Chapter 7. Sampling Distributions and Point Estimation of Parameters

If $X_1, \cdots, X_n$ is a random sample from a population with normal distribution $N(\mu, \sigma^2)$, then the sample mean $\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$ has a normal distribution $N(\mu, \frac{\sigma^2}{n})$.

Central Limit Theorem:
If $X_1, \cdots, X_n$ is a random sample of size $n$ taken from a population (either finite or infinite) with mean $\mu$ and finite variance $\sigma^2$, and if $\bar{X}$ is the sample mean, the limiting distribution of
$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$
is the standard normal distribution as $n \to \infty$.

Unbiased Estimators:
The point estimator $\hat{\theta}$ is an unbiased estimator for the parameter $\theta$ if $E(\hat{\theta}) = \theta$. If the estimator is not unbiased, then the difference $E(\hat{\theta}) - \theta$ is called the bias of the estimator $\hat{\theta}$.

The mean squared error of an estimator $\hat{\theta}$ of the parameter $\theta$ is defined as
$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = V(\hat{\theta}) + (bias)^2,$$ where $V(\hat{\theta})$ is the variance of $\hat{\theta}$. For two estimator $\hat{\Theta}_1, \hat{\Theta}_2$, the relative efficiency of $\hat{\Theta}_2$ to $\hat{\Theta}_1$ is defined as
$$\frac{MSE(\hat{\Theta}_1)}{MSE(\hat{\Theta}_2)}.$$ If this relative efficiency is less than 1, then $\hat{\Theta}_1$ is a more efficient (better) estimator of $\theta$ than $\hat{\Theta}_2$.

Methods of Moments:
Let $X_1, X_2, \cdots, X_n$ be a random sample from the probability distribution $f(x)$ (pdf or pmf). The $k$th population moment is $E(X^k), k = 1, 2, \ldots$. The corresponding $k$th sample moment is
$$\frac{1}{n} \sum_{i=1}^{n} X_i^k, \quad k = 1, 2, \cdots.$$ Let the population distribution have $m$ unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$. The moment estimators $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m$ are found by equating the first $m$ population moments to the first $m$ sample moments and solving the resulting equations for the unknown parameters.

Maximum Likelihood Estimator:
Suppose that $X$ is a random variable with probability distribution $f(x; \theta)$, where $\theta$ is a single unknown parameter. Let $x_1, x_2, \cdots, x_n$ be the observed values in a random sample of size $n$. Then the likelihood function of the sample is
$$L(\theta) = f(x_1; \theta) \cdot f(x_2; \theta) \cdots f(x_n; \theta)$$
The maximum likelihood estimator (MLE) of $\theta$ is the value of $\theta$ that maximizes the likelihood function $L(\theta)$. Basically, we solve the equation
$$\frac{d \ln L(\theta)}{d\theta} = 0.$$
and its solution gives the MLE of $\theta$.

Chapter 8. Statistical Intervals for a Single Sample

Confidence interval for $\mu$ when $\sigma$ is known:
If $\bar{x}$ is the sample mean of a random sample of size $n$ from a normal population with known variance $\sigma^2$, a $100(1 - \alpha)\%$ CI on $\mu$ is given by

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}},$$

where $z_{\alpha/2}$ is the upper $100\alpha/2$ percentage point of the standard normal distribution.

If $\bar{x}$ is used as an estimate of $\mu$, we can be $100(1 - \alpha)\%$ confident that the error $|\bar{x} - \mu|$ will not exceed a specified amount $E$ when the sample size is

$$n = \left( \frac{z_{\alpha/2}\sigma}{E} \right)^2$$

A $100(1 - \alpha)\%$ upper-confidence bound for $\mu$ is

$$\mu \leq u = \bar{x} + z_{\alpha} \frac{\sigma}{\sqrt{n}}$$

and a $100(1 - \alpha)\%$ lower-confidence bound for $\mu$ is

$$\bar{x} - z_{\alpha} \frac{\sigma}{\sqrt{n}} = l \leq \mu$$

Confidence interval on $\mu$ when $\sigma^2$ is unknown:
Let $X_1, \ldots, X_n$ be a random sample from a normal distribution $N(\mu, \sigma^2)$ with unknown $\mu$ and $\sigma^2$. The random variable

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

has a $t$ distribution with $n - 1$ degrees of freedom.

A $100(1 - \alpha)\%$ CI on $\mu$ (with unknown $\sigma^2$) is given by

$$\bar{x} - t_{\alpha/2,n-1} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2,n-1} \frac{s}{\sqrt{n}},$$

where $t_{\alpha/2,n-1}$ is the upper $100\alpha/2$ percentage point of the $t$ distribution with $n - 1$ degrees of freedom.

A $100(1 - \alpha)\%$ upper-confidence bound for $\mu$ is

$$\mu \leq u = \bar{x} + t_{\alpha,n-1} \frac{s}{\sqrt{n}}$$

and a $100(1 - \alpha)\%$ lower-confidence bound for $\mu$ is

$$\bar{x} - t_{\alpha,n-1} \frac{s}{\sqrt{n}} = l \leq \mu$$
CI for population proportion \( p \):
If \( \hat{p} \) is the proportion of observations in a random sample of size \( n \) that belongs to a class of interest, an approximate 100\((1 - \alpha)\)% confidence interval on \( p \) is
\[
\hat{p} - z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \leq p \leq \hat{p} + z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}
\]
where \( z_{\alpha/2} \) is the upper \( \alpha/2 \) percentage point of the standard normal distribution.

**The Choice of Sample Size:**
\[
n = \left( \frac{z_{\alpha/2}}{E} \right)^2 \hat{p}(1 - \hat{p})
\]
The approximate 100\((1 - \alpha)\)% lower and upper confidence bounds are
\[
\hat{p} - z_{\alpha} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \leq p \quad \text{and} \quad p \leq \hat{p} + z_{\alpha} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}
\]

**Prediction and Tolerance Interval:**
A 100\((1 - \alpha)\)% prediction interval on a single future observation from a normal distribution is given by
\[
\bar{x} - t_{\alpha/2, n-1} s \sqrt{\frac{1 + \frac{1}{n}}{n}} \leq x_{n+1} \leq \bar{x} + t_{\alpha/2, n-1} s \sqrt{\frac{1 + \frac{1}{n}}{n}}
\]
A tolerance interval for capturing at least 100\(\gamma\)% of the values in a normal distribution with confidence level 100\((1 - \alpha)\)% is
\[
\bar{x} - ks, \bar{x} + ks
\]
where \( k \) is a tolerance interval factor found in Appendix Table XII. Values are given for \( \gamma = 90\%, 95\% \), and 99% and for 90%, 95% and 99% confidence.

**Chapter 9. Tests of Hypotheses for a Single Sample**
\( \alpha = P(\text{type I error}) = P(\text{reject } H_0 \text{ when } H_0 \text{ is true}). \)
\( \beta = P(\text{type II error}) = P(\text{fail to reject } H_0 \text{ when } H_0 \text{ is false}). \)

Tests on the mean with known variance:
Null Hypothesis : \( H_0 : \mu = \mu_0 \)
Test Statistics:
\[
Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}
\]

<table>
<thead>
<tr>
<th>Alternative hypothesis</th>
<th>Rejection region</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_1 : \mu \neq \mu_0 )</td>
<td>( z_0 &gt; z_{\alpha/2} ) or ( z_0 &lt; -z_{\alpha/2} )</td>
</tr>
<tr>
<td>( H_1 : \mu &gt; \mu_0 )</td>
<td>( z_0 &gt; z_{\alpha} )</td>
</tr>
<tr>
<td>( H_1 : \mu &lt; \mu_0 )</td>
<td>( z_0 &lt; -z_{\alpha} )</td>
</tr>
</tbody>
</table>

**P-value Method:**
\[
P = \begin{cases} 
2[1 - \Phi(|z_0|)], & \text{for a two-tailed test : } H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu \neq \mu_0 \\
1 - \Phi(z_0), & \text{for a upper-tailed test : } H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu > \mu_0 \\
\Phi(z_0), & \text{for a lower-tailed test : } H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu < \mu_0 
\end{cases}
\]
If \( P \leq \alpha \), then \( H_0 \) would be rejected. If \( P > \alpha \), we will fail to reject \( H_0 \).