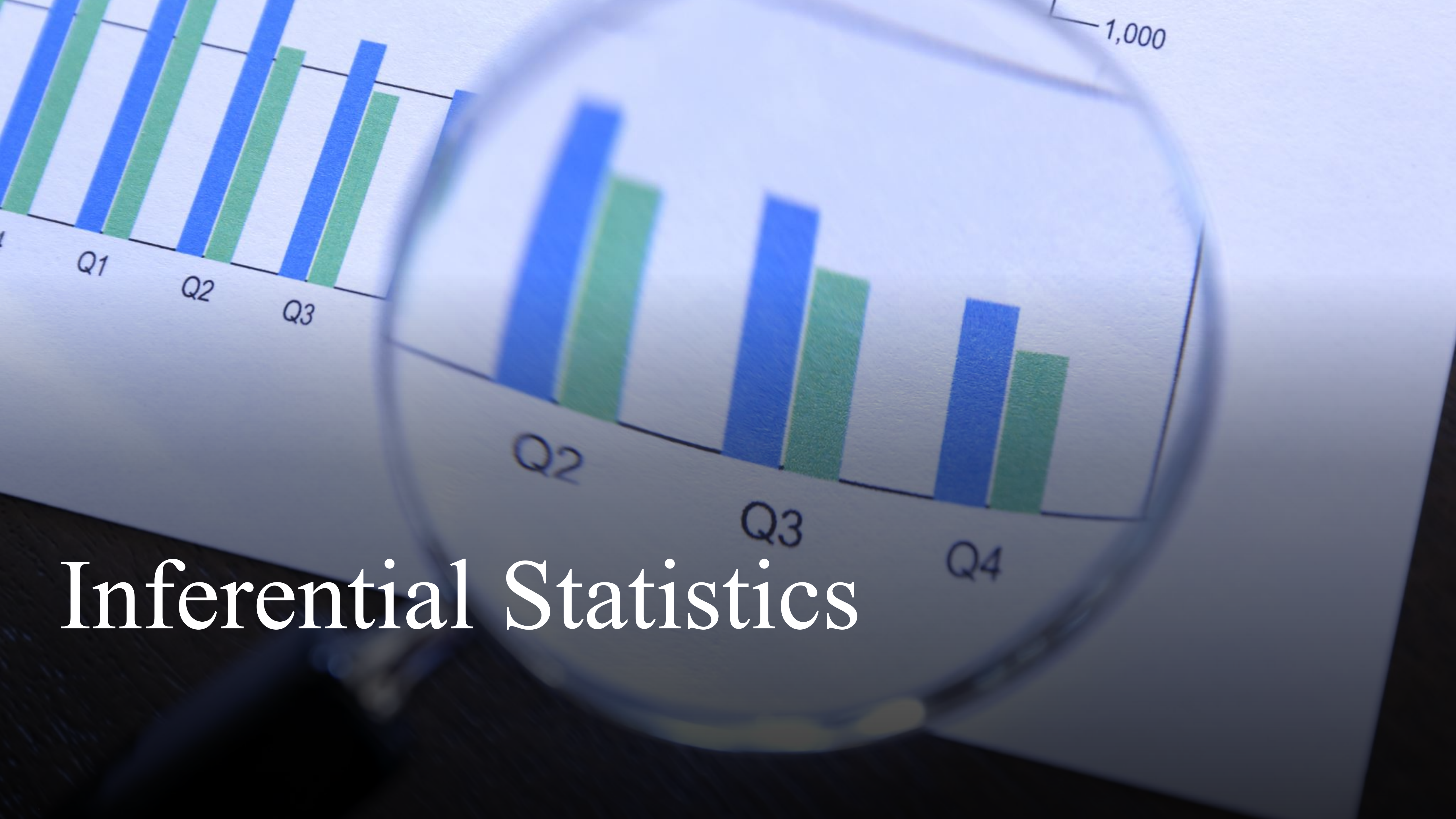


Inferential Statistics

Understanding Inferential Statistics

There are two main areas of inferential statistics:

- Estimating parameters. This means taking a statistic from your sample data (for example the sample mean) and using it to say something about a population parameter (i.e. the population mean).
- Hypothesis tests. This is where you can use sample data to answer research questions. For example, you might be interested in knowing if a new cancer drug is effective. Or if breakfast helps children perform better in schools.

Descriptive and Inferential Statistics

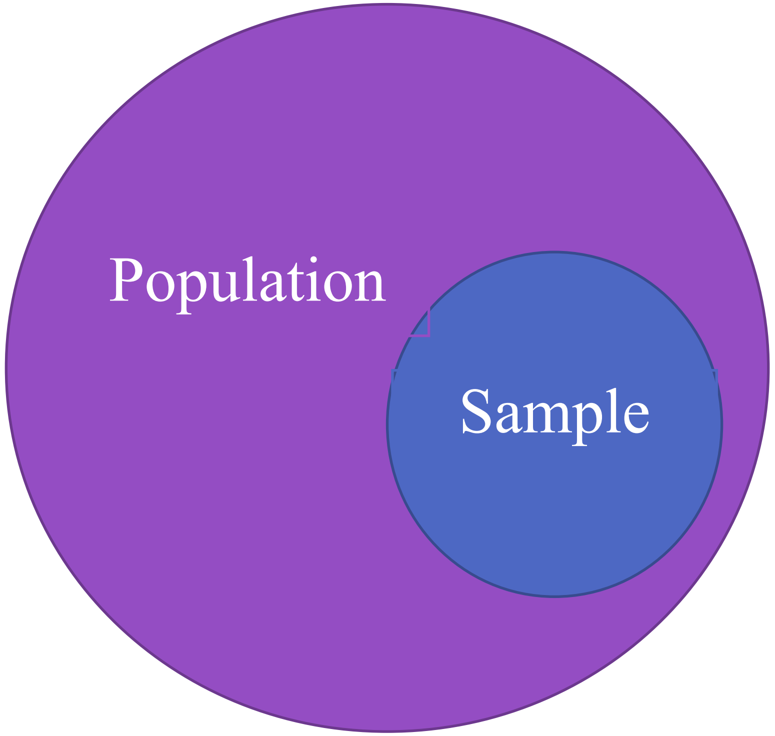
BASIS FOR COMPARISON	DESCRIPTIVE STATISTICS	INFERENTIAL STATISTICS
Meaning	Descriptive Statistics is that branch of statistics which is concerned with describing the population under study.	Inferential Statistics is a type of statistics, that focuses on drawing conclusions about the population, on the basis of sample analysis and observation.
What it does?	Organize, analyze and present data in a meaningful way.	Compares, test and predicts data.
Form of final Result	Charts, Graphs and Tables	Probability
Usage	To describe a situation.	To explain the chances of occurrence of an event.
Function	It explains the data, which is already known, to summarize sample.	It attempts to reach the conclusion to learn about the population, that extends beyond the data available.

What is sampling variability?

- Sampling variability is how much an estimate varies between samples. “Variability” is another name for range; Variability between samples indicates the *range of values* differs between samples.
- Sampling variability is the extent to which the measures of a sample differ from the measure of the population. The variance ($\sigma 2$) and standard deviation (σ) are common measures of variability.

Variability and Sample Sizes

- Increasing or decreasing sample sizes leads to changes in the variability of samples. For example, a sample size of 10 people taken from the same population of 1,000 will very likely give you a very different result than a sample size of 100.
- There is no "perfect" sample size that will give you accurate estimates for the sample mean, variance and other statistics. Instead, you take your best "guess" — using standardized statistical procedures. In general, estimates will change from sample to sample and will probably never exactly match the population parameter.



Central Limit Theorem

- The central limit theorem (CLT) states that the distribution of sample means approximates a normal distribution as the sample size gets larger.
- Sample sizes equal to or greater than 30 are considered sufficient for the CLT to hold.
- A key aspect of CLT is that the average of the sample means, and standard deviations will equal the population mean and standard deviation.
- A sufficiently large sample size can predict the characteristics of a population accurately.

Central Limit Theorem

- Central limit theorem is applicable for a sufficiently large sample sizes ($n \geq 30$). The formula for central limit theorem can be stated as follows:

$$\mu_{\bar{x}} = \mu \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Where,

μ = Population mean

σ = Population standard deviation

$\mu_{\bar{x}}$ = Sample mean

$\sigma_{\bar{x}}$ = Sample standard deviation

n = Sample size

Central Limit Theorem Example

- Q The record of weights of male population follows normal distribution. Its mean and standard deviation are 70 kg and 15 kg respectively. If a researcher considers the records of 50 males, then what would be the mean and standard deviation of the chosen sample?

Ans.

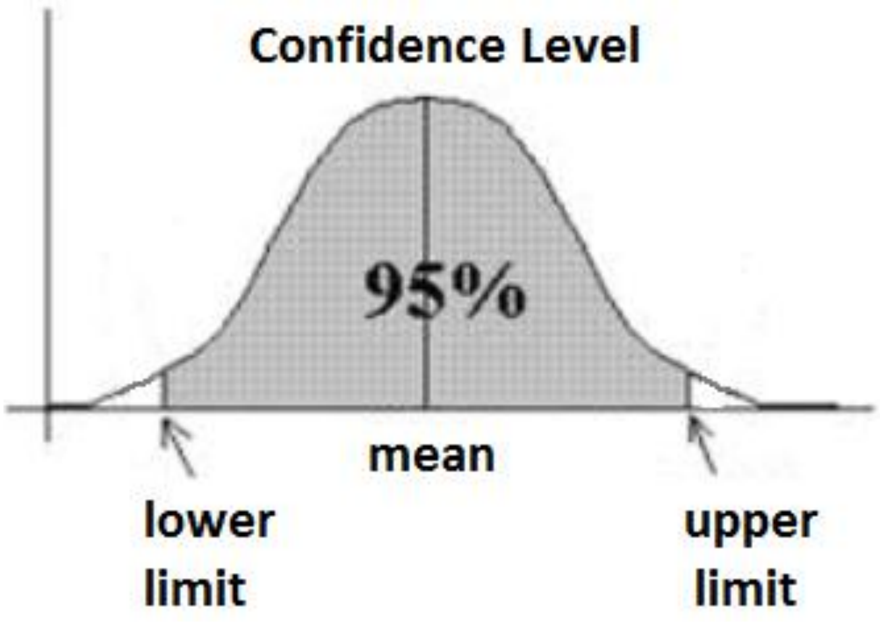
Mean of the population μ = 70 kg
Standard deviation of the population = 15 kg
sample size n = 50

Mean of the sample is given by:
$\mu_{\bar{x}} = 70$ kg

Standard deviation of the sample is given by:
$\sigma_{\bar{x}} = \sigma/\sqrt{n}$
$\sigma_{\bar{x}} = 15/\sqrt{50}$
$\sigma_{\bar{x}} = 2.1$ kg

Confidence Interval

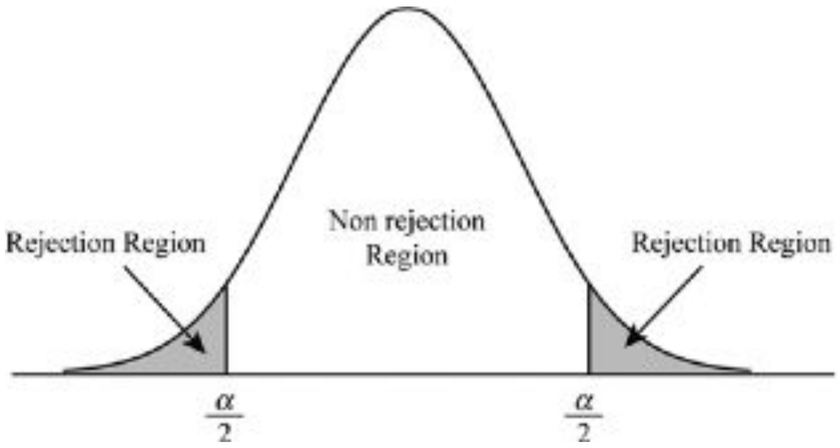
A confidence interval is how much uncertainty there is with any statistic. Confidence intervals are often used with a margin of error. Confidence intervals are intrinsically connected to confidence levels. Confidence levels are expressed as a percentage (for example, a 95% confidence level).



Hypothesis Testing

The process of hypothesis testing is to draw inferences or some conclusion about the overall population or data by conducting some statistical tests on a sample.

Hypothesis Testing



- Setting up a hypothesis
- Setting up a suitable level of significance
- Setting up a test criterion
- Performing calculations and making decisions

Statistical Hypothesis

The best way to determine whether a statistical hypothesis is true would be to examine the entire population. Since that is often impractical, researchers typically examine a random sample from the population. If sample data are not consistent with the statistical hypothesis, the hypothesis is rejected.

There are two types of statistical hypotheses:

- **Null Hypothesis**. The null hypothesis, denoted by H_o, is usually the hypothesis that sample observations result purely from chance.
- **Alternative Hypothesis**. The alternative hypothesis, denoted by H_1, is the hypothesis that sample observations are influenced by some non-random cause.

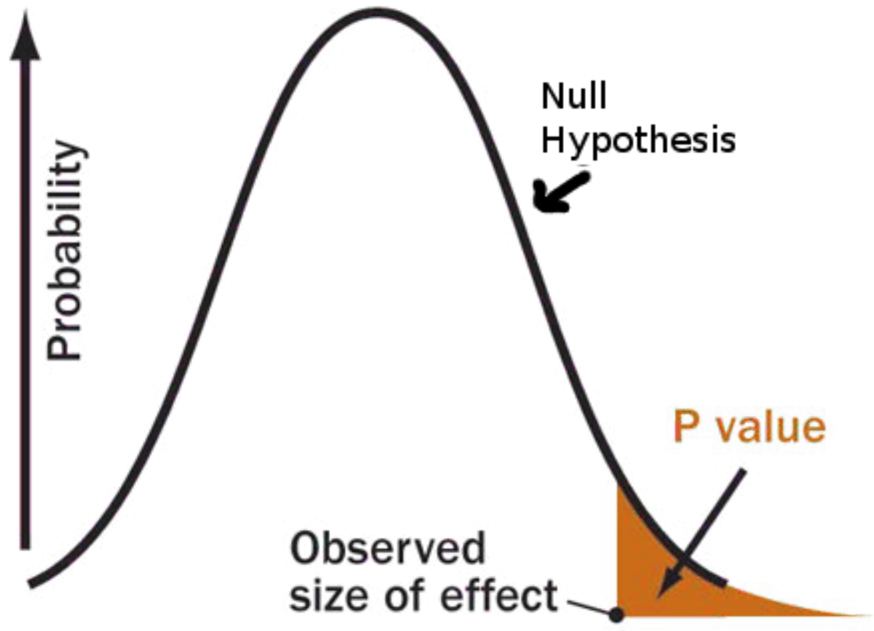
Hypothesis Testing Process

- The first step is to state the relevant null and alternative hypotheses. This is important, as mis-stating the hypotheses will muddy the rest of the process.
- The second step is to consider the statistical assumptions being made about the sample in doing the test; for example, assumptions about the statistical independence or about the form of the distributions of the observations. This is equally important as invalid assumptions will mean that the results of the test are invalid.
- Decide which test is appropriate, and state the relevant test statistic T.
- Derive the distribution of the test statistic under the null hypothesis from the assumptions. In standard cases this will be a well-known result. For example, the test statistic might follow a Student's t distribution or a normal distribution.
- Select a significance level (α), a probability threshold below which the null hypothesis will be rejected. Common values are 5% and 1%.
- The distribution of the test statistic under the null hypothesis partitions the possible values of T into those for which the null hypothesis is rejected—the so-called critical region—and those for which it is not. The probability of the critical region is α.
- Compute from the observations the observed value tobs of the test statistic T.
- Decide to either reject the null hypothesis in favour of the alternative or not reject it. The decision rule is to reject the null hypothesis H0 if the observed value tobs is in the critical region, and to accept or "fail to reject" the hypothesis otherwise.

p-value

The *p*-value is the probability that a given result (or a more significant result) would occur under the null hypothesis.

For example, say that a fair coin is tested for fairness (the null hypothesis). At a significance level of 0.05, the fair coin would be expected to (incorrectly) reject the null hypothesis in about 1 out of every 20 tests. The *p*-value does not provide the probability that either hypothesis is correct (a common source of confusion).



Example: Courtroom Trial

A statistical test procedure is comparable to a criminal trial; a defendant is considered not guilty if his or her guilt is not proven. The prosecutor tries to prove the guilt of the defendant. Only when there is enough evidence for the prosecution is the defendant convicted.

H0: "the defendant is not guilty"

H1: "the defendant is guilty"

The hypothesis of innocence is rejected only when an error is very unlikely, because one doesn't want to convict an innocent defendant.

Table of error types		Null hypothesis (H_0) is	
		True	**False**
Decision about null hypothesis (H_0)	**Don't reject**	Correct inference (true negative) (probability = $1-\alpha$)	Type II error (false negative) (probability = β)
	Reject	Type I error (false positive) (probability = α)	Correct inference (true positive) (probability = $1-\beta$)

Parametric Tests

Parametric tests assume a normal distribution of values, or a "bell-shaped curve." These types of tests include Student's T tests and ANOVA tests, which assume data is from a normal distribution.

Z-test

In a z-test, the sample is assumed to be normally distributed. A z-score is calculated with population parameters such as "population mean" and "population standard deviation" and is used to validate a hypothesis that the sample drawn belongs to the same population.

The formula for Z – Test is given as:

$$z = \frac{x - \mu}{\sigma / \sqrt{n}}$$

When do we use Z-test?

- When samples are drawn at random.
- When the samples are taken from population are independent.
- When standard deviation is known.
- When number of observation is large ($n \geq 30$)

Question

A principal at a school claims that the students in his school are above average intelligence. A random sample of thirty students' IQ scores have a mean score of 112.5. Is there sufficient evidence to support the principal's claim? The mean population IQ is 100 with a standard deviation of 15.

Solution

- Step 1: State the Null hypothesis. The accepted fact is that the population mean is 100, so: H0: $\mu=100$.
- Step 2: State the Alternate Hypothesis. The claim is that the students have above average IQ scores, so: H1: $\mu > 100$.
- Step 3: Draw a picture to help you visualize the problem.
- Step 4: State the alpha level. If you aren't given an alpha level, use 0.05, An alpha level of 0.05 is equal to a z-score of 1.645. To calculate z-score use TI-83 calculator.
- Step 5: Find the Z using this formula: For this set of data: $Z= (112.5-100) / (15/\sqrt{30}) = 4.56$.
- Step 6: If Step 5 (4.56) is greater than Step 4 (1.645), reject the null hypothesis. If it's less than Step 4, you cannot reject the null hypothesis. In this case, it is greater, so you can reject the null and principal's claim is right.

T-test

A t-test is used to compare the mean of two given samples. Like a z-test, a t-test also assumes a normal distribution of the sample. A t-test is used when the population parameters (mean and standard deviation) are not known.

Formula of t test:

$$t = \frac{m - \mu}{s/\sqrt{n}}$$

When do we use T-test

A t-test can only be used when comparing the means of two groups (a.k.a. pairwise comparison). If you want to compare more than two groups, or if you want to do multiple pairwise comparisons, use an ANOVA test.

The t-test is a parametric test of difference, meaning that it makes the same assumptions about your data as other parametric tests. The t-test assumes your data:

- are independent
- are (approximately) normally distributed.
- have a similar amount of variance within each group being compared (a.k.a. homogeneity of variance)

What type of t-test should I use?

When choosing a t-test, you will need to consider two things: whether the groups being compared come from a single population or two different populations, and whether you want to test the difference in a specific direction.

One-sample, two-sample, or paired t-test?

01

If there is one group being compared against a standard value (e.g. comparing the acidity of a liquid to a neutral pH of 7), perform a **one-sample t-test**.

02

If the groups come from two different populations (e.g. two different species, or people from two separate cities), perform a **two-sample t-test** (a.k.a. **independent t-test**).

03

If the groups come from a single population (e.g. measuring before and after an experimental treatment), perform a **paired t-test**.

One-tailed or two-tailed t-test?

If you only care whether the two populations are different from one another, perform a **two-tailed t-test**.

If you want to know whether one population mean is greater than or less than the other, perform a **one-tailed t-test.**

ANOVA

ANOVA, also known as analysis of variance, is used to compare multiple samples with a single test.

- One-way ANOVA between groups: used when you want to test two groups to see if there’s a difference between them.
- Two-way ANOVA without replication: used when you have one group and you’re double-testing that same group. For example, you’re testing one set of individuals before and after they take a medication to see if it works or not.
- Two-way ANOVA with replication: Two groups, and the members of those groups are doing more than one thing. For example, two groups of patients from different hospitals trying two different therapies.

ANOVA

The statistics used to measure the significance, in this case, is called F-statistics. The F value is calculated using the formula,

$$F= ((SSE1 — SSE2)/m)/ SSE2/n\text{-}k$$

where,

- SSE = residual sum of squares
- m = number of restrictions
- k = number of independent variables

Assumptions of ANOVA

- **Independence of observations**: the data were collected using statistically-valid methods, and there are no hidden relationships among observations. If your data fail to meet this assumption because you have a confounding variable that you need to control for statistically, use an ANOVA with blocking variables.
- **Normally-distributed response variable**: The values of the dependent variable follow a normal distribution.
- **Homogeneity of variance**: The variation within each group being compared is similar for every group. If the variances are different among the groups, then ANOVA probably isn't the right fit for the data.

ANOVA vs T-test

A Student's T-test will tell you if there is a significant variation between groups. A t-test compares means, while the ANOVA compares variances between populations.

You could technically perform a series of t-tests on your data.

However, as the groups grow in number, you may end up with a lot of pair comparisons that you need to run. ANOVA will give you a single number (the f-static) and one p-value to help you support or reject the null hypothesis.

Nonparametric Tests

Nonparametric tests don’t require that your data follow the normal distribution. They’re also known as distribution-free tests and can provide benefits in certain situation. Nonparametric tests include Chi-Square.

Chi-Square Test

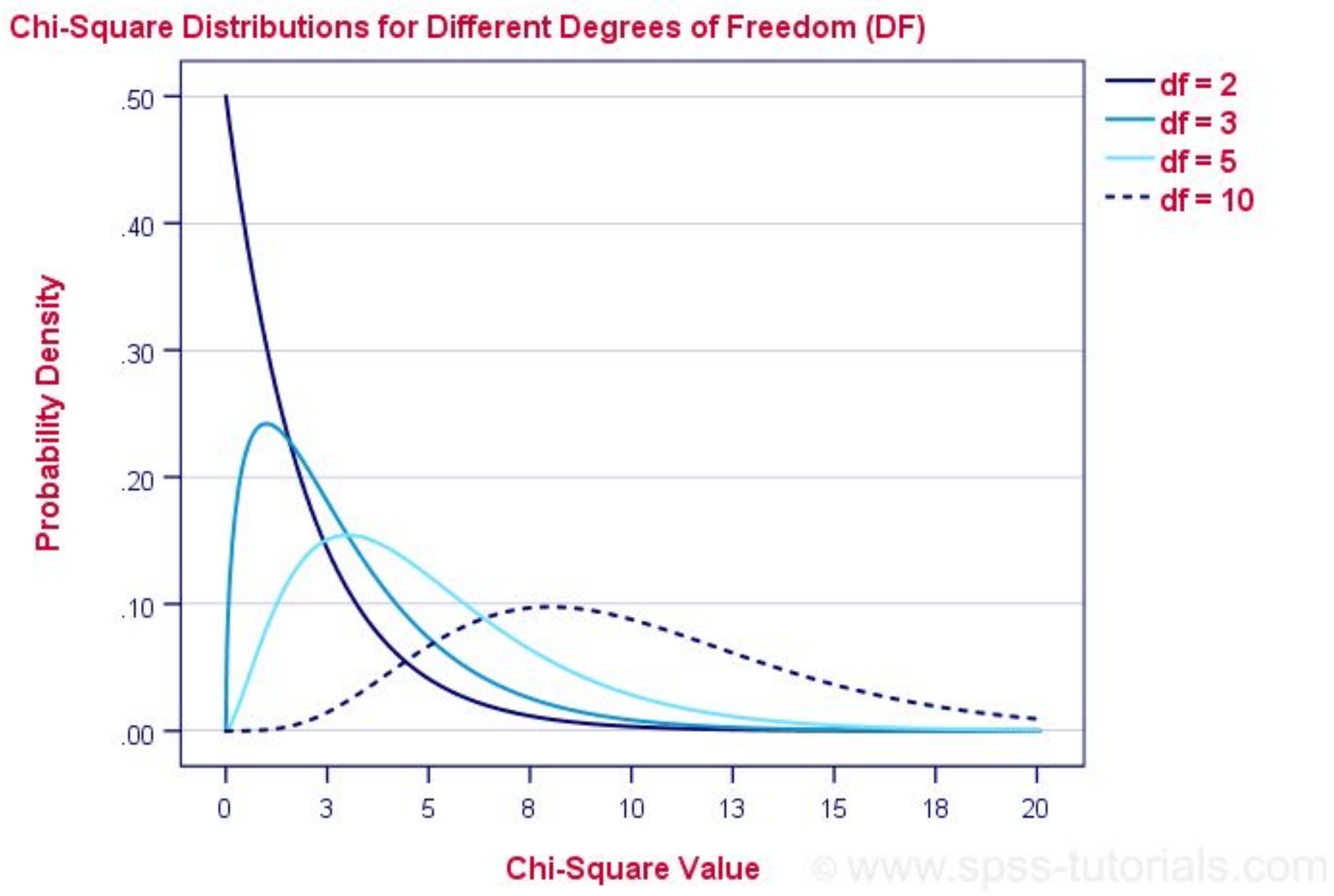
$$\chi_c^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Chi-square test is used to compare categorical variables. There are two types of chi- square tests. Both use the chi-square statistic and distribution for different purposes:

- A chi-square goodness of fit test determines if a sample data matches a population. For more details on this type, see: Goodness of Fit Test.
- A chi-square test for independence compares two variables in a contingency table to see if they are related. In a more general sense, it tests to see whether distributions of categorical variables differ from each another.
- A very small chi square test statistic means that your observed data fits your expected data extremely well. In other words, there is a relationship.
- A very large chi square test statistic means that the data does not fit very well. In other words, there isn't a relationship.
- The formula for the chi-square statistic used in the chi square test is:

Chi-Square Test - Degrees of Freedom

The degrees of freedom is basically a number that determines the exact shape of our distribution. The figure below illustrates this point.



Uses of Chi-Square Test

- Confidence Interval estimation for a population standard deviation of a normal distribution from a sample standard deviation.
- Independence of two criteria of classification of qualitative variables.
- Relationships between categorical variables (contingency tables).
- Sample variance study when the underlying distribution is normal.
- Tests of deviations of differences between expected and observed frequencies (one-way tables).
- The chi-square test (a goodness of fit test).

Thank You