$ of a statistic $\hat{\Theta}$.

Ex.1: The value $\bar{x}$ of the statistic $\overline{X}$ is a point estimate of μ, and the value s^2 of the statistic S^2 is a point estimate of σ^2.

Definition: A statistic $\hat{\Theta}$ is said to be an **unbiased estimator** of θ iff $\mu_{\hat{\theta}} = E(\hat{\theta}) = \theta$.

Ex.2: Show that $\overline{X}$ and S^2 are unbiased estimators for μ and σ^2 respectively.

Definition: Let $\hat{\theta}_1$ and $\hat{\theta}_2$ are two unbiased estimators of θ. If $\sigma^2_{\hat{\theta}_1} < \sigma^2_{\hat{\theta}_2}$, then $\hat{\theta}_1$ is **more efficient** estimator of θ than $\hat{\theta}_2$.

Note: An unbiased estimator of θ which has the smallest variance is called the Most Efficient Estimator of θ.

Definition: (**Interval Estimation**) It is a method of estimating the unknown parameter using intervals. For θ, we want to find limits of an interval, $\hat{\theta}_L$ and $\hat{\theta}_U$, such that $P(\hat{\theta}_L < \theta < \hat{\theta}_U) = 1 - \alpha$.

Note:

1. The interval $\hat{\theta}_L < \theta < \hat{\theta}_U$ computed from the sample is called a $(1 - \alpha)$ 100% confidence interval (**C.I.**). for θ.
2. $1 - \alpha$ is called the confidence coefficient or degree of confidence.

3. $\hat{\theta}_L$ and $\hat{\theta}_U$ are called confidence limits.

9.4 Estimating the Mean

Objectives:

1. **To calculate a C.I. for μ for both large and small n.**
2. **To find the error in estimation and required sample size for a given confidence level.**

Cases of finding C.I.'s for μ

*Case.*1: For small n, normal population, and known σ: $\bar{X} : N(\mu, \frac{\sigma^2}{n})$ →

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha = P(-z_{\alpha/2} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2})$$ →

$$P(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$ → A (1 – α) 100% C.I. for μ is given by $[\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$.

Notes:

1. The lower and upper limits of the C.I. are denoted by: $\bar{X}_L = \bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ and $\bar{X}_U = \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ respectively.
2. If the sample is **large** then by CLT, since $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$, a (1 – α) 100% C.I. for μ is given by $[\bar{X} \pm z_{\alpha/2}\frac{\sigma(s)}{\sqrt{n}}]$, i.e. if σ is unknown, it is simply replaced by s.
3. The C.I. is interpreted as; if we have a very large number of samples, each of size n, then (1 – α) 100% of all the C.I. based on these samples will contain the actual value of the unknown μ.

Ex.1 (9.2/236): $n = 36$, $\bar{x} = 2.6$, $\sigma = 0.3$. Find a 95% & a 99% C.I.'s for μ.

Theorem: If $\bar{X}$ is used as an estimate of μ, then the estimation **error** (e) ≤ $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ with a confidence level of (1 – α) 100%.

Ex.2 (9.2/236): Find the maximum error in estimating μ if the significance level is 5% and 1%.

Theorem: If $\bar{X}$ is used as an estimate of μ, then the required sample size n, given an estimation error ≤ e and with a confidence level of (1 – α) 100%, is given by $n \geq \left(\frac{z_{\alpha/2}\sigma}{e}\right)^2$.

Ex.3 (9.3/238): Find the minimum n for Ex.2 if the required confidence level is 95% and the maximum estimation error is 0.05.

*Case.*2: For small n, normal population, and unknown σ: $\frac{\overline{X}-\mu}{s/\sqrt{n}}$:t_{n-1} →

$$P(-t_{\alpha/2} < T < t_{\alpha/2}) = 1-\alpha = P(-t_{\alpha/2} < \frac{\overline{X}-\mu}{s/\sqrt{n}} < t_{\alpha/2}) \quad \rightarrow$$

$$P(\overline{X} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \overline{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}) = 1-\alpha$$ → A (1 – α) 100% C.I.

for μ is given by $[\overline{X} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}]$.

Ex.4 (9.4/239): n = 7, the sample is 9.8, 10.2, 10.4, 9.8, 10.0, 10.2, and 9.6. Assume normality and find a 95% C.I. for μ.

Ex.5: Problem (7/245). n = 100, $\bar{x}$ = 23,500, s = 3,900, and σ = ?. Find a 99% C.I. for μ and the maximum estimation error.

Ex.6: Problem (14/246). n = 10, $\bar{x}$ = 230, s = 15, σ = ?, and assume normality. Find a 99% C.I. for μ.

9.5 Standard Error of a Point Estimate
Objective: To define the standard error of a point estimate.

Definition: The **standard error** of a point estimate is simply its standard deviation and is denoted by SE($\hat{\theta}$).

Ex.1: $\overline{X}$ is a point estimate of μ, then the standard error of $\overline{X}$ is given by:

SE($\overline{X}$) = $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$, but SE($\overline{X}$) = $\sigma_{\overline{X}} = \frac{s}{\sqrt{n}}$ if σ is unknown.

Definition: When n < 30, the population is normal, and σ is unknown, then A (1 – α) 100% C.I. for μ is given by $[\overline{X} \pm t_{\alpha/2,n-1}SE(\overline{X})]$.

9.6 Standard Error of a Point Estimate
Objectives:
1. **To define the prediction interval.**
2. **To calculate a prediction interval for a future value.**
3. **To apply P.I. for detecting outliers.**

Definition: A **prediction interval**, abbreviated by **P.I.**, is simply a confidence interval for a specific individual value of a future observation.

Note:

1. For small n, normal population, and known σ:

Since $\overline{X}$: N(μ , $\frac{\sigma^2}{n}$), then for a new **random** observation X_o, independent from $\overline{X}$, the new r.v. $X_o - \overline{X}$: N(0 , $\sigma^2 + \frac{\sigma^2}{n}$). We are looking for two numbers $-z_{\alpha/2}$ and $z_{\alpha/2}$ such that:

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1-\alpha = P(-z_{\alpha/2} < \frac{X_o - \overline{X}}{\sqrt{\sigma^2 + \frac{\sigma^2}{n}}} < z_{\alpha/2}) \quad \rightarrow$$

$P(\overline{X} - z_{\alpha/2}\sigma\sqrt{1+\frac{1}{n}} < X_o < \overline{X} + z_{\alpha/2}\sigma\sqrt{1+\frac{1}{n}}) = 1-\alpha$ → A (1 – α) 100% P.I. for x_o is given by $[\overline{X} \pm z_{\alpha/2}\sigma\sqrt{1+\frac{1}{n}}]$.

2. For large n. Since $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$ and $X_o - \overline{X} \sim N(0, \sigma^2 + \frac{\sigma^2}{n})$ → A (1 – α) 100% P.I. for x_o is given by $[\overline{X} \pm z_{\alpha/2}\sigma(s)\sqrt{1+\frac{1}{n}}]$.

3. For small n, normal population, and **unknown** σ: Since $\frac{X_o - \overline{X}}{s\sqrt{1+\frac{1}{n}}} : t_{n-1}$ → A (1 – α) 100% P.I. for x_o is given by $[\overline{X} \pm t_{\alpha/2,n-1} s\sqrt{1+\frac{1}{n}}]$.

Ex.1 (9.6/242): $n = 30$, $\bar{x} = 96.2$, $s = 0.8$, $\sigma = ?$, and assume normality. Find a 99% P.I. for a new pack.

Note: A nice application of the prediction interval for a single future observation is the outlier detection, following the **rule** that an **observation is an outlier if it falls outside the prediction interval constructed without including the observation of interest in the sample**.

9.10 Estimating the Proportion

Objectives:

1. **To define the estimate of the binomial parameter P.**
2. **To compute a C.I. for P.**
3. **To find the error in estimation and required sample size for a given confidence level.**

Definition: A point estimate of the **proportion P** in a binomial experiment is given by: $\hat{P} = \frac{X}{n}$, X: number of successes in n trials, where the r.v. X:B(n,p). The sample proportion $\hat{p} = \frac{x}{n}$, x: number of successes in a sample of size n.

Definition: For sufficiently **large** n (as $n \to \infty$), $\hat{P} = \frac{X}{n}$, by the CLT has approximately the normal distribution with a mean $\mu_{\hat{P}} = E(\hat{P}) = E(\frac{X}{n}) = \frac{np}{n} = p$ and a variance $\sigma^2_{\hat{P}} = V(\hat{P}) = V(\frac{X}{n}) = \frac{V(X)}{n^2} = \frac{\sigma_X^2}{n^2} = \frac{npq}{n^2} = \frac{pq}{n}$.

Definition: A (1 – α) 100% C.I. for P is given by $[\hat{p} \pm z_{\alpha/2} SE(\hat{p})]$ = $[\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}]$, given that $n \geq 30$ (large), or $np \geq 5$ & $nq \geq 5$.

Theorem: If $\hat{P}$ is used as an estimate of P, then the estimation **error** (e) $\leq$ $z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$ with a confidence level of $(1 - \alpha)$ 100%.

Theorem: If $\hat{P}$ is used as an estimate of P, then the required sample size n, given an estimation error $\leq e$ and with a confidence level of $(1 - \alpha)$ 100%, is given by $n \geq \frac{z^2_{\alpha/2}\hat{p}\hat{q}}{e^2}$.

Ex.1 (9.12/258): If n = 500, x = 340, then find a 95% C.I. for P, the **maximum** estimation error with a confidence level 95%, and the required sample size needed for estimating P with an error of 0.02 and a 95% confidence level.

Theorem: If $\hat{P}$ is used as an estimate of P, then the **maximum** required sample size n, given an estimation error $\leq e$ and with a confidence level of $(1 - \alpha)$ 100%, is given by $n = \frac{z^2_{\alpha/2}}{4e^2}$.

Ex.2 (9.14/260): If n = 500, x = 340, then find a 95% C.I. for P, the **maximum** estimation error with a confidence level 95%, and the required sample size needed for estimating P with an error of 0.02 and a 95% confidence level.

Ex.3: Problem (4/263). If n = 100, x = 8, then find a 98% C.I. for P.

Ex.4: Problem (11/263). If n = 100, x = 8, then find the required sample size needed for estimating P with an error of 0.05 and a 98% confidence level.

Ex.5: Problem (12/263). Find the required sample size needed for estimating P with an error of 1% and at least 95% confidence level.

9.12 Estimating the Variance

Objectives:

1. **To define the point estimate of the variance σ^2.**
2. **To compute a C.I. for σ^2.**
3. **To find the error in estimation and required sample size for a given confidence level.**

Definition: A point estimate of the **variance σ^2** is the sample variance S^2, where the statistic $\frac{(n-1)S^2}{\sigma^2} : \chi^2_{n-1}$, given that the sample of size n is drawn from a normal population.

Definition: If S^2 is the variance of a sample of size n drawn from a normal population, then a $(1 - \alpha)$ 100% C.I. for σ^2 is given by $\frac{(n-1)S^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}$ with n-1 degrees of freedom.

Ex.1 (9.16/265): If the sample is 46.4, 46.1, 45.8, 47.0, 46.1, 45.9, 45.8, 46.9, 45.2, and 46.0, then find a 95% C.I. for σ^2. Write the necessary assumption(s).

Ex.2: Problem (2/268). If n = 20, $\bar{x}$ = 72, s = 4, then find a 98% C.I. for σ^2, assuming normality.