CHAPTER 9

Hypothesis Testing

Similar to a courtroom trial, the jury needs to decide between one of two possibilities:

- The person is guilty.
- The person is innocent.

To begin with, the person is assumed innocent. The prosecutor presents evidence, trying to convince the jury to reject the original assumption of innocence, and conclude that the person is guilty.

Parts of a Statistical Test

- The null hypothesis, H_0
- The alternative hypothesis, H_a
- The test statistic and its p-value
- The rejection region
- The conclusion

The two competing hypotheses are the alternative hypothesis H_a, generally the hypothesis that the researcher wishes to support, and the null hypothesis H_0, a contradiction of the alternative hypothesis.
The researcher uses the sample data to

- Reject H_0 and conclude that H_a is true.
- Accept (do not reject) H_0 as true.

Test statistic: A single number calculated from the sample data.
p**-value**: A probability calculated using the test statistic.
Rejection region: One set, consisting of values that support the alternative hypothesis and lead to rejecting H_0.
Accepting region: One set, consisting of values that support the null hypothesis.

Critical values: The value that separate the acceptance and rejection regions.
A **Type I error** for a statistical test is the error of rejecting the null hypothesis when it is true.
A **level of significance** (significance level) α: for a statistical test of hypothesis is

$$\alpha = P(\text{Type I error}) = P(\text{falsely rejecting } H_0) = P(\text{rejecting } H_0 \text{ when it is true})$$

A **Type II error** for a statistical test is the error of accepting the null hypothesis when it is false. And

$$\beta = P(\text{Type II error}) = P(\text{falsely accepting } H_0) = P(\text{accepting } H_0 \text{ when it is fale})$$

The **power** of a statistical test, given as

$$1 - \beta = P(\text{reject } H_0 \text{ when } H_a \text{ is true})$$

measures the ability of the test to perform as required.

p**-value**: The p-value or observed significant level of a statistical test is the smallest value of α for which H_0 can be rejected. It is the actual risk of committing a Type I error, if H_0 is rejected based on the observed value of the test statistic. The p-value measures the strength of the evidence against H_0. If the p-value is less than or equal to a preassigned significance level α, then the null hypothesis can be rejected, and you can report that the results are statistically significant at level α.

Large-Sample Statistical Test for the population mean μ

When n observations in the sample are randomly selected from the population with mean μ and n is large ($n \geq 30$), then we conduct the following hypothesis testing about μ based on the given large sample.

1. Null hypothesis: $H_0 : \mu = \mu_0$

2. Alternative hypothesis
One-Tailed: $H_a : \mu > \mu_0$ (or; $H_a : \mu < \mu 0$)
Two-Tailed Test: $H_a : \mu \neq \mu_0$

3. The test statistic is

$$z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

4. p-value
For the One-Tailed: $H_a : \mu > \mu_0$, the p-value is defined as p-value$=P(Z > z)$
For the One-Tailed: $H_a : \mu < \mu_0$, the p-value is defined as p-value$=P(Z < z)$
For the Two-Tailed Test: $H_a : \mu \neq \mu_0$, the p-value is defined as

$$p - \text{value} = P(Z < -|z|) + P(Z > |z|) = 2P(Z < -|z|) = 2P(Z > |z|).$$

5. Conclusions at level α:

One-Tailed: (critical value approach):
Reject H_0 in favor of $H_a : \mu > \mu_0$ if $z > z_\alpha$
Reject H_0 in favor of $H_a : \mu < \mu_0$ if $z < -z_\alpha$

One-Tailed: (p-value approach):
Reject H_0 in favor of $H_a : \mu > \mu_0$ if p-value$\leq \alpha$
Reject H_0 in favor of $H_a : \mu < \mu_0$ if p-value$\leq \alpha$

Two-Tailed: (critical value approach):
Reject H_0 in favor of $H_a : \mu \neq \mu_0$ if $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$.

Two-Tailed: (p-value approach):
Reject H_0 in favor of $H_a : \mu \neq \mu_0$ if p-value$\leq \alpha$.

Large-Sample Statistical Test for the population proportion p
When a sample of n is obtained from a population with proportion p_0 in such a way that $np_0 \geq 5$ and $nq_0 \geq 5$ then we can conduct the following hypothesis testing problems based on this sample.

1. Null hypothesis: $H_0 : p = p_0$

2. Alternative hypothesis
One-Tailed: $H_a : p > p_0$ (or; $H_a : p < p_0$)
Two-Tailed Test: $H_a : p \neq p_0$

3. The test statistic is

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}}$$

4. p-value
For the One-Tailed: $H_a : p > p_0$, the p-value is defined as p-value$=P(Z > z)$
For the One-Tailed: $H_a : p < p_0$, the p-value is defined as p-value$=P(Z < z)$
For the Two-Tailed Test: $H_a : p \neq p_0$, the p-value is defined as

$$p - \text{value} = P(Z < -|z|) + P(Z > |z|) = 2P(Z < -|z|) = 2P(Z > |z|).$$

5. Conclusions at level α:

One-Tailed: (critical value approach):
Reject H_0 in favor of $H_a : p > p_0$ if $z > z_\alpha$
Reject H_0 in favor of $H_a : p < p_0$ if $z < -z_\alpha$

One-Tailed: (p-value approach):
Reject H_0 in favor of $H_a : p > p_0$ if p-value$\leq \alpha$
Reject H_0 in favor of $H_a : p < p_0$ if p-value$\leq \alpha$

Two-Tailed: (critical value approach):
Reject H_0 in favor of $H_a : p \neq p_0$ if $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$.

Two-Tailed: (p-value approach):
Reject H_0 in favor of $H_a : p \neq p_0$ if p-value$\leq \alpha$.

Examples: 1. Suppose a scheduled flight must average at least 60% occupancy in order to be profitable, and an examination of the occupancy rate for 120 flights from Atlanta to Dallas showed a mean occupancy per flight of 58% and a standard deviation of 11
a. If μ is the mean occupancy per flight and if the company wishes to determine whether or not this scheduled flight is unprofitable, give the alternative and the null hypotheses for the test.
b. Does the alternative hypothesis in part a imply a one or two-tailed test?
c. Do the occupancy data for the 120 flights suggest that this scheduled flight is unprofitable?

2. A random sample of 120 observations was selected from a binomial population, and 72 successes were observed. Do the data provide sufficient evidence to indicate that p is greater than 0.5?

Large-Sample Statistical Test for $(\mu_1 - \mu_2)$

When two independent samples of n_1 and n_2 observations are obtained from two populations with means μ_1 and μ_2 in such a way that $n_1 \geq 30$ and $n_g eq30$ then we can conduct the following hypothesis testing problems about $\mu_1 - \mu_2$.

1. Null hypothesis: $H_0 : \mu_1 - \mu_2 = D_0$

2. Alternative hypothesis
One-Tailed: $H_a : \mu_1 - \mu_2 > D_0$ (or; $H_a : \mu_1 - \mu_2 < D_0$)
Two-Tailed Test: $H_a : \mu_1 - \mu_2 \neq D_0$

3. The test statistic is

$$z = \frac{\bar{X}_1 - \bar{X}_2 - D_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

4. p-value
For the One-Tailed: $H_a : \mu_1 - \mu_2 > D_0$, the p-value is defined as p-value$=P(Z > z)$
For the One-Tailed: $H_a : \mu_1 - \mu_2 < D_0$, the p-value is defined as p-value$=P(Z < z)$
For the Two-Tailed Test: $H_a : \mu_1 - \mu_2 \neq D_0$, the p-value is defined as

$$p - \text{value} = P(Z < -|z|) + P(Z > |z|) = 2P(Z < -|z|) = 2P(Z > |z|).$$

5. Conclusions at level α:

One-Tailed: (critical value approach):
Reject H_0 in favor of $H_a : \mu_1 - \mu_2 > D_0$ if $z > z_\alpha$
Reject H_0 in favor of $H_a : \mu_1 - \mu_2 < D_0$ if $z < -z_\alpha$

One-Tailed: (p-value approach):
Reject H_0 in favor of $H_a : \mu_1 - \mu_2 > D_0$ if p-value$\leq \alpha$

Reject H_0 in favor of $H_a : \mu_1 - \mu_2 < D_0$ if p-value$\leq \alpha$

Two-Tailed: (critical value approach):
Reject H_0 in favor of $H_a : \mu_1 - \mu_2 \neq D_0$ if $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$.

Two-Tailed: (p-value approach):
Reject H_0 in favor of $H_a : \mu_1 - \mu_2 \neq D_0$ if p-value$\leq \alpha$.

Large-Sample Statistical Test for $(p_1 - p_2)$

When two independent samples of n_1 and n_2 observations are obtained from two populations with proportions p_1 and p_2 in such a way that $n_1\hat{p}_1 \geq 5$, $n_1\hat{q}_1 \geq 5$, $n_2\hat{p}_2 \geq 5$ and $n_2\hat{p}_2 \geq 5$, then we can conduct the following hypothesis testing problems about $p_1 - p_2$.

1. Null hypothesis: $H_0 : p_1 - p_2 = 0$

2. Alternative hypothesis
One-Tailed: $H_a : p_1 - p_2 > 0$ (or; $H_a : p_1 - p_2 < 0$)
Two-Tailed Test: $H_a : p_1 - p_2 \neq 0$

3. The test statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}}$$

where $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$.

4. p-value
For the One-Tailed: $H_a : p_1 - p_2 > 0$, the p-value is defined as p-value$=P(Z > z)$
For the One-Tailed: $H_a : p_1 - p_2 < 0$, the p-value is defined as p-value$=P(Z < z)$
For the Two-Tailed Test: $H_a : p_1 - p_2 \neq 0$, the p-value is defined as

$$p - \text{value} = P(Z < -|z|) + P(Z > |z|) = 2P(Z < -|z|) = 2P(Z > |z|).$$

5. Conclusions at level α:

One-Tailed: (critical value approach):
Reject H_0 in favor of $H_a : p_1 - p_2 > 0$ if $z > z_\alpha$
Reject H_0 in favor of $H_a : p_1 - p_2 < 0$ if $z < -z_\alpha$

One-Tailed: (p-value approach):
Reject H_0 in favor of $H_a : p_1 - p_2 > 0$ if p-value$\leq \alpha$
Reject H_0 in favor of $H_a : p_1 - p_2 < 0$ if p-value$\leq \alpha$

Two-Tailed: (critical value approach):
Reject H_0 in favor of $H_a : p_1 - p_2 \neq 0$ if $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$.

Two-Tailed: (p-value approach):
Reject H_0 in favor of $H_a : p_1 - p_2 \neq 0$ if p-value$\leq \alpha$.

Example: Independent random samples of 280 and 350 observations were selected from binomial populations 1 and 2 respectively. Sample 1 had 132 successes, and sample 2 had 178 successes. Do the data present sufficient evidence to indicate that the proportion of successes in population 1 is smaller than the proportion in population 2?