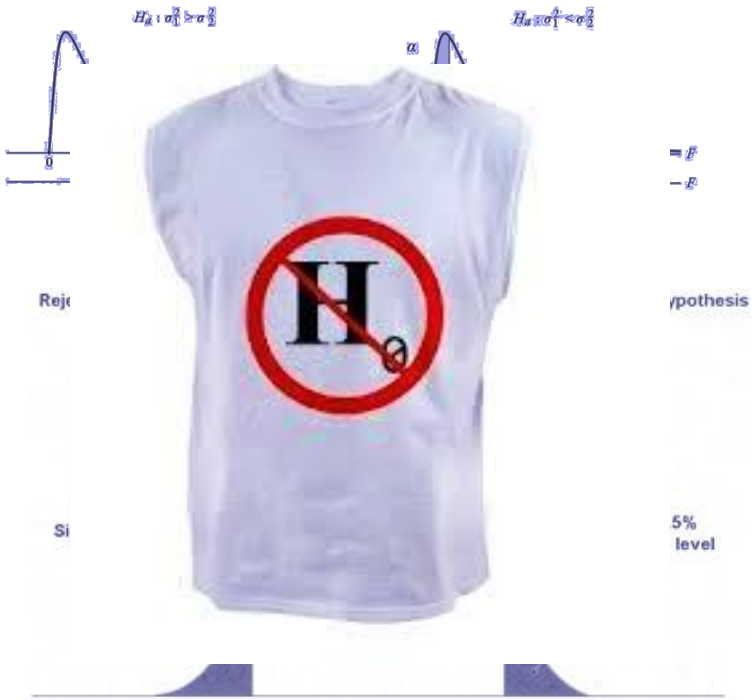


ESIC EDITORIAL

Chapter 13

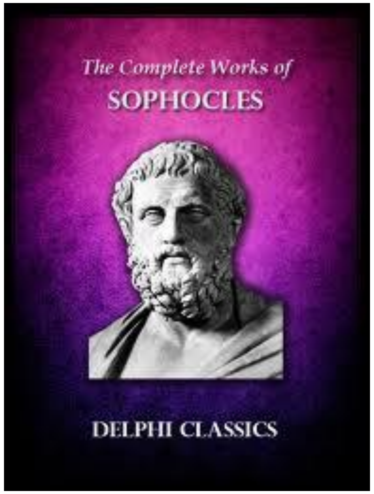
Hypothesis Testing

Business Research Methods
Verónica Rosendo Ríos
Enrique Pérez del Campo

Verónica Rosendo Ríos ©

"It is horrible to speak well and be wrong"

SOPHOCLES



Verónica Rosendo Ríos ©

CONTENTS

1. Hypothesis testing (Part I)
2. Testing process
    1. Proportion Tests
    2. Mean Tests

Verónica Rosendo Ríos ©

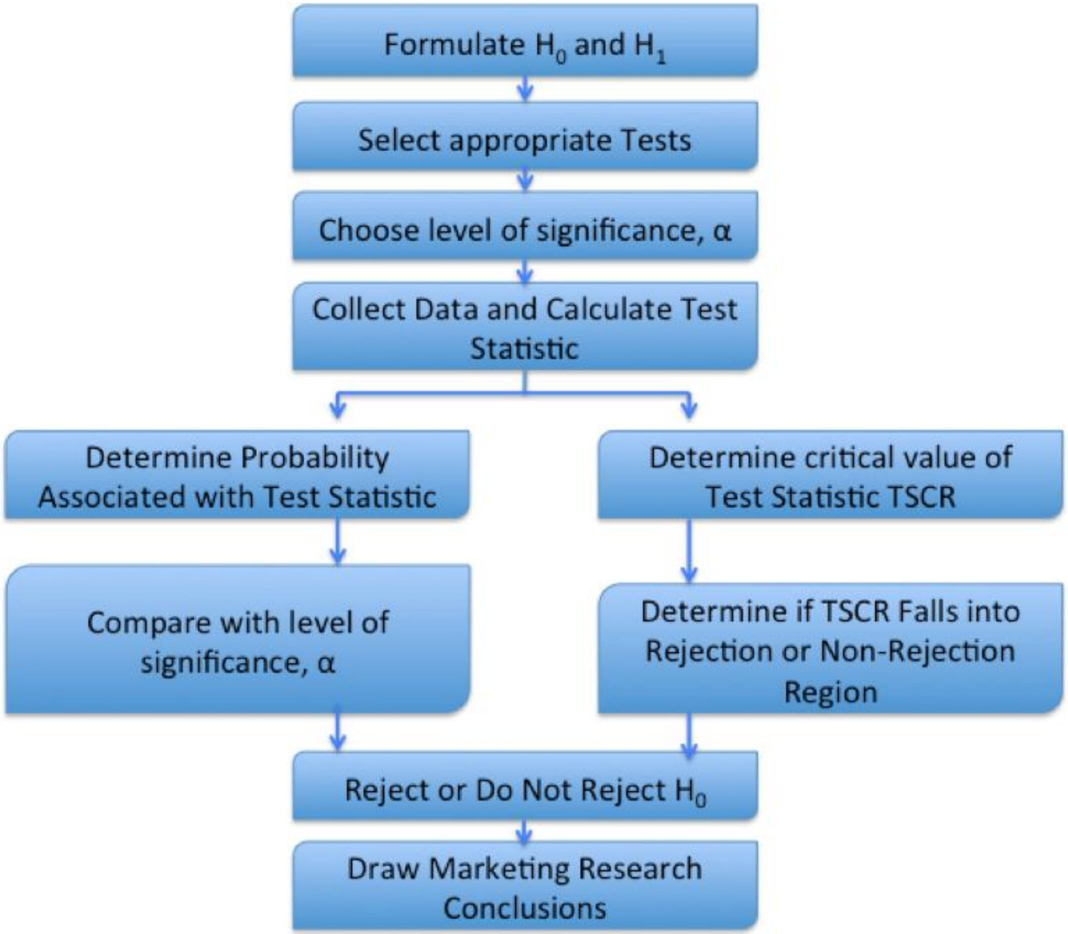
★ FREQUENCY DISTRIBUTION:

Statistical analysis can be divided into several groups:

- **Univariate statistical analysis** tests hypotheses involving only one variable.
- **Bivariate statistical analysis** test hypotheses involving two variables.
- **Multivariate statistical analysis** test hypotheses involving multiple (three or more) variables or sets of variables.

Verónica Rosendo Ríos ©

★ Hypothesis testing process:



Verónica Rosendo Ríos ©

★ **Step 1**: Formulating the Hypothesis

-A **null hypothesis** is a statement of the status quo, one of no difference or no effect. If the null hypothesis is not rejected, no changes will be made.

-An **alternative hypothesis** is one in which some difference or effect is expected. Accepting the alternative hypothesis will lead to changes in opinions or actions. Thus, the alternative hypothesis is exactly the opposite of the null hypothesis.

<u>The null hypothesis is always the hypothesis that is tested</u>. The null hypothesis refers to a specified value of the population parameter (e.g. μ, σ, or π).

<u>A null hypothesis may be rejected, but it can never be accepted based on a single test</u>!!

Verónica Rosendo Ríos ©

✓ Statistical Tests

A statistical test can have one or two outcomes:

1. that the null hypothesis is rejected and the alternative hypothesis is accepted or
2. that the null hypothesis is NOT rejected based on the evidence.

☑ It would be incorrect, however, to conclude that since the null hypothesis is not rejected it can be accepted as valid. In classical hypothesis testing there is no way to determine whether the null hypotheses are true.

Verónica Rosendo Ríos ©

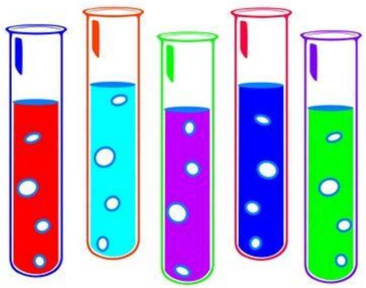
✓ Statistical Tests

In marketing research, the null hypothesis is formulated in such a way that *its rejection leads to the acceptance of the desired conclusion*. *The alternative hypothesis represents the conclusion for which evidence is sought*.

For example, an industrial marketing firm is considering the introduction of a new servicing plan for hydraulic parts. The plan will be introduced if it is preferred by more than 40% of the customers. The appropriate way to formulate the hypotheses is:

$$H_0 = \pi \leq 0.40$$

$$H_1 = \pi > 0.40$$



✓ If the null hypothesis H_0 is rejected, then the alternative hypothesis H_1 will be accepted and the new service plan introduced. On the other hand, if H_0 is not rejected, then the new service plan should not be introduced unless additional evidence is obtained.

Verónica Rosendo Ríos ©

✓ Statistical Tests

➢ One tail Test

The alternative hypothesis is expressed directionally: the proportion of customers who express a preference is greater than 0.4.

➢ Two tail Test

On the other hand, suppose that the researcher wanted to determine whether the new service plan is different (superior or inferior) from the current plan, which is preferred by 40% of the customers. Then a **two tailed test** would be required, and the hypotheses would be expressed as:

$$H_0: \ \pi = 0.40$$

$$H_1 \ \ \pi \neq 0.40$$

In commercial marketing research, the one-tailed test is the most common test. Typically, there is some preferred direction for the conclusion for which evidence is sought. For example, the higher the profits, sales, and product quality, the better.

Verónica Rosendo Ríos ©

★ **Step 2**: Selecting an Appropriate Statistical Technique

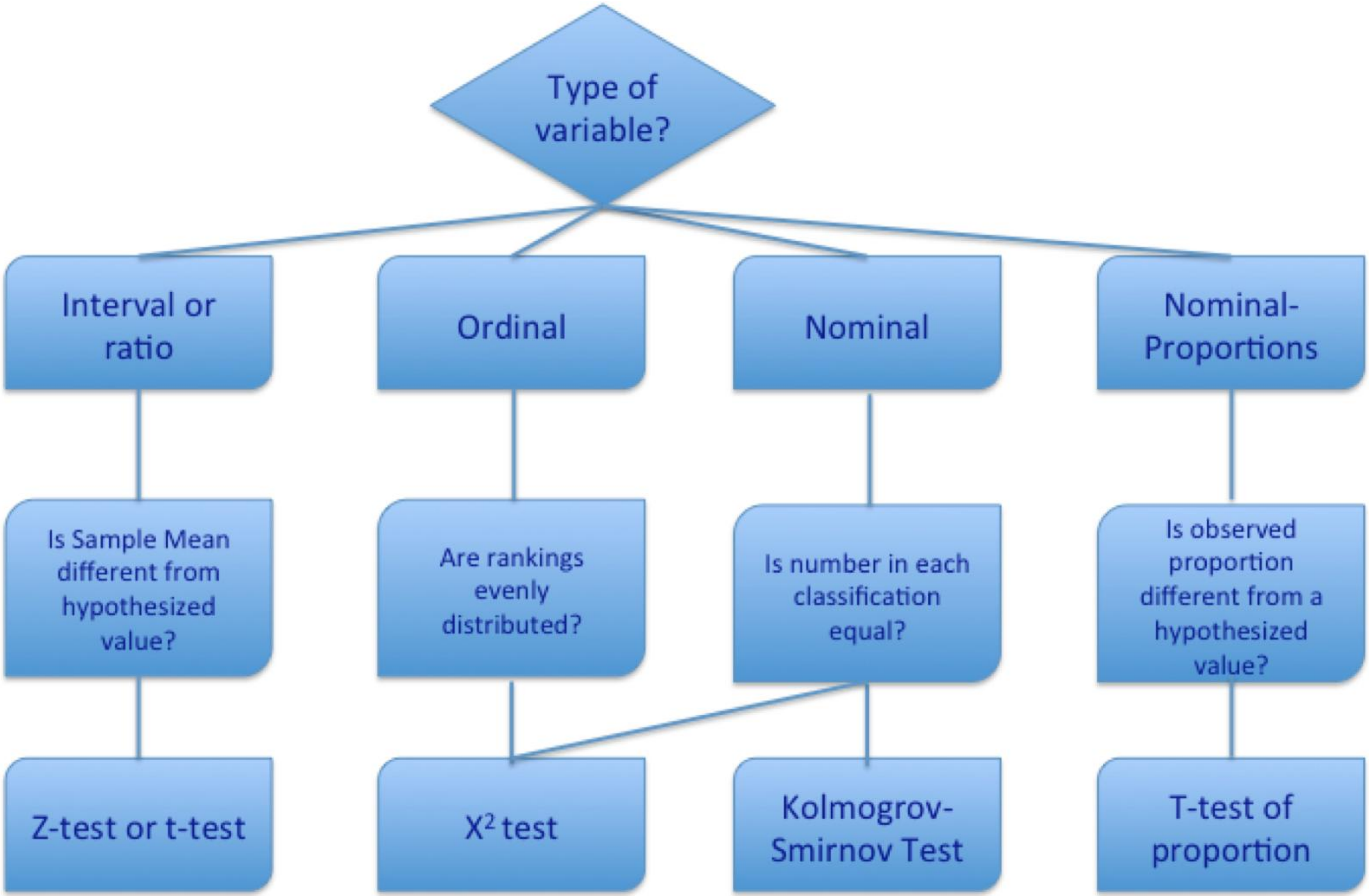
The test statistic often follows a well-known distribution, such as the *normal*, *t*, or *chi-square* distribution.

Numerous statistical techniques are available to assist the researcher in interpreting the data. The more difficult task, however, is determining when to use each method. Making the correct choice can be determined by considering:

- ✓ The type of question to be answered.
- ✓ The number of variables involved
- ✓ The level of scale measurement
- ✓ Type of distribution: Parametric vs non-parametric (or distribution free) tests

Verónica Rosendo Ríos ©

Univariate Statistical Choice:



Verónica Rosendo Ríos ©

✓ The T distribution

- ✓ Symetrical bell-shaped distribution with a mean=0 and std dev.=1
- ✓ When sample size (n) is larger than 30, the t-distribution and the z-distribution are almost identical.
- ✓ The z-distribution and the t-distribution are very similar, and thus the z-test and t-test will provide much the same results in most situations.
- ✓ However, when the population standard deviation (σ) is known, the Z-test is most appropriate.
- ✓ When σ is unknown (the situation in most business research studies), and the sample size is greater than 30, the Z-test can also be used.
- ✓ When σ is unknown and the sample size is small, the t-test is more appropriate.

Verónica Rosendo Ríos ©

★ Step 3: Choosing the Level of Significance

There is always a risk that when we draw inferences of a population based on a sample, an incorrect conclusion is made. The researcher using sampling runs the risk of committing *two types of errors*:

	Decision: Accept H_0	Decision: Reject H_0
Realitiy: H_0 is true	Correct-no error ☺	Type I error FALSE POSITIVE
Realitiy: H_0 is false	Type II error FALSE NEGATIVE	Correct-no error ☺

Verónica Rosendo Ríos ©

➢ Type I Error

FALSE POSITIVE (Rejecting H_0 when it is true).

Examples of Type I errors:

-H_0= The patient is healthy/innocent. We reject H_0 and consider the patient to be ill/guilty when he is in fact healthy/innocent.

The probability of Type I error (α) is called the Level of Significance (i.e. the acceptance level of type I error).

➢ Type II Error.

FALSE NEGATIVE (Accepting H_0 when it is false)

-H_0= The patient is healthy/innocent. We accept H_0 and consider the patient to be healthy/innocent when he is in fact ill/guilty.

The probability of type II error is denoted by ß (the incorrect decision is called ß). Unlike α, which is specified by the researcher, the magnitude of ß depends on the actual value of the population parameter (proportion). The complement (1-ß) of the probability of a type II error is called the Power of a statistical test.

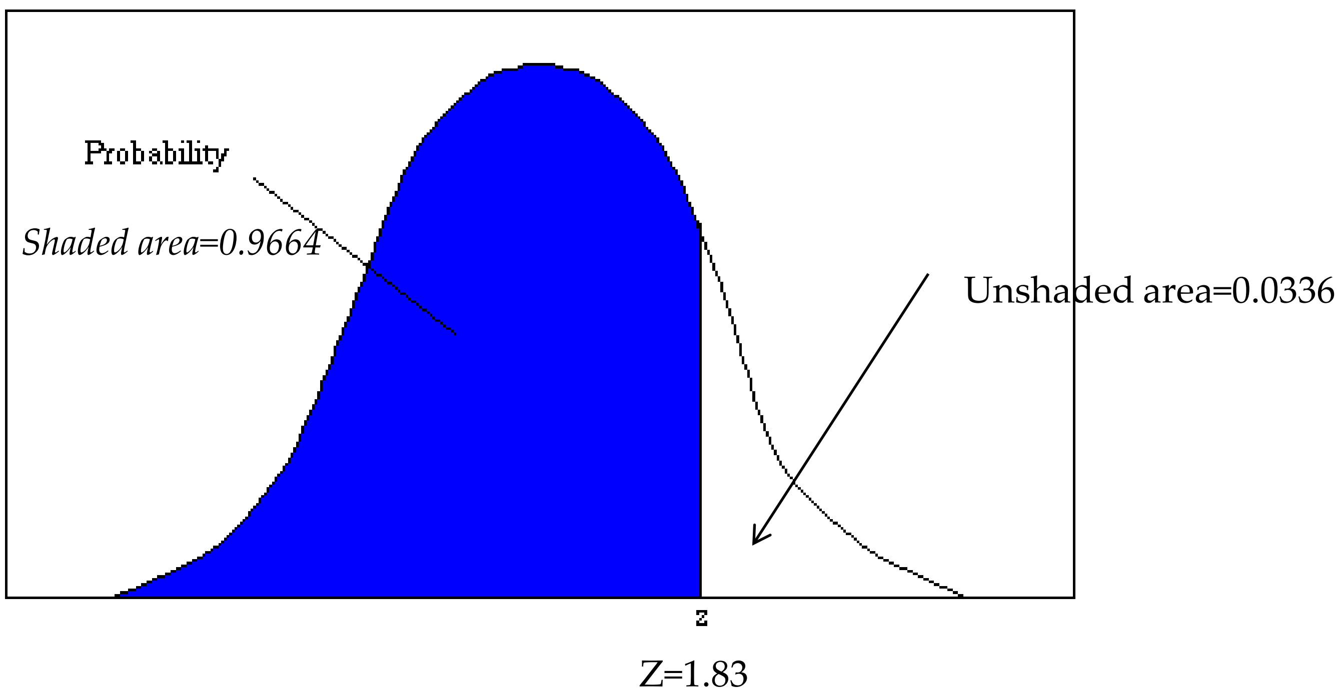
Verónica Rosendo Ríos ©

★ Power of a test:

Probability 1-ß

Although ß is unknown, it is related to α.

Probability of Z with a one-tail test



Verónica Rosendo Ríos ©

★ **Step 4**: Data Collection

- ✓ Sample size is determined after taking into account the desired α and other considerations, such as budget constraints.
- ✓ Then, the required data are collected and the value of the test statistic computed.
- ✓ Suppose, in our previous example, that 500 customers were surveyed and 220 expressed a preference for the new service plan. Thus, the value of the sample proportion is p=220/500= 0.44 (44%).
- The value of the standard deviation of the sample proportion $\boldsymbol{\sigma_p}$ can be determined as follows:

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} = \sqrt{\frac{(0.40)(0.6)}{500}} = 0.0219$$

The critical value can be calculated as:

$$z\text{-value} = \frac{p-\pi}{\sigma_p} = \frac{0.44-0.40}{0.0219} = 1.83$$

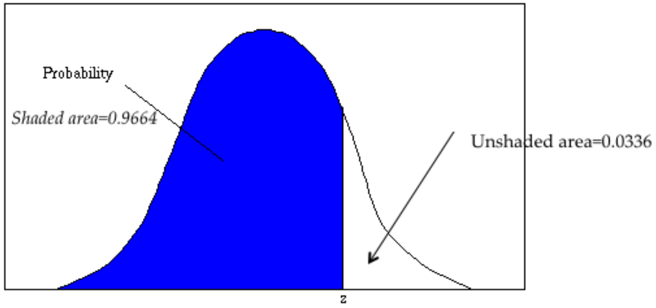
Verónica Rosendo Ríos ©

★ **Step 5**: Determining the Probability (Critical Value)

There are two ways:

1. Using standard normal tables: the probability of obtaining a *z* value of 1.83 can be calculated using standard normal tables (see figure below). The shaded area between -∞ and 1.83 is 0.9664 (see table 1 in the appendix). Therefore, the area to the right of z= 1.83 is 1.0000-0.9664= 0.0336.

Probability of Z with a one-tail test



- 2. Alternatively, the critical value of z, which will give an area to the right side of the critical value of $\alpha/2$=0.05 or α=0.10 (90% confidence interval), is between 1.64 and 1.65 and equals 1.645. (Note that in determining the critical value of the test statistic, the area to the right of the critical value is either α or $\alpha/2$. It is α for a one tailed test and $\alpha/2$ for a two-tailed test).

Verónica Rosendo Ríos ©

★ Step 6 & 7: Comparing the Probability (Critical Value) and Making the Decision

There are two ways of testing the null hypothesis:

E.g. an industrial marketing firm is considering the introduction of a new servicing plan. The plan will be introduced if it is preferred by more than 40% of the customers.

$$H_0: \ \pi \leq 0.40$$

$$H_1 \ \ \pi > 0.40$$

1. The probability associated with the calculated or observed value of the test statistic is 0.0336. This is the probability of getting a p value of 0.44 when π=0.40. This is less than the level of significance of 0.05. Hence, the null hypothesis is rejected. (0.0336<0.05; Reject $H_{0;}$; The new plan is introduced because it is preferred by more than 40% of the customers)

2. Alternatively, the calculated value of the test statistic z= 1.83 lies in the rejection region, beyond the value of 1.645. Again, the same conclusion to reject the null hypothesis is reached. (1.83>1.645; Reject $H_{0;}$; The new plan is introduced)

Verónica Rosendo Ríos ©

★ **Step 8**: Marketing Research Conclusion

E.g. an industrial marketing firm is considering the introduction of a new servicing plan. The plan will be introduced if it is preferred by more than 40% of the customers.

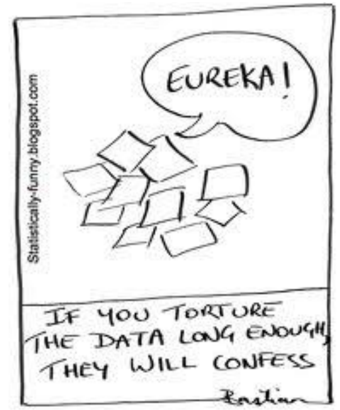
$$H_0: \ \pi \leq 0.40$$

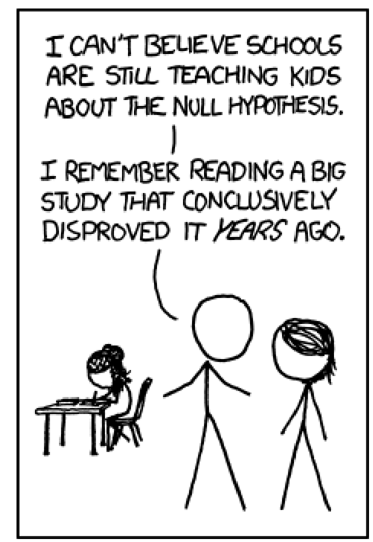
$$H_1 \ \ \pi > 0.40$$

- ✓ The conclusion reached by hypothesis testing must be expressed in terms of the marketing research problem.
- ✓ In our examples, we conclude that there is evidence that the proportion of customers preferring the new service plan is significantly greater than 0.4. Hence, the recommendation would be to introduce a new service plan.

Verónica Rosendo Ríos ©

CONTENTS





Verónica Rosendo Ríos ©

★ Hypothesis testing can be broadly classified as: Parametric VS Non-Parametric

Parametric Statistics

Are based on the assumption that the data in the study are drawn from a population with a normal (bell-shaped) distribution and/or normal sampling distribution. E.g: t-test, z-test…

Non-Parametric Statistics

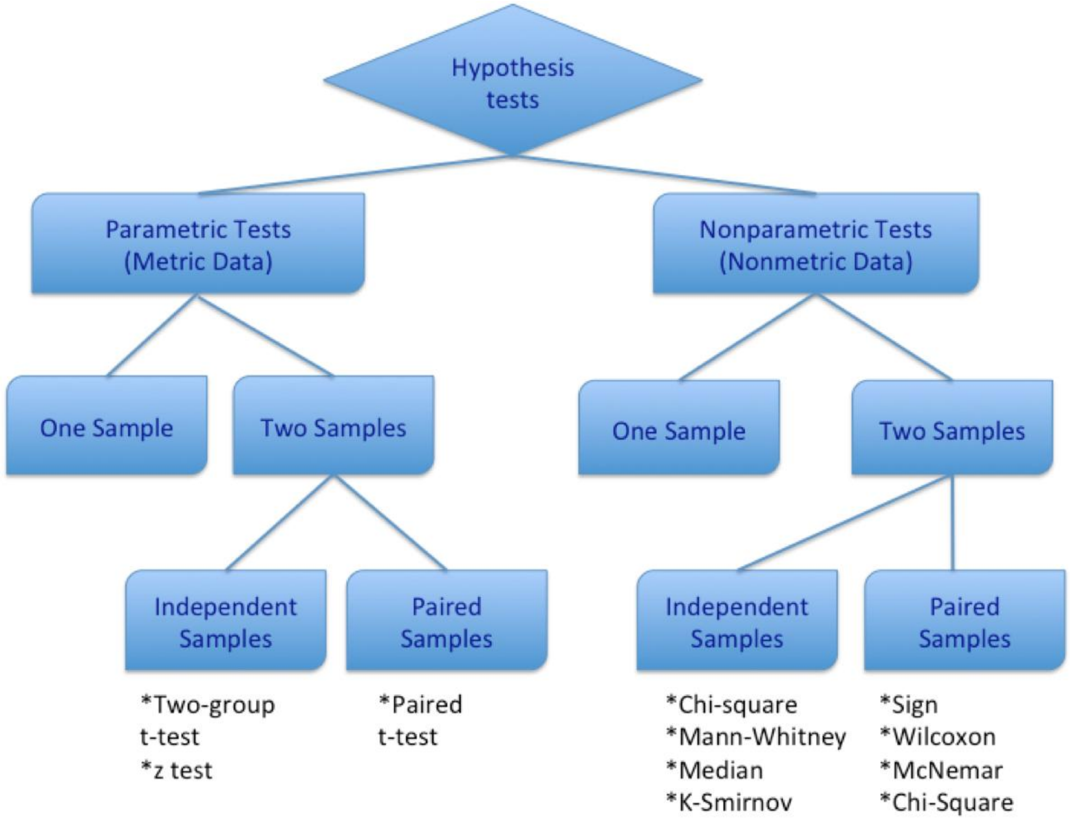
Are used when the researcher *does not know how the data are distributed*. Making the assumption that the population distribution or sampling distribution is normal generally is inappropriate when data are either nominal or ordinal.

Thus, nonparametric statistics are referred to as *distribution free*. Data analysis for both nominal and ordinal scales typically uses nonparametric statistical tests. E.g. Kolmogorov-Smirnov, Chi-Square…

Verónica Rosendo Ríos ©

★ Hypothesis testing can be broadly classified as: Parametric VS Non-Parametric

Hypothesis Testing Procedures



Verónica Rosendo Ríos ©

★ Parametric Tests: Means and Proportions

Parametric Statistics

✓ One Sample

The researcher is often interested in making statements about a single variable against a known or given standard.

E.g. At least 65% of customers will like a new package design, and 80% of dealers will prefer the new pricing policy.

These statements can be translated to null hypotheses that can be tested using a one-sample test, such as the t test or the z test.

In the case of a t test for a single mean the researcher is interested in testing whether the population mean conforms to a given hypothesis (H_0).

★ Imagine that a new machine attachment would be introduced if it receives a mean of at least 7 on a ten-point likert-scale. A sample of 20 purchase engineers is shown the attachment to evaluate it. The results indicate a mean rating of 7.9 with a standard deviation of 1.6. A significance level of α=0.05 is selected. Should the part be introduced?

Verónica Rosendo Ríos ©

★ Parametric Tests: Means and Proportions

Parametric Statistics

✓ One Sample

- ★ Imagine that a new machine attachment would be introduced if it receives a mean of at least 7 on a ten-point likert-scale. A sample of 20 purchase engineers is shown the attachment to evaluate it. The results indicate a mean rating of 7.9 with a sample standard deviation of 1.6. A significance level of α=0.05 is selected.
- ★ Should the part be introduced?

$$H_0 = \mu \leq 7.0$$

$$H_1 = \mu > 7.0$$

$$t = \frac{(X_{\text{high bar}} - \mu)}{S_{x\_\text{high bar}}}$$

$$S_{x\_\text{high bar}} = S/\sqrt{n}\ ;\ \ S_{x\_\text{high bar}} = 1.6/\sqrt{20} \quad 1.6/4.472 = 0.358$$

$$t = (7.9-7.0)/0.358 = 0.9/0.358 = \mathbf{2.514}$$

Verónica Rosendo Ríos ©

★ Parametric Tests: Means and Proportions

Parametric Statistics

✓ One Sample

★ Should the part be introduced?

t= (7.9-7.0)/0.358 = 0.9/0.358 = **2.514**

2.54>1.7291 We reject H_0. The new part is introduced.

- The degrees of freedom for the t statistic to test hypothesis about one mean are n-1. In our case 20-1=19 degrees of freedom.
- Two ways of checking:

1. From the table (t-distribution) a probability of getting a more extreme value than 2.514 is less than 0.05.
2. Alternatively, the critical t value for 19 degrees of freedom and a significance level of 0.05 is 1.7291, which is less than the calculated value of 2.514. Hence, the null hypothesis is rejected, favoring the introduction of the part.

Verónica Rosendo Ríos ©

★ Parametric Tests: Means and Proportions

Parametric Statistics

✓ One Sample

- ★ Should the part be introduced?
- ★ If we know the population standard deviation we have to use **z test** instead of a t test:

$$z = \frac{X_{\_high\ bar} - \mu}{\sigma_{X\_high\ bar}}$$

$$\sigma_{X\_high\ bar} = 1.5/\sqrt{20} = 1.5/\ 4.472 = 0.335$$

$$\text{and } z = 7.9 - 7.0\ /0.358 = 0.9\ /\ 0.358 = 2.514$$

Verónica Rosendo Ríos ©

★ Parametric Tests: Means and Proportions

Parametric Statistics

✓ One Sample

★ Should the part be introduced?

and $z = 7.9 - 7.0 / 0.358 = 0.9 / 0.358 = 2.514$

$1.645 < 2.514$ Therefore, we reject H_0

- From the table (Table 2. Area under the normal curve), the probability of getting a more extreme value of z than 2.514 is less than 0.05. (Alternatively, the critical z value for a one-tailed test and a significance level of 0.05 is 1.645, which is less than the calculated value). Therefore, the null hypothesis is rejected, reaching the same conclusion arrived at earlier by the t test (Malhotra, 2012).
-

Verónica Rosendo Ríos ©

★ Parametric Tests: Means and Proportions

Parametric Statistics

✓ Two Independent Samples

- Several hypothesis in marketing relate to parameters from two different populations:

E.g.:

users and non users of a brand may differ in terms of their perceptions of the brand, or the proportion of brand loyal users in segment I is more than the proportion in segment II.

- Samples drawn randomly from different populations are termed **independent samples**.
- As in the case for one sample, the hypotheses could relate to means or proportions.

Verónica Rosendo Ríos ©

★ Parametric Tests: Means and Proportions

Parametric Statistics

✓ Two Independent Samples

Means

In the case of means for two independent samples, the hypotheses take the following form:

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

- The two populations are sampled, and the means and variances are computed based on samples of sizes n_1 and n_2.

Verónica Rosendo Ríos ©

★ Parametric Tests: Means and Proportions

Parametric Statistics

✓ Two Independent Samples

Means

✓ If both populations are found to have ***the same variance***, a pooled variance estimate is computed from the two sample variances. The standard deviation of the test statistic and the appropriate value of t are then estimated. The degrees of freedom in this case are (n_1+n_2-2).

Verónica Rosendo Ríos ©

★ Parametric Tests: Means and Proportions

Parametric Statistics

✓ Two Independent Samples

Means

- ✓ If the two populations have *unequal variances*, and exact t cannot be computed for the difference in sample means. Instead, an approximation to t is computed.
- An **F test** of sample variance may be performed if it is not known whether the two populations have equal variance. In this case the hypotheses are:

$$H_0: \boldsymbol{\sigma}_1 = \boldsymbol{\sigma}_2$$

$$H_1: \boldsymbol{\sigma}_1 \neq \boldsymbol{\sigma}_2$$

Verónica Rosendo Ríos ©

★ Parametric Tests: Means and Proportions

Parametric Statistics

✓ Two Independent Samples

Means

- An **F test** of sample variance may be performed if it is not known whether the two populations have equal variance. In this case the hypotheses are:

$$H_0: \ \boldsymbol{\sigma}_1 = \boldsymbol{\sigma}_2$$
$$H_1: \boldsymbol{\sigma}_1 \neq \boldsymbol{\sigma}_2$$

If the probability of F is greater than the significance level α, H_0 is not rejected and t based on the pooled variance estimate can be used.

On the other hand, if the probability of F is less than or equal to α, H_0 is rejected and t based on a separate variance estimate is used.

Verónica Rosendo Ríos ©

✓ Two Independent Samples

Means

E.g.: Imagine that the researcher wanted to determine whether respondents who are familiar with a store attach different importance to store credit and billing policies than those who are unfamiliar with the store. As before, respondents are divided into two familiarity groups based on a median split. A two independent sample test was conducted, and the results are presented in the following table:

	Summary Statistics Number of Cases	Mean	Standard Deviation
Unfamiliar Group	135	3.9778	1.604
Familiar Group	132	4.3712	1.627

F Test for Equality of Variances

F Value	Two-tailed probability
1.03	0.871

t test

Pooled Variance Estimate			Separate Variance Estimate		
t value	Degrees of freedom	Two-tailed probability	t value	Degrees of freedom	Two tailed probability
-1.99	265	0.048	-1.99	264.56	0.048

)87>0.05, use pooled
timate

Verónica Rosendo Ríos ©

✓ Two Independent Samples

Means

	Summary Statistics Number of Cases	Mean	Standard Deviation
Unfamiliar Group	135	3.9778	1.604
Familiar Group	132	4.3712	1.627

F Test for Equality of Variances

F Value	Two-tailed probability
1.03	0.871

t test

Pooled Variance Estimate			Separate Variance Estimate		
t value	Degrees of freedom	Two-tailed probability	t value	Degrees of freedom	Two tailed probability
-1.99	265	0.048	-1.99	264.56	0.048

.087>0.05, use pooled
stimate

- The t value is -1.99, and with 265 degrees of freedom, that gives a probability of 0.048, which is less than the significance level of 0.05. Therefore, the null hypothesis of equal means is rejected.
- <u>**Conclusion**</u>: Since the mean importance for the unfamiliar group is 3.9778 and that for the familiar group is 4.3712, those who are familiar attach significantly greater importance to store credit and billing policies when selecting a store than those who are not familiar.

Verónica Rosendo Ríos ©

✓ Two Independent Samples

Proportions

Comparison of proportions of jean users in USA and Hong Kong

	Users	Non users	Row Totals
United States	160	40	200
Hong Kong	120	80	200
Column Totals	280	120	

Is the proportion of users the same in the US and Hong Kong examples? The null and alternative hypotheses are:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

A z test is used as in testing the proportion for one sample. In this case, however, the test statistic is given by:

$$Z = \underline{P_1 - P_2}$$

$$S_{P1-P2\ (\_high\ bars)}$$

Verónica Rosendo Ríos ©

✓ Two Independent Samples

Proportions

Comparison of proportions of jean users in USA and Hong Kong

	Users	Non users	Row Totals
United States	160	40	200
Hong Kong	120	80	200
Column Totals	280	120	

A z test is used as in testing the proportion for one sample. In this case, however, the test statistic is given by:

$$Z = \frac{P_1 - P_2}{S_{P1-P2\ (\_high\ bars)}}$$

$S_{P1-P2\ (\_high\ bars)} = \sqrt{P(1-P)\ (1/n_1 + 1/n_2)}$ where:

$P = n_1P_1 + n_2P_2 / n_1 + n_2$ A significance level of $\alpha = 0.05$ is selected.

$P_1 - P_2 = 0.8 - 0.6 = 0.2$

$$P = \frac{200x0.8 + 200x\ 0.6}{200+200}$$

$S_{P1-P2\ (\_high\ bars)} = \sqrt{0.7x0.3\ (1/200 + 1/200)} = 0.04583$

$$Z = \frac{0.2}{0.04583} = 4.36$$

Verónica Rosendo Ríos ©

✓ Two Independent Samples

Proportions

Comparison of proportions of jean users in USA and Hong Kong

	Users	Non users	Row Totals
United States	160	40	200
Hong Kong	120	80	200
Column Totals	280	120	

A z test is used as in testing the proportion for one sample. In this case, however, the test statistic is given by:

$$Z = \frac{0.2}{0.04583} = 4.36$$

Given a two-tailed test, the area to the right of the critical value is $\alpha/2$, or 0.025. Hence, the critical value of the test statistic is z=1.96. Since the calculated value of 4.36> 1.96, the null hypothesis is rejected.

Conclusion: The proportion of users (0.80 for the United States, and 0.60 for Hong Kong) is significantly different for the two samples.

Verónica Rosendo Ríos ©

✓ Paired Samples

The observations for the two groups are not selected from independent samples.

Rather, the information relate to the same respondents.

E.g.: A sample of respondents may rate two competing brands, may indicate the relative importance of two attributes of a product, or may evaluate a brand at two different times.

The difference in these cases is examined by a paired samples t test.

To compute t for paired samples, the paired difference variable, denoted by D, is formed and its mean and variance calculated.

Verónica Rosendo Ríos ©

✓ Paired Samples

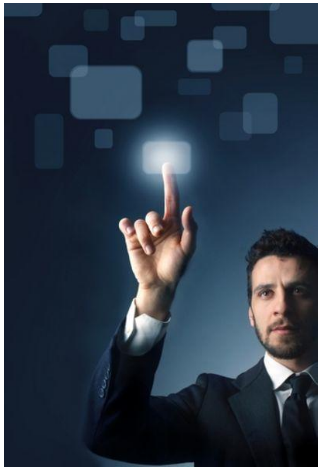
To compute t for paired samples, the paired difference variable, denoted by D, is formed and its mean and variance calculated.

Then the t statistic is computed. The degrees of freedom are n-1, where n is the number of pairs.

The hypotheses are:

$$H_0: \mu_D = 0$$

$$H_1: \mu_D \neq 0$$



Verónica Rosendo Ríos ©

✓ Paired Samples

E.g.: in the case of the department store, a paired t test could be used to determine if the respondents attached more importance to the quality of merchandise than to store credit and billing policies.

Variable	Number of cases	Mean	Standard Deviation	Standard Error
Quality	271	5.6273	0.670	0.041
Store credit	271	4.1882	1.619	0.098

(Difference) Mean	Standard Deviation	Standard Error	Correlation	Two-tailed Probability	t value	Degrees of Freedom	Two-tail Probability
1.431	1.6087	0.0977	0.222	0.000	14.73	270	0.000

The mean difference between the variables is 1.431, with a standard deviation of 1.6087 and a standard error of 0.0977. This results in a t value of (1.4391/0.0977)=14.73, with 271-1=270 degrees of freedom and a probability of less than 0.001.

Conclusion: the quality of merchandise is more important than store credit and billing policies in selecting a department store.

Verónica Rosendo Ríos ©

★ Non-Parametric Tests: Means and Proportions

Non-Parametric Statistics

Non-parametric tests are used **when the variables are nonmetric**. Like parametric tests, nonparametric tests are available for testing variables from one sample, two independent samples, or two related samples.

✓ One Samples

1. The Kolmogorov Smirnov (K-S) one sample test. The decision to reject the null hypothesis is based on the value of K. The larger K is, the more confidence we have that H_0 is false.

For example, in the context of the store department, imagine that one wanted to test whether the distribution of the importance attached to the store credit and billing policy was normal. A K-S one-sample test is conducted, obtaining the data shown in the following table:

Verónica Rosendo Ríos ©

★ Non-Parametric Tests: Means and Proportions

✓ One Samples

1. The Kolmogorov Smirnov (K-S) one sample test.

For example, in the context of the store department, imagine that one wanted to test whether the distribution of the importance attached to the store credit and billing policy was normal. A K-S one-sample test is conducted, obtaining the data shown in the following table:

	Test Distribution, Normal
Mean:	4.19
Standard Deviation:	1.62
Cases:	271

Most Extreme Differences				
Absolute	Positive	Negative	K-S z	Two-tailed p
0.19754	0.13150	-0.19754	3.252	0.000

K=0.1975. The critical value for K is 1.36/271=0.005.

Two ways of checking:

1. Since the calculated value of K (0.1975) >than the critical value (0.005), the null hypothesis is rejected.
2. Alternatively, the previous table indicates that the probability of observing a K value of 0.1975, is less than 0.001 (p=0.000). Since this is less than the significance level of 0.05, the null hypothesis is rejected.

Hence, the <u>distribution of the importance attached to store credit and billing policies variable deviates significantly from the normal distribution</u>.

Verónica Rosendo Ríos ©

★ Non-Parametric Tests: Means and Proportions

✓ One Samples

2. The chi-square test can also be performed on a single variable from one sample. In this context, the chi-square serves as a goodness of fit test. It tests whether a significant difference exists between the observed number of cases in each category and the expected number.

3. Other one-sample nonparametric tests include the run test and the binomial test.

- The runs test is a test of randomness for the dichotomous variables. This test is conducted by determining whether the order or sequence in which observations are obtained is random.
- The binomial test is also a goodness of fit test for dichotomous variables. It tests the goodness of fit of the observed number of observations in each category to the number expected under a specified binomial distribution.

Verónica Rosendo Ríos ©

★ Non-Parametric Tests: Means and Proportions

✓ Two Independent Samples

1. The Mann-Whitney U test. When the difference in the location of two populations is to be compared based on observations from two independent samples and the variable is measured on an *ordinal scale*, the Mann-Whitney U test can be used.

2. The two sample median test determinates whether the two groups are drawn from populations with the same median. It is not as powerful as the Mann-Whitney U test because it merely uses the location of each observation relative to the median, and not the rank, of each observation.

3. The Kolmogorov-Smirnov two-sample test examines whether the two distributions are the same. It takes into account any differences between the two distributions, including the median, dispersion, and skewness.

Verónica Rosendo Ríos ©

★ Non-Parametric Tests: Means and Proportions

✓ Paired Samples

1. Wilcoxon matched-pairs signed-ranks test.

This test analyzes the differences between the paired observations, taking into account the magnitude of the differences

For example, let´s see whether the respondents attached more importance to the quality of merchandise or to store credit and billing policies. Suppose that we assume that both these variables are measured on ordinal rather than interval scales

	Quality with Store Credit	
Mean Rank	Cases	(Store Credit Quality)
99.88	177	• Ranks
46.89	14	+ Ranks
	80	Ties
	271	Total
z=-11.1262		2-tailed p= 0.0000

◎The probability associated with the z statistic is less than 0.05, indicating that the difference is indeed significant.

Verónica Rosendo Ríos ©

★ Non-Parametric Tests: Means and Proportions

✓ Paired Samples

1. Another paired sample nonparametric test is the sign test. This test is not as powerful as the Wilcoxon matched-pairs signed-ranks test because it only compares the signs of the differences between pairs of variables without taking into account the magnitude of the differences.
2. In the special case of binary variables where the researcher whishes to test differences in proportions, the McNemar test can be used.
3. Alternatively, the chi-square test can also be used for binary variables.

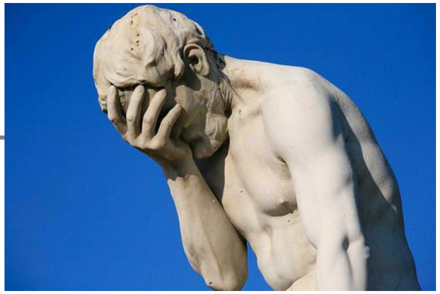
Verónica Rosendo Ríos ©

The various parametric and nonparametric tests are summarized in:

Sample	Application	Level of Scaling	Test/Comments
One Sample			
One Sample	Distributions	Nonmetric	K-S and chi-square for goodness of fit Runs test for randomness Binomial test for goodness of fit for dichotomous variables
One Sample	Means	Metric	t test, if variance is unknown z test, if variance is known
One Sample	Proportions	Metric	z test
Two independent samples			
Two independent samples	Distributions	Nonmetric	K-S two-sample test for examining the equivalence of two distributions
Two independent samples	Means	Metric	Two-group test F test for equality of variances
Two independent samples	Proportions	Metric Nonmetric	Z test Chi-square test
Two independent samples	Rankings/medians	Nonmetric	Mann-Whitney U test is more powerful than the median test
Paired Samples			
Paired Samples	Means	Metric	Paired t test
Paired Samples	Proportions	Nonmetric	McNemar test for binary variables Chi-square test
Paired Samples	Rankings/medians	Nonmetric	Wilcoxon matched-pairs ranked-signs test is more powerful tan the sign test

Verónica Rosendo Ríos ©

EXAMPLE



"Well, Are They Safisfied or Not?"

Ed Bond had worked for PrecisionMetals for six years, but had really only served as an analyst for the production facility. This was the first corporate-level opportunity to showcase his research skills. His corporate contacts are Rob Baer, who currently serves as Chief Operations Officer, and Kathy Hahn, the Chief Executive Officer for PrecisionMetals. Rob and Kathy specifically asked to meet with Ed about the employee satisfaction survey conducted a month ago.

"Ed, we continue to worry about losing metalwork employees at our Madison plant, but our Richmond plant seems to be improving in terms of turnover" Rob stated. "What is your take on our employee satisfaction?" Ed replied "**We put together an index of three questions that asked about job satisfaction. We have analyzed the data from the Richmond plant, and our average satisfaction is 3.9**"

Kathy asked "**What does 3.9 mean**? **What am I supposed to take away from that**? "

Verónica Rosendo Ríos ©

EXAMPLE

"Well, Are They Safisfied or Not?"

Ed replied, "I am sorry, I should have explained this better. We asked our employees on a scale with five categories, with 1 meaning 'Strongly Disagree' and 5 meaning 'Strongly Agree'. When the scores are averaged for Richmond, our overall satisfaction is 3.9 on this 5.0 point scale."

Rob continued, "**Is that good or bad**? It sounds OK… I guess. What about Madison?" Ed, realizing that he was not communicating the information well, responded, "The satisfaction score from Madison is 3.5. Historically, both plants had a satisfaction score of 3.5"

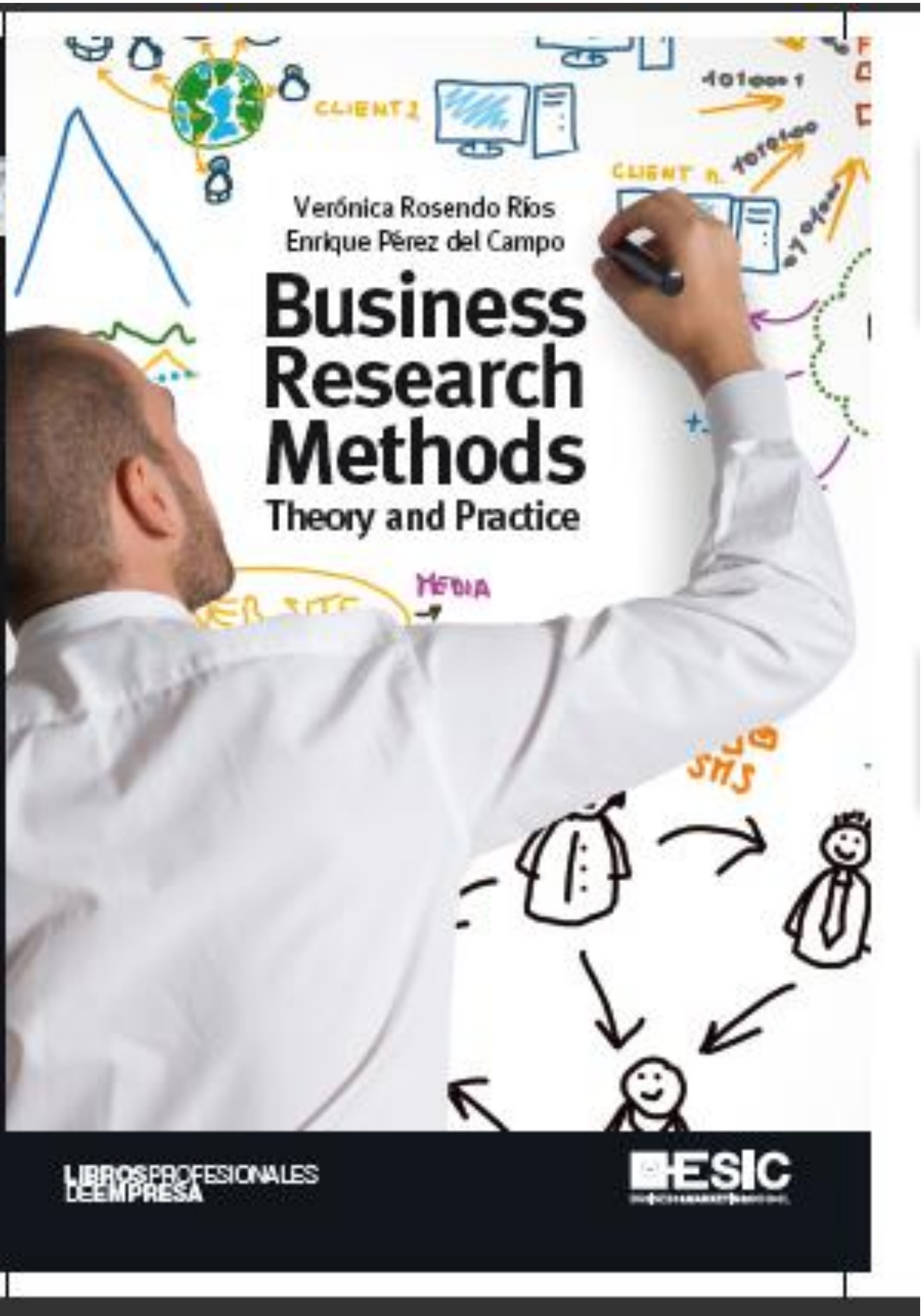
Kathy realized that Ed was getting flustered. It was time to reassure him. "Ed, we appreciate what you are doing. I'm sorry but I don't know exactly what the scores mean. **Is 3.9 good? Is 3.5 good? Is the difference between Richmond this year and the scores we have seen in the past significant? Is the difference between those scores enough to explain the difference in our turnover? I just want to know if the survey shows if our employees are satisfied or not**". Ed went back to the research section, with two things on his mind. "I have to actually compare the Richmond satisfaction score with the old scores", he thought. And as he walked into his office and shut the door quietly, he said to himself "**I just cannot speak about scores of 3.9 and 3.5. I'm here to help them understand what the scores really mean**". It was time to get to work!

Verónica Rosendo Ríos ©

REFERENCES

- Brown, T.J., Suter, T. (2012). *MR*. South Western, Cenage Learning. USA.
- Hair, J.; Bush, R. & Ortinau, D. (2006). *Marketing research. Within a changing environment*. Revised International Edition (3rd Ed.). McGraw Hill, New York, USA.
- Malhotra, N.K. (1996) Marketing Research. An Applied Orientation. 2nd Edn. Prentice-Hall International. USA
- Rosendo-Rios, V., de Esteban, J., Antonovica, A. (2012). MR: Development ofTheoretical Concepts for Market Research I and II. South Western, Cenage Learning. USA
- Zikmund, W. G., Babin, B.J., Carr, J.C., Griffin, M.(2013) Business Research Methods. 9th Edition. South Western, Cenage Learning. USA

Verónica Rosendo Ríos ©



Chapter 13

Hypothesis Testing

Business Research Methods
Verónica Rosendo Ríos
Enrique Pérez del Campo