LM08 · QUANTITATIVE METHODS · CFA LEVEL I

Hypothesis Testing

Six-step process, Type I/II errors, power, z and t tests, p-values

What Is a Hypothesis Test?

Statistical inference = making judgments about a larger group (population) based on a smaller group (sample). Hypothesis testing checks whether a **sample statistic** is likely to come from a population with the hypothesized parameter.

i.e., Does $\bar{X} = \mu_0$?

Learning Outcome Statements

LOS a Explain hypothesis testing and its components: statistical significance, Type I and Type II errors, power of a test

LOS b Construct hypothesis tests and determine statistical significance, Type I/II errors, and power given a significance level

LOS c Compare and contrast parametric and nonparametric tests, and describe situations where each is more appropriate

LOS A, B · THE SIX STEPS

Six-Step Hypothesis Testing Process

1. **State the hypothesis** (H_0 and H_a)
2. **Identify the appropriate test statistic** and its probability distribution (z, t, χ^2, F)
3. **Specify the level of significance** (α)
4. **State the decision rule** — reject or fail to reject
5. **Collect data and calculate the test statistic**
6. **Make a decision** on the hypothesis

Key Principle

A **hypothesis** is a statement about a population parameter. The **null (H_0)** is assumed true unless we can reject it. Typically, we *want to reject H_0* to support H_a.

STEP 1 · STATE THE HYPOTHESIS

Null vs Alternative Hypothesis

The **null (H_0)** is the hypothesis to be tested; it always includes the **equality sign** (=, ≤, or ≥). The **alternative (H_a)** is what is concluded if H_0 is rejected.

TEST TYPE	H_0	H_A	REJECT H_0 WHEN
Two-tailed	$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert\text{test stat}\rvert > \lvert t_{\alpha/2}\rvert$ or $\lvert z_{\alpha/2}\rvert$
Right one-tailed	$\mu \leq \mu_0$	$\mu > \mu_0$	test stat $> t_\alpha$ or z_α
Left one-tailed	$\mu \geq \mu_0$	$\mu < \mu_0$	test stat $< -t_\alpha$ or $-z_\alpha$

Exam Tip: The null hypothesis **always contains the equality sign**. Testing H_0 is always done *at the equality*. Put what you want to prove in **H_a**.

Example Formulations

Test if average S&P 500 returns are **greater than** 10%: H_0: $\mu \leq 10\%$, H_a: $\mu > 10\%$.

Test if **less than** 10%: H_0: $\mu \geq 10\%$, H_a: $\mu < 10\%$.

Test if **not equal to** 10%: H_0: $\mu = 10\%$, H_a: $\mu \neq 10\%$.

STEP 2 · TEST STATISTIC

Choosing the Test Statistic

z-statistic (σ^2 known)

$$z = (\bar{X} - \mu_0)/(\sigma/\quad n)$$

Distributed normally.

t-statistic (σ^2 unknown)

$$t = (\bar{X} - \mu_0)/(s/\quad n)$$

t-distributed with df = n − 1.

SAMPLING FROM	SMALL SAMPLE (N < 30)	LARGE SAMPLE (N ≥ 30)
Normal distribution	Variance known: **z**	Variance known: **z**
	Variance unknown: **t**	Variance unknown: **t** (or z)
Non-normal distribution	Variance known: **N/A**	Variance known: **z**
	Variance unknown: **N/A**	Variance unknown: **t** (or z)

CHECK YOUR UNDERSTANDING · QUIZ 1

Quiz 1

Select the best answer. Feedback appears after each click.

Q1. You believe the average return of all stocks in the S&P 500 is greater than 10%. The correct hypotheses are:

Q2. If the population variance is unknown, the appropriate test statistic for a mean is:

Q3. In a two-tailed test, we reject H_0 when:

Q4. The null hypothesis:

STEP 3 · SIGNIFICANCE, ERRORS, POWER

Type I, Type II Errors & Power of a Test

	H_0 TRUE	H_0 FALSE
Fail to reject H_0	Correct decision $(1 - \alpha)$ = confidence level	**Type II error** probability = β
Reject H_0	**Type I error** probability = α (significance)	Correct decision $(1 - \beta)$ = **power of test**

Level of Significance (α)

Probability of a **Type I error** — rejecting a true H_0. Common values: 0.01, 0.05, 0.10.

$\alpha = 1 -$ confidence level. As $\alpha \downarrow$, critical value $\uparrow$, harder to reject H_0.

Power of a Test = 1 − β

Probability of correctly **rejecting a false H_0**. A powerful test detects effects that are really there.

Trade-off

As **α ↓, β ↑**. The only way to decrease BOTH is to **increase n**. Larger samples reduce SE (denominator of t-stat ↓), raising power.

Exam Tip: Mnemonic — "H_0: not pregnant / H_a: pregnant". A Type I error: say pregnant when not; Type II: say not pregnant when actually pregnant.

STEPS 4 & 6 · DECISION RULE

Decision Rule and p-values

Decision Rules by Test Type

- **Two-tail:** reject H_0 when |test-statistic| > |critical value| ($t_{\alpha/2}$ or $z_{\alpha/2}$)
- **Right-tail:** reject H_0 when test-statistic > t_α or z_α
- **Left-tail:** reject H_0 when test-statistic < $-t_\alpha$ or $-z_\alpha$

p-value

The **smallest level of significance at which H_0 can be rejected**.

- If p-value < α ⇒ **reject H_0**
- If p-value > α ⇒ **do not reject H_0**

Example: p-value = 4%. Reject at 5% significance, but NOT at 1% significance.

COMPLETE WORKED EXAMPLE

Hypothesis Test — σ Known (z-test)

WORKED EXAMPLE

S&P 500 Returns > 10%?

GIVEN

You believe avg return of all S&P 500 stocks > 10%. Sample: $n = 49$ stocks, $\bar{X} = 12\%$, population $\sigma = 4\%$. $\alpha = 5\%$.

STEPS

1. Hypotheses: H_0: $\mu \leq 10\%$; H_a: $\mu > 10\%$

2. Test statistic: σ known ⇒ use z

3. α = 5%, one-tail right

CALCULATION

$z = (\bar{X} - \mu_0) / (\sigma/\sqrt{n}) = (12 - 10) / (4/\sqrt{49}) = 2 / (4/7) = \mathbf{3.5}$

Critical value: $z_{0.05} = \mathbf{1.65}$

Since $3.5 > 1.65$ ⇒ **Reject H_0**.

ANSWER

At 5% significance, we conclude the average S&P 500 return is greater than 10%. Belief is supported.

CHECK YOUR UNDERSTANDING · QUIZ 2

Quiz 2

Select the best answer. Feedback appears after each click.

Q1. A Type I error occurs when:

Q2. The power of a test equals:

Q3. The p-value of a test is 0.04. At $\alpha = 5\%$, you should:

Q4. A sample $n = 25$, $\bar{X} = 12\%$, $s = 7\%$, testing H_0: $\mu \leq 10\%$. The t-statistic is:

Q5. To reduce BOTH Type I and Type II errors, you should:

LOS B · TESTS BEYOND THE MEAN

Difference of Means, Paired, Variance Tests

Difference of Two Means (Independent Samples)

When two populations are normally distributed with equal (unknown) variances:

$$t = [(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)]/ \quad (s_p^2/n_1 + s_p^2/n_2)$$

df = $n_1 + n_2 - 2$. Pooled variance:

$$s_p^2 = [(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2]/(n_1 + n_2 - 2)$$

Paired Comparisons (Dependent Samples)

Compute sample mean difference $\bar{d} = (1/n)\Sigma d_i$. Standard error $s_{\bar{d}} = s_d/\sqrt{n}$.

$$t = (\bar{d} - \mu_{d0})/s_{\bar{d}} \quad df = n - 1$$

Variance Tests

- **Single variance — chi-square test:** $\chi^2 = (n-1)s^2/\sigma_0^2$, df = n−1
- **Equality of two variances — F-test:** $F = s_1^2/s_2^2$

LOS C · PARAMETRIC VS NON-PARAMETRIC

Parametric vs Non-parametric Tests

	PARAMETRIC	NON-PARAMETRIC
Targets	Population parameters (μ or σ^2) using sample statistics ($\bar{X}$ or S^2)	No parameters tested
Distributional assumptions	Yes	No
Examples	z, t, F, χ^2	Spearman rank, sign test, runs test

Use Non-parametric When:

1. Data do NOT meet distributional assumptions ($n < 30$, non-normal)
2. There are **outliers** — test of median instead of mean
3. Data are given in **ranks** or use an **ordinal scale** (ordered categorical)
4. Hypothesis does NOT concern a parameter (e.g., "Is the sample random?")

Exam Tip: Parametric tests are more powerful when their assumptions hold. Switch to non-parametric when assumptions are violated or data are qualitative.

CHECK YOUR UNDERSTANDING · QUIZ 3

Quiz 3

Select the best answer. Feedback appears after each click.

Q1. A test that is MORE likely to reject a false H_0 is said to have:

Q2. You should use a NON-parametric test when:

Q3. For testing equality of two population variances, the appropriate test is the:

MODULE COMPLETE

LM08 · Hypothesis Testing Summary

Final Score

0 / 0

Cheat Sheet

$$z = (\bar{X} - \mu_0)/(\sigma/\quad n)(\sigma\ known)$$

$$t = (\bar{X} - \mu_0)/(s/\quad n), df = n - 1(\sigma\ unknown)$$

$$TypeI = \alpha; TypeII = \beta; Power = 1 - \beta$$

$$H_0 always contains equality (=,\ \leq,\ \geq)$$

$$Reject if p-value < \alpha$$

$$\chi^2 = (n-1)s^2/\sigma_0^2; F = s_1^2/s_2^2$$

Last-Minute Exam Checklist

z-test example: n=49, $\bar{X}$=12%, σ=4%, z=3.5, reject at 5%

t-test example: n=25, $\bar{X}$=12%, s=7%, t=1.43, do not reject

Critical values at 5%: $z_{0.05}$=1.65 (1-tail), $z_{0.025}$=1.96 (2-tail)

Power: 1−β, increase via larger n

Non-parametric: ranks, outliers, non-normal, small n

F-test: equality of two variances

Step 1 of 13