WILEY

Reading 11: Hypothesis Testing

Learning Outcome Statements

- Covered
  - 11a, 11b, 11c, 11d, 11e, 11f, 11g, 11h, 11i, 11j, 11k, 11l

- Not Covered
  - None

Steps in Hypothesis Testing

1. State the hypotheses: null and alternative.
2. Identify the appropriate test statistic and its probability distribution.
3. Specify the level of significance.
4. State the decision rule.
5. Collect the data and calculate the test statistic.
6. Make the statistical decision regarding the hypotheses.
7. Make an economic or investment decision.

Null and Alternative Hypotheses

- Null hypothesis—H_0
  - Hypothesis to be tested
  - Considered true unless evidence suggests otherwise
  - The hypothesis we would like to reject
  - Includes the equal sign
- Alternative hypothesis—H_a
  - The complement of H_0
  - The hypothesis we would like to *accept*
  - Does not include the equal sign
- Formulations of the null and alternative hypotheses
  - H_0: $\theta = \theta_0$ versus H_a: $\theta \neq \theta_0$—a two-tailed test
  - H_0: $\theta \leq \theta_0$ versus H_a: $\theta > \theta_0$—a one-tailed test
  - H_0: $\theta \geq \theta_0$ versus H_a: $\theta < \theta_0$—a one-tailed test

Selecting the Test Statistic—I

The appropriate test statistic depends on the nature of the hypotheses being tested:

- Often the test statistic is of the form

$$\text{Test statistic} = \frac{\text{Sample statistic} - \text{Hypothesized value}}{\text{Standard error of sample statistic}}$$

- For tests concerning a population mean, the sample statistic will be the sample mean.
- For tests concerning a population variance, the sample statistic will be the sample variance.
- For tests concerning two population means, the sample statistic will be the difference of the sample means.
- For tests concerning two population variances, the sample statistic will be the ratio of the sample variances.

Selecting the Test Statistic—II

The probability distribution for the test statistic depends on the nature of the test statistic:

- For tests concerning a population mean, the sample statistic will follow a *z*-distribution or a *t*-distribution.
- For tests concerning a population variance, the sample statistic will follow a chi-square (X^2) distribution.
- For tests concerning two population means, the sample statistic will follow a *z*-distribution or a *t*-distribution.
- For tests concerning two population variances, the sample statistic will follow an *F*-distribution.

Significance Level—Type I and II Errors

Selecting the appropriate level of significance means balancing two possible errors:

	Reject H_0	Fail to Reject H_0
H_0 True	Type I Error	Correct (No Error)
H_0 False	Correct (No Error)	Type II Error

The level of significance (α) is P(Type I Error).

Generally, as P(Type I Error) decreases, P(Type II Error) increases, and vice versa. The only way to decrease both probabilities is to increase *n*: to get more data.

The **power of a test** is P(correctly rejecting a false H_0):

$$\text{power of a test} = 1 - \text{P(Type II Error)}$$

Collect the Data

- Choose an appropriate sample size
  - Required precision for statistics
  - Avoiding sampling from more than one population
  - Cost of additional data
- Avoid sample biases
  - Data-mining bias
  - Sample selection bias
  - Survivorship bias
  - Look-ahead bias
  - Time-period bias

Calculate Test Statistic – I

- Testing one mean, *z*-test:

$$z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \quad \text{or} \quad z = \frac{\bar{X} - \mu_0}{s / \sqrt{n}}$$

- Testing one mean, *t*-test:

$$t = \frac{\bar{X} - \mu_0}{s / \sqrt{n}}$$

- Testing one variance, normal population, chi-square test, $n - 1$ *df*:

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

Calculate Test Statistic—II

- Testing difference of two means, normal populations, *t*-test
  - Population variances unknown but assumed equal:

$$t = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\left(\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}\right)^{1/2}}$$

where:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

is a pooled estimator of the common variance

  - *df* is $n_1 + n_2 - 2$

Calculate Test Statistic—III

- Testing difference of two means, independent normal populations, *t*-test (cont.)
  - Population variances unknown but ***not*** assumed equal:

$$t = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^{1/2}}$$

where:

$$df = \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(s_1^2 / n_1\right)^2}{n_1} + \dfrac{\left(s_2^2 / n_2\right)^2}{n_2}}$$

Calculate Test Statistic—IV

- Testing mean difference of two dependent normal populations (**paired comparison test**), *t*-test

$$t = \frac{\bar{d} - \mu_{d0}}{s_{\bar{d}} / \sqrt{n}}$$

- Testing equality of variances of two normal populations, *F*-test

$$F = \frac{s_1^2}{s_2^2}$$

  - $df_1 = n_1 - 1$, $df_2 = n_2 - 1$
  - By convention, $s_1^2 > s_2^2$

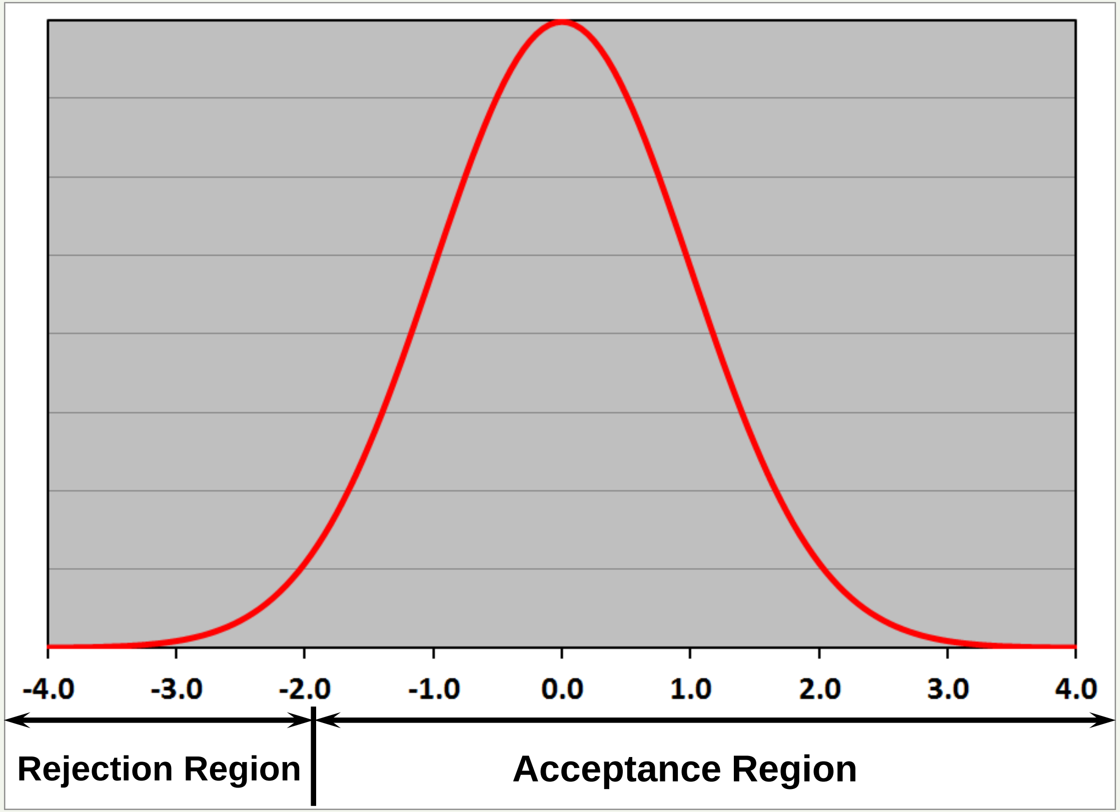
Calculate Test Statistic—V

- Testing the significance of the correlation coefficient, *t*-test

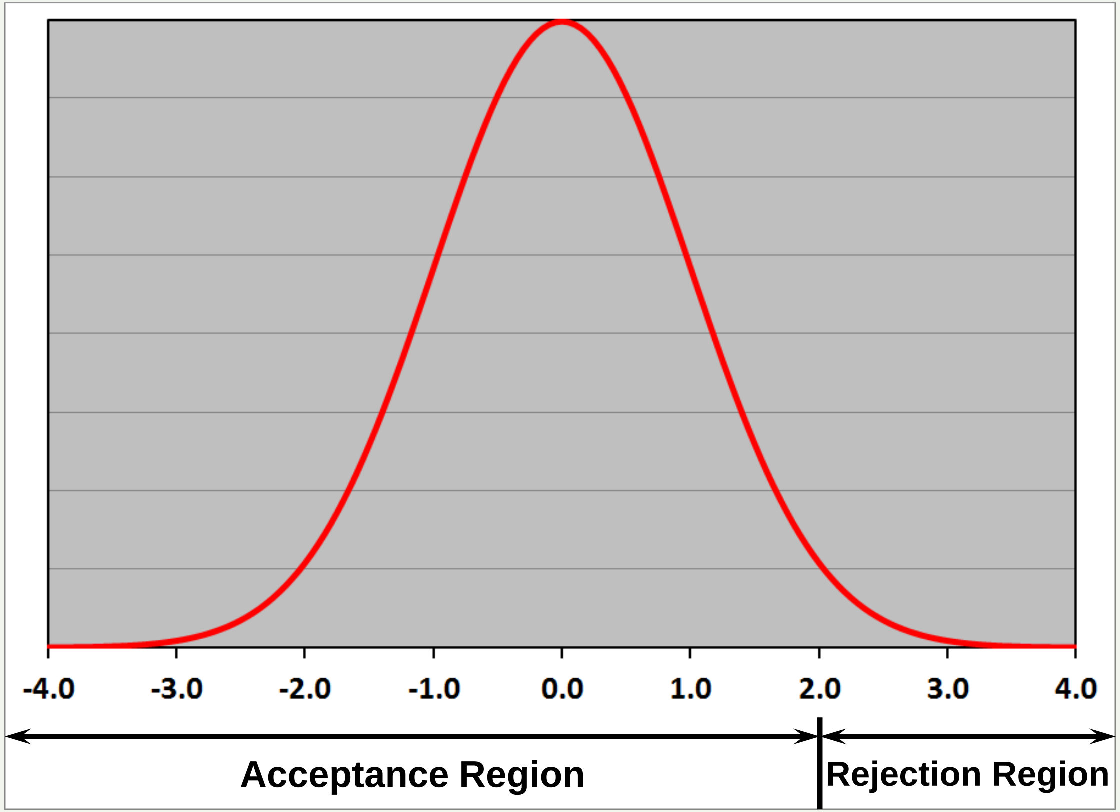
$$\text{Test-stat} = t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

  - $df = n - 2$

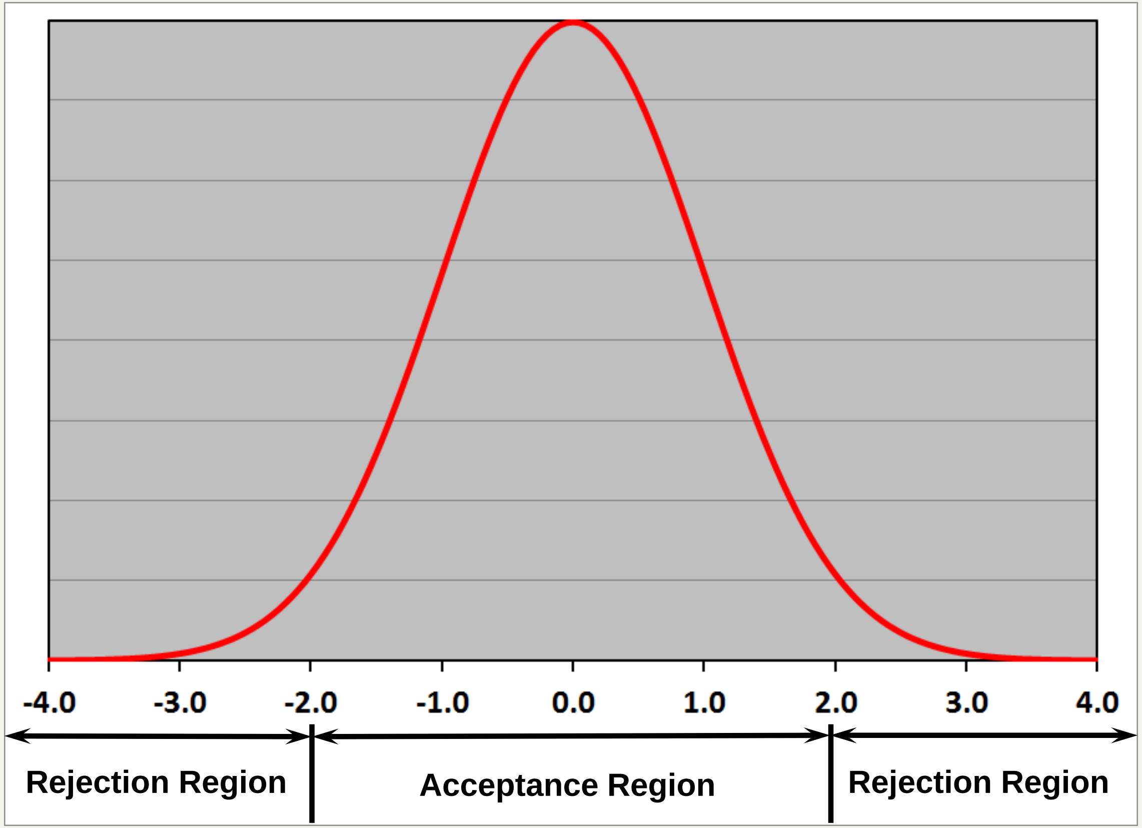
Decision Rule: One-Tail Test—I



Decision Rule: One-Tail Test—II



Decision Rule: Two-Tail Test



Practice Question

The standard deviation of returns for ABC stock for the last 61 months is 7.8%, and the standard deviation of returns for XYZ stock for the last 41 months is 10.1%. The null hypothesis that the overall standard deviations of returns are equal should:

A. Not be rejected at the 90% confidence level.
B. Be rejected at the 5% level of significance.
C. Be rejected at the 10% level of significance.

Statistical vs. Economic Significance

A hypothesis test may lead to a statistically meaningful result but not an economically meaningful result.

The statistical test may not consider:

- Transaction costs
- Taxes
- Risk
- Changing conditions that may make the strategy less useful in the future

p-Value

The ***p*-value** is the smallest level of significance at which the null hypothesis can be rejected.

- A p-value is a level of significance, just like α.
- The p-value can be thought of as the ***observed*** level of significance, while α is the *chosen* level of significance.
- As the p-value is the <u>smallest</u> level of significance at which H_0 can be rejected, compare it to α.
    - If $p \leqslant \alpha$, do not reject H_0.
    - If $p > \alpha$, reject H_0 in favor of H_a.

Practice Question

In a test to determine whether the average large-cap stock's Sharpe ratio is equal to 0.60, a sample of 40 Sharpe ratios for large-cap stocks results in a *p*-value of 7.9%. The null hypothesis should:

A. Be rejected at a 90% confidence level.

B. Be rejected at a 5% level of significance.

C. Not be rejected at a 10% level of significance.

Hypothesis Testing Example—I

A fund manager at your firm claims that a mutual fund's average monthly return is 1%. You decide to test that claim at the 95% confidence level. The information you have is:

- 60 months of returns
- Sample mean: 1.23%
- Sample standard deviation: 1.05%
- Population is nonnormal, with an unknown variance

Step 1: State the hypotheses

As the manager is at your firm, you'd prefer to show that the monthly returns are greater than 1%; thus,

$$H_0: \mu \leq 1\% \text{ versus } H_a: \mu > 1\%$$

Hypothesis Testing Example—II

Step 2: Identify the appropriate test statistic and its probability distribution

As the test concerns a single population mean and the sample size is large, the appropriate statistic is a *z*-statistic:

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

Step 3: Specify the level of significance

A confidence level of 95% has been given to us, so the level of significance is 100% – 95% = 5%.

Hypothesis Testing Example—III

Step 4: State the decision rule

The critical *z*-value is 1.645. The decision rule is that we will fail to reject H_0 if the test statistic is less than or equal to 1.645; we will reject H_0 in favor of H_a if the test statistic is greater than 1.645.

Step 5: Collect the data and calculate the test statistic

We've been given a sample size of 60, a sample mean of 1.23%, and a sample standard deviation of 1.05%, so the value of the test statistic is:

$$z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{1.23\% - 1.0\%}{1.05\%/\sqrt{60}} = 1.697$$

Hypothesis Testing Example —IV

Step 6: Make the statistical decision regarding the hypotheses

As the value of the test statistic, 1.697, is greater than the critical *z*-value, 1.645, we reject H_0 in favor of H_a and conclude that the mean monthly return on the fund is greater than 1%.

Step 7: Make an economic or investment decision

Having concluded that the mean monthly return is greater than 1%, we may decide to have our clients invest in the fund, assuming that other factors (e.g., diversification benefits) are not unfavorable.

Nonparametric Tests

A test that does not concern a population parameter or makes minimal assumptions about the population from which a sample comes is called a **nonparametric test**.

Situations for which a nonparametric test is appropriate:

- Distributional assumptions needed for parametric tests aren't satisfied (e.g., a small sample from a nonnormal population)
- Data are merely ranked; e.g., Spearman rank-order correlation
- The question doesn't address a parameter (e.g., testing whether a sample is random, testing whether a sample comes from a population with a specific probability distribution)

WILEY

Practice Questions with Solutions

Practice Question

The standard deviation of returns for ABC stock for the last 61 months is 7.8%, and the standard deviation of returns for XYZ stock for the last 41 months is 10.1%. The null hypothesis that the overall standard deviations of returns are equal should:

A. Not be rejected at the 90% confidence level.
B. Be rejected at the 5% level of significance.
C. Be rejected at the 10% level of significance.

Correct answer: C. be rejected at the 10% level of significance

The test statistic is $s^2_{XYZ}/s^2_{ABC} = (10.1\%)^2/(7.8\%)^2 = 1.6767$.

The critical *F*-value at the 90% confidence level (10% level of significance, 5% in the upper tail) with 40 *df* in the numerator and 60 *df* in the denominator is 1.59. As $1.6767 > 1.59$, reject H_0 at the 10% level of significance. The critical *F*-value at the 95% confidence level (5% level of significance, 2.5% in the upper tail) is 1.74; as $1.6767 < 1.74$, do not reject H_0 at the 5% level of significance.

Practice Question

In a test to determine whether the average large-cap stock's Sharpe ratio is equal to 0.60, a sample of 40 Sharpe ratios for large-cap stocks results in a *p*-value of 7.9%. The null hypothesis should:

A. Be rejected at a 90% confidence level.

B. Be rejected at a 5% level of significance.

C. Not be rejected at a 10% level of significance.

Correct answer: A. Be rejected at a 90% confidence level.

A 90% confidence level corresponds to $\alpha = 10\%$. As $p = 7.9\% < 10\%$, H_0 should be rejected.

At a 5% level of significance, H_0 should not be rejected, as $p = 7.9\% > 5\%$.