LEVEL 1: QUANTITATIVE METHODS

Reading 6 (6th out of 7): HYPOTHESIS TESTING

Difficulty: **hard**

Benchmark Study Time: **4.25h**

2022

INTENTIONALLY EMPTY

THIS E-BOOK:

- ❖ is a selective summary of the corresponding Reading in your CFA® Program Curriculum,
- ❖ provides place for your own notes,
- ❖ helps you structure your study and revision time!

How to use this e-book to maximize your knowledge retention:

1. **Print** the e-book in duplex and bind it to keep **all** important info for this Reading **in one place.**
2. **Read** this e-book, best twice, to grasp the idea of what this Reading is about.
3. **Study** the Reading from your curriculum. **Here add** your notes, examples, formulas, definitions, etc.
4. **Review** the Reading using this e-book, e.g. write your summary of key concepts or revise the formulas at the end of this e-book (if applicable).
5. **Done?** Go to your study plan and change the Reading's status to **green**: (it will make your Chance-to-Pass-Score™ grow ☺).
6. **Come back** to this e-book from time to time to **regularly review for knowledge retention!**

NOTE: While studying or reviewing this Reading, you can use the tables at the end of this e-book and mark your study/review sessions to hold yourself accountable.

INTRODUCTION TO HYPOTHESIS TESTING

Steps

A hypothesis is a statement about the values of parameters of one or more populations.

Testing a hypothesis involves the following steps:

1. Formulate two hypotheses – the null hypothesis and the alternative hypothesis.
2. Identify the appropriate test statistic and its probability distribution.
3. Specify the significance level.
4. Formulate the decision rule.
5. Gather the data and calculate the test statistic.
6. Make the statistical decision.
7. Make the investment or economic decision based on the statistical decision and relevant data.

Step 1 – Formulate two hypotheses

The null hypothesis and the alternative hypothesis form a pair of hypotheses that complement each other and exhaust all possible values of the parameter or parameters. The **null hypothesis** ($\mathbf{H_0}$) is the one that is tested. We assume it is true unless there is convincing evidence that it is, in fact, false.

The **alternative hypothesis** ($\mathbf{H_1}$, $\mathbf{H_a}$) is accepted when the null hypothesis is rejected.

<u>Types of hypothesis tests</u>

There are two types of hypothesis tests:

- **a two-sided** hypothesis test (two-tailed hypothesis test):
  $\mathrm{H_0}: \theta = \theta_0 \;\; \mathrm{H_1}: \theta \neq \theta_0$

- **one-sided** hypothesis tests (one-tailed hypothesis tests):
  a. a right-tailed hypothesis test: $\mathrm{H_0}: \theta \leq \theta_0 \;\; \mathrm{H_1}: \theta > \theta_0$
  b. a left-tailed hypothesis test: $\mathrm{H_0}: \theta \geq \theta_0 \;\; \mathrm{H_1}: \theta < \theta_0$

HERE KNOWLEDGE RETENTION HAPPENS | WRITE: notes, examples, formulas, definitions, relations, etc.

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

Test statistics regarding the population mean are calculated according to the following formula:

$$\text{test statistic} = \frac{(\text{sample statistic}) - (\text{value of the population parameter under } H_0)}{(\text{standard error of the sample statistic})}$$

standard error:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \text{ or } s_{\bar{X}} = \frac{s}{\sqrt{n}}$$

t-test	t-distribution
z-test	standard normal distribution or z-distribution
chi-square test	chi-square distribution
F-test	F-distribution

Step 3 – Specify the significance level

Decision rule

A decision rule is about determining conditions under which we can reject the null hypothesis.

Decision	H_0 true	H_0 false
reject H_0	Type I error	Correct decision
do not reject H_0	Correct decision	Type II error

In reaching a statistical decision, we can make the following errors:

1. Type I error, when we reject the true null hypothesis, and
2. Type II error, when we don't reject the false null hypothesis.

α = level of significance of the test = probability of a Type I error

β = probability of a Type II error

The power of a test is the probability of not committing a Type II error.

The power of a test is the probability of correctly rejecting the null hypothesis that is false.

$$\text{power of a test} = 1 - \beta$$

HERE KNOWLEDGE RETENTION HAPPENS | WRITE: notes, examples, formulas, definitions, relations, etc.

As far as Type I and Type II errors are concerned, in practice we focus on the level of significance, i.e. on the probability of a Type I error.

Step 4 – Formulate the decision rule

Critical values

If we are dealing with a two-sided hypothesis test, we've got two critical values:

- one negative, under the left tail, and
- one positive, under the right tail.

For a right-tailed test, we have one critical value that is under the right tail.

For a left-tailed test, we have one critical value that is under the left tail.

Step 5 – Gather the data and calculate the test statistic

Conclusions we draw depend not only on the appropriateness of the statistical model but also on the quality of the data we use when carrying out the test.

Steps 6 & 7 – Make the statistical decision and economic decision

The statistical decision does not determine the actual decision.

Before making the actual decision, many other nonstatistical factors should be considered.

P-value

In practice, we often encounter the so-called p-value or the marginal significance level. The p-value is the smallest level of significance at which the null hypothesis can be rejected.

If the p-value is higher than α, we have no reason to reject the null hypothesis and the test is not statistically significant. When the p-value is lower than α, the test is statistically significant and the null hypothesis is rejected.

$\text{p-value} > \alpha$ → there are no reasons to reject the null hypothesis ←→

←→ the test is statistically insignificant

$\text{p-value} \leq \alpha$ → reject the null hypothesis ←→the test is statistically significant

HERE KNOWLEDGE RETENTION HAPPENS | WRITE: notes, examples, formulas, definitions, relations, etc.

TEST CONCERNING POPULATION MEAN

Tests concerning a single population mean

	Statistic for small sample $(\mathbf{n} < \mathbf{30})$	Statistic for large sample $(\mathbf{n} \geq \mathbf{30})$
Normal distribution	t-test	t-test or z-test
Nonnormal distribution	–	t-test or z-test

<u>T-test</u>

If we have the distribution with unknown variance but:

1. the sample is large, or
2. the sample is small, but the population sampled is normally distributed or approximately normally distributed,

then we can apply the following test statistic for hypothesis tests concerning a single population mean:

$$t_{n-1} = \frac{\overline{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Where:

- t_{n-1} – t-statistic with (n minus 1) degrees of freedom,
- $\overline{X}$ – sample mean,
- μ_0 – hypothesized value of the population mean,
- n – sample size,
- s – sample standard deviation.

HERE KNOWLEDGE RETENTION HAPPENS | WRITE: notes, examples, formulas, definitions, relations, etc.

Z-test

If the sample is drawn from a normally distributed population with known variance, then the z-test concerning a single population mean is as follows:

$$z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Where:

- z – z-statistic,
- $\overline{X}$ – sample mean,
- μ_0 – hypothesized value of the population mean,
- n – sample size,
- σ – population standard deviation.

Basing on the central limit theorem, we may also use the z-test if the sample is large and we don't know the population variance:

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

Where:

- z – z-statistic,
- $\overline{X}$ – sample mean,
- μ_0 – hypothesized value of the population mean,
- n – sample size,
- s – sample standard deviation.

Tests concerning differences between population means

When we assume that two populations are distributed normally, for tests concerning differences between the population means, we use a t-test (**assumption**: <u>samples are independent of each other</u>).

Depending on data, we can assume two cases, namely:

- unknown population variances are equal, or
- unknown population variances are not equal.

HERE KNOWLEDGE RETENTION HAPPENS | WRITE: notes, examples, formulas, definitions, relations, etc.

If **unknown population variances are assumed equal**, then the formula for the t-test is:

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\left(\frac{s_P^2}{n_1} + \frac{s_P^2}{n_2}\right)^{\frac{1}{2}}}$$

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

Where:

- s_P^2 – pooled estimator.
- $(n_1 + n_2 - 2)$ – number of degrees of freedom,
- $\overline{X}$ – sample mean,
- μ – population mean,
- n – sample size,
- s – sample standard deviation.

If **unknown variances are assumed unequal**, then the formula for the t-test is:

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

Where:

$$\text{number of degrees of freedom} = df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2}}$$

- $\overline{X}$ – sample mean,
- μ – population mean,
- n – sample size,
- s – sample standard deviation.

HERE KNOWLEDGE RETENTION HAPPENS | WRITE: notes, examples, formulas, definitions, relations, etc.

Tests concerning mean differences between paired values

When samples are not independent of each other → use the paired-comparisons test.

Assumptions:

- populations are normally distributed,
- variances are unknown, and
- observations are paired.

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

$$s_d^2 = \frac{\sum_{i=1}^{n}(d_i - \bar{d})^2}{n - 1}$$

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$$

$$s_{\bar{d}} = \frac{s_d}{\sqrt{n}}$$

$$\text{number of degrees of freedom} = n - 1$$

Where:

- n – number of pairs in the sample,
- $\bar{d}$ – sample mean difference,
- s_d^2 – sample variance of differences,
- $s_{\bar{d}}$ – standard error of the sample mean difference.

HERE KNOWLEDGE RETENTION HAPPENS | WRITE: notes, examples, formulas, definitions, relations, etc.

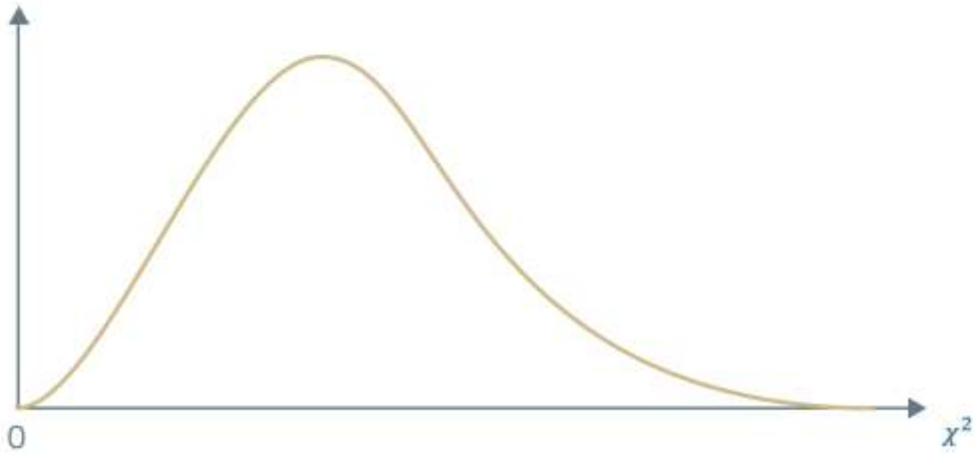
TEST CONCERNING POPULATION VARIANCE

Tests concerning a single variance of a normally distributed population

In tests concerning a single variance of a normally distributed population, we make use of a chi-square test statistic.

A **chi-square distribution:**

1. is defined by one parameter (the number of degrees of freedom),
2. is asymmetrical,
3. is bounded below by zero.



The chi-square test statistic for a normally distributed population:

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

Where:

- χ^2 – chi-square test statistic,
- σ_0^2 – hypothesized value of the variance,
- n – sample size,
- $(n-1)$ – number of degrees of freedom,
- $\overline{X}$ – sample mean,
- s – sample standard deviation.

Values of the chi-square test are usually given in the table for probability in the right tail.

$\boldsymbol{\alpha}$ probability in the right tail = $(1 - \boldsymbol{\alpha})$ probability in the left tail

HERE KNOWLEDGE RETENTION HAPPENS | WRITE: notes, examples, formulas, definitions, relations, etc.

For a two-sided hypothesis test, we reject the null hypothesis if the test statistic is:

- greater than the value for the chi-square distribution
  with a probability of $\boldsymbol{\alpha/2}$ in the right tail

OR

- less than the value for the chi-square distribution
  with a probability of $\boldsymbol{\alpha/2}$ in the left tail

In the table for probability in the right tail, check if the value of the test is:

- greater than the value for the chi-square distribution
  with a probability of $\boldsymbol{\alpha/2}$ in the right tail

OR

- less than the value for the chi-square distribution
  with a probability of $\boldsymbol{1-\alpha/2}$ in the right tail

For a right-tailed hypothesis test, we reject the null hypothesis if the test statistic is greater than the value for the chi-square distribution with a probability of $\boldsymbol{\alpha}$ in the right tail.

For a left-tailed hypothesis test, we reject the null hypothesis if the test statistic is less than the value for the chi-square distribution with a probability of $\boldsymbol{\alpha}$ in the left tail (=probability of $\boldsymbol{1-\alpha}$ in the right tail).

Tests concerning the equality or inequality of two variances of normally distributed populations

Assuming that the samples drew from two normally distributed populations are independent, we can test them with an F-test. The F-test is the ratio of sample variances:

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

With:

- $df_1 = n_1 - 1$ numerator degrees of freedom
- $df_2 = n_2 - 1$ denominator degrees of freedom

HERE KNOWLEDGE RETENTION HAPPENS | WRITE: notes, examples, formulas, definitions, relations, etc.

The F-test concerning the equality or inequality of variances of two populations makes use of the F-distribution. Like the chi-square distribution, the F-distribution is a family of asymmetrical distributions bounded from below by 0. Each F-distribution is defined by two values of degrees of freedom, called the numerator and denominator degrees of freedom. While conducting the F-test, we have to compute the sample variances for both populations.

For the F-test, put the greater variance in the numerator

and the lower one in the denominator!

Rejection points for hypothesis tests

For a two-sided hypothesis test: we reject the null hypothesis if the test statistic is greater than the value for the F-distribution with a significance level of $\alpha/2$ and the numerator and denominator degrees of freedom given.

For a right- and left-tailed test: we reject the null hypothesis if the test statistic is greater than the value for the F-distribution with a significance level of α and the numerator and denominator degrees of freedom given.

Tests concerning correlation

For normally distributed variables, we can use the sample correlation coefficient and t-test with **n-2** degrees of freedom to test whether the two variables are linearly correlated.

We use a two-tailed test with:

- the null hypothesis stating that the correlation coefficient in the population is equal to 0, and
- the alternative hypothesis stating that the correlation coefficient is different than 0.

The formula for the t-test:

$$t = \frac{r \times \sqrt{n-2}}{\sqrt{1-r^2}}$$

Where:

- r – sample correlation coefficient,
- n – sample size,
- $(n-2)$ – degrees of freedom,

We reject the null hypothesis if the test statistic is either larger or smaller than the critical level we can find it the t-table, e.g. if $t_c = 2.678$, then we reject the null if the test statistic is larger than **2.678** or smaller than **-2.678**.

HERE KNOWLEDGE RETENTION HAPPENS | WRITE: notes, examples, formulas, definitions, relations, etc.

Parametric tests vs Nonparametric tests

The division of tests into parametric and nonparametric is said to be one of the most important in hypothesis testing. Parametric tests are used in hypothesis testing concerning population parameters such as the mean or the variance.

Nonparametric tests are used mainly in the following situations:

- when the hypothesis is not related to any parameter,
- when we make no or minimal assumptions about the population distribution, or
- when data are given in ranks.

Parametric tests are generally better and more accurate than nonparametric ones. They are characterized by a broader range of assumptions that need to be satisfied. Also, they are more powerful and their results are easy to interpret.

An example of a nonparametric test is a test based on the **Spearman rank correlation coefficient**.

It is one of the most frequently used tests to examine the correlation between two variables. We often employ it when we cannot use a t-test because random variables don't meet assumptions about the distribution. To calculate the Spearman rank correlation coefficient, first we have to give a number to each observation from a sample. 1 is for the largest observation, 2 is for the second largest one, 3 is for the third largest one, and so on. The process of giving each observation a number is carried out for both random variables.

We calculate the Spearman rank correlation coefficient (r_S) in the following way:

$$r_S = 1 - \frac{6 \times \left(\sum_{i=1}^{n} d_i^2\right)}{n \times (n^2 - 1)}$$

Where:

- n – number of observations,
- d_i – the difference between the ranks of each pair of observations of the two random variables.

To decide whether to reject the null or not, we use a t-test if the sample is large enough. In other cases, we use special tables to find critical values.

HERE KNOWLEDGE RETENTION HAPPENS | WRITE: notes, examples, formulas, definitions, relations, etc.

Summarizing key concepts:

☐ 7 steps in hypothesis testing

My summary:

☐ Null hypothesis ($\mathbf{H_0}$) vs Alternative hypothesis ($\mathbf{H_1}$, $\mathbf{H_a}$), Two-sided hypothesis test vs One-sided hypothesis test, Significance level

My summary:

☐ Type I error vs Type II error

My summary:

☐ Power of a test

My summary:

☐ P-value

My summary:

☐ Tests concerning a single population mean

My summary:

☐ Tests concerning differences between population means

My summary:

☐ Tests concerning mean differences between paired values

My summary:

☐ Tests concerning a single variance of a normally distributed population

My summary:

☐ Tests concerning the equality or inequality of two variances of normally distributed populations

My summary:

☐ Tests concerning correlation

My summary:

☐ Parametric tests vs Nonparametric tests

My summary:

Reviewing formulas:

$$\text{test statistic} = \frac{(\text{sample statistic}) - (\text{value of the population parameter under } H_0)}{(\text{standard error of the sample statistic})}$$

Write down the formula:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} \text{ or } s_{\overline{X}} = \frac{s}{\sqrt{n}}$$

Write down the formula:

$$t_{n-1} = \frac{\overline{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

$$z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

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

Write down the formula:

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\left(\frac{s_P^2}{n_1} + \frac{s_P^2}{n_2}\right)^{\frac{1}{2}}}$$

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

Write down the formula:

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

Where:

$$\text{number of degrees of freedom} = \text{df} = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2}}$$

Write down the formula:

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

$$s_d^2 = \frac{\sum_{i=1}^{n}(d_i - \bar{d})^2}{n - 1}$$

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$$

$$s_{\bar{d}} = \frac{s_d}{\sqrt{n}}$$

$$\text{number of degrees of freedom} = n - 1$$

Write down the formula:

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

Write down the formula:

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

Write down the formula:

$$t = \frac{r \times \sqrt{n-2}}{\sqrt{1-r^2}}$$

Write down the formula:

$$r_S = 1 - \frac{6 \times \left(\sum_{i=1}^{n} d_i^2\right)}{n \times (n^2 - 1)}$$

Write down the formula:

Keeping myself accountable:

TABLE 1 | STUDY

When you sit down to study, you may want to **try the Pomodoro Technique** to handle your study sessions: study for 25 minutes, then take a 5-minute break. Repeat this 25+5 study-break sequence all throughout your daily study session.

✔ Tick off as you proceed.

POMODORO TIMETABLE: study-break sequences (25′ + 5′)													
date		date		date		date		date		date		date	
25′		25′		25′		25′		25′		25′		25′	
5′		5′		5′		5′		5′		5′		5′	
25′		25′		25′		25′		25′		25′		25′	
5′		5′		5′		5′		5′		5′		5′	
25′		25′		25′		25′		25′		25′		25′	
5′		5′		5′		5′		5′		5′		5′	
25′		25′		25′		25′		25′		25′		25′	
5′		5′		5′		5′		5′		5′		5′	

TABLE 2 | REVIEW

Never ever neglect revision! Though it's not the most popular thing among CFA candidates, regular revision is what makes the difference. If you want to pass your exam, **schedule & do your review sessions.**

REVIEW TIMETABLE: When did I review this Reading?													
date		date		date		date		date		date		date	
date		date		date		date		date		date		date	