Statistical Tests

Statistics are the arrangement of statistical tests which analysts use to make inference from the data given. These tests enables us **to make decisions on the basis of observed pattern from data**. There is a wide range of statistical tests. The choice of which statistical test to utilize relies upon the **structure of data**, **the distribution of the data, and variable type**. There are many different types of tests in statistics like t-test, Z-test, chi-square test, ANOVA test , binomial test, one sample median test etc.

Statistical tests are used in **hypothesis testing.** *Hypothesis testing is an objective method of making decisions or inferences from sample data (evidence). Sample data is used to choose between two choices i.e. hypotheses or statements about a population. Typically this is carried out by comparing what we have observed to what we expected if one of the statements (Null Hypothesis) was true;*

- **NULL HYPOTHESIS (H0): is a statement about the population & sample data used to decide whether to reject that statement or not. Typically the statement is that there is no difference between groups or association between variables.**
- **ALTERNATIVE HYPOTHESIS (H1): is often the research question and varies depending on whether the test is one or two tailed.**

They can be used to

- Determine whether a predictor variable has a statistically significant relationship with an outcome variable.
- Estimate the difference between two or more groups.

Statistical tests assume a **null hypothesis** of no **relationship** or **no difference between groups**. Then they determine whether the observed data fall outside of the range of values predicted by the null hypothesis.
If you already know what types of variables you're dealing with, you can use the flowchart to choose the right statistical test for your data.

A step-by-step guide to hypothesis testing

Hypothesis testing is a formal procedure for investigating our ideas about the world using statistics. It is most often used by scientists to test specific predictions, called hypotheses, that arise from theories. There are 5 main steps in hypothesis testing:

1. State your research hypothesis as a null (H_o) and alternate (H_a) hypothesis.
2. Collect data in a way designed to test the hypothesis.
3. Perform an appropriate statistical test.
4. Decide whether the null hypothesis is supported or refuted.
5. Present the findings in your results and discussion section.

Statistical tests work by calculating a test statistic – a number that describes how much the relationship between variables in your test differs from the null hypothesis of no relationship.

p-value (probability value). The p-value estimates how likely it is that you would see the difference described by the test statistic if the null hypothesis of no relationship were true.

- If the value of the test statistic **is more extreme than** the statistic calculated from the null hypothesis, then you can infer a statistically significant relationship between the predictor and outcome variables.
- If the value of the test statistic **is less extreme than** the one calculated from the null hypothesis, then you can infer no statistically significant relationship between the predictor and outcome variables.

Perform A Statistical Test

To perform statistical tests on data that have been collected in a statistically valid manner – either through an experiment, or through observations made using probability sampling methods. For a **statistical test to be valid**, your sample size **needs to be large enough** to approximate the true distribution of the population being studied. To **determine which statistical test to use**, you need to know:

- Whether your data meets certain assumptions.
- The types of variables that you're dealing with.

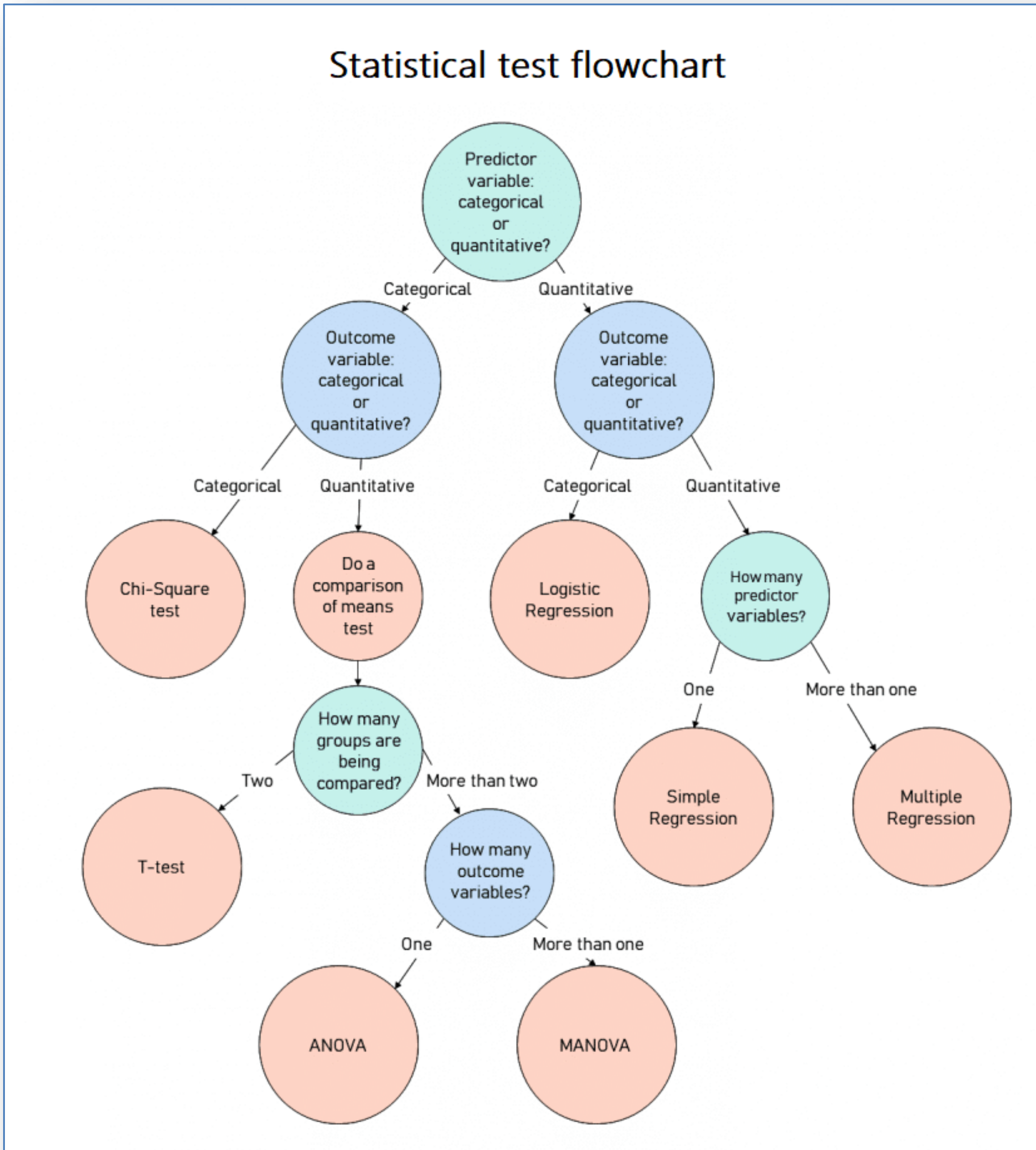
Interpreting Test Statistics

For any combination of sample sizes and number of predictor variables, a statistical test will produce a predicted distribution for the test statistic. This shows the most likely range of values that will occur if your data follows the null hypothesis of the statistical test.

The more extreme your test statistic the further to the edge of the range of predicted test values it is the less likely it is that your data could have been generated under the null hypothesis of that statistical test.

The agreement between your calculated test statistic and the predicted values is described by the p-value. The smaller the p-value, the less likely your test statistic is to have occurred under the null hypothesis of the statistical test.

Because the test statistic is generated from your observed data, this ultimately means that the smaller the p-value, the less likely it is that your data could have occurred if the null hypothesis was true.

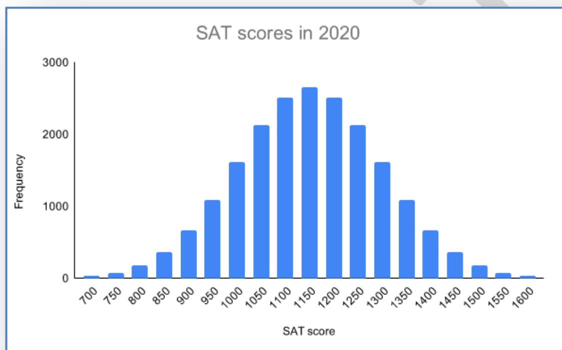


Statistical assumptions جدا مهمة كخطوة اساسية لبدأ التحليل الاحصائي

Statistical tests make some common assumptions about the data they are testing:

1. **Determine type of variable:** the types of variables you have usually **determine what type of statistical test** you can use

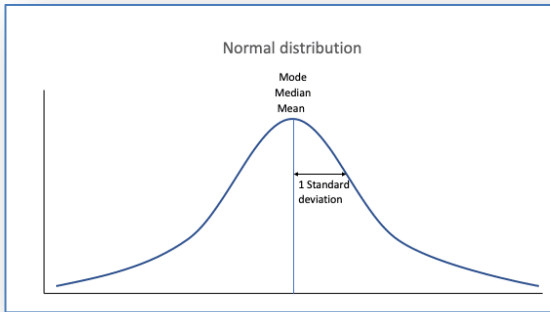
2. **Independence of observations** (a.k.a. no autocorrelation): The observations/variables you include in your test are not related (for example, multiple measurements of a single test subject are not independent, while measurements of multiple different test subjects are independent).
3. Homogeneity of variance: **the variance(is a measure of variability)** within each group being compared is similar among all groups. If one group has much more variation than others, it will limit the test's effectiveness.
4. Normality of data: the data follows **a normal distribution** (a.k.a. a bell curve). In a **normal distribution**, data is symmetrically distributed **with no skew**. When plotted on a graph, the data follows **a bell shape**, with most values clustering around a central region and tapering off as they go further away from the center as show in figure bellow.
   Normal distributions are also called **Gaussian distributions** or bell curves because of their shape. This assumption applies only to quantitative data.



The properties of normal distributions

Normal distributions have key characteristics that are easy to spot in graphs:

- The mean, median and mode are exactly the same.
- The distribution is symmetric about the mean—half the values fall below the mean and half above the mean.
- The distribution can be described by two values: the **mean and the standard deviation** as show in figure bellow;



If your data do not meet the assumptions of normality or homogeneity of variance, you may be able to perform a **nonparametric statistical test**, which allows you to make **comparisons without any assumptions about the data distribution**.

❖ A Parametric And A Nonparametric Test

1. Parametric tests assume underlying statistical distributions in the data. Therefore, several conditions of validity must be met so that the result of a parametric test is reliable. For example, Student's t-test for two independent samples is reliable only if each sample follows a normal distribution and if sample variances are homogeneous.
   - ❖ Parametric analyses to assess group means.
   - ❖ Parametric tests can provide trustworthy results with distributions that are skewed and non-normal

Parametric analyses	Sample size requirements
1-sample t-test	Greater than 20
2-sample t-test	Each group should have more than 15 observations
One-Way ANOVA	○ For 2-9 groups, each group should have more than 15 observations ○ For 10-12 groups, each group should have more than 20 observations

2. Nonparametric tests do not rely on any distribution. They can thus be applied even if parametric conditions of validity are not met.
   - ❖ Nonparametric analyses to assess group medians
   - ❖ Nonparametric tests are valid when our sample size is small and your data are potentially nonnormal

- Nonparametric tests can analyze ordinal data, ranked data, and outliers.
  An outlier is an unusually large or small observation. Outliers can have a disproportionate **effect** on statistical results, such as the **mean,** which can result in **misleading interpretations.**

Parametric tests of means	Nonparametric tests of medians
1-Sample t-test	1-sample Sign, 1-sample Wilcoxon
2-Sample t-test	Mann-Whitney test
One-Way ANOVA	Kruskal-Wallis, Mood's median test

Types of Statistical Tests

In parametric statistical test; parametric tests are used when data is normally distributed; parametric tests include

1. **Z-test:**
   A z-test is a statistical test used to determine whether two population means are different when the variances are known and the sample size is large. In z-test mean of the population is compared. The parameters used are population mean and population standard deviation. Z-test is used to validate a hypothesis that the sample drawn belongs to the same population.

$$z = (x — \mu) / (\sigma / \sqrt{n}),$$

where , x=sample mean, u=population mean, $\sigma / \sqrt{n}$ = population standard deviation.
If z value is less than critical value accept null hypothesis else reject null hypothesis.

2. **T-test:** In t-test the mean of the two given samples are compared. A t-test is used when the population parameters (mean and standard deviation) are not known.
3. **Paired T-Test:** Tests for the difference between two variables from the same population(pre- and posttest score). For example- In a training program performance score of the trainee before and after completion of the program.
4. **Independent T-test**: The independent t-test which is also called the two sample t-test or student's t-test, is a statistical test that determines whether there is a statistically significant

difference between the means in two unrelated groups. For example -comparing boys and girls in a population.

5. **One sample t-test:** The mean of a single group is compared with a given mean. For example- to check the increase and decrease in sales if the average sales is given.

   $t = (x1 — x2) / (\sigma / \sqrt{n1} + \sigma / \sqrt{n2})$,

   where x1 and x2 are mean of sample 1 and sample 2 respectively.

6. ANOVA Test- Analysis of variance (ANOVA) is a statistical technique that is used to check if the means of two or more groups are significantly different from each other. ANOVA checks the impact of one or more factors by comparing the means of different samples. If we use a t-test instead of ANOVA test it won't be reliable as number of samples are more than two and it will give error in the result. The hypothesis being tested in ANOVA is

   Ho: All pairs of samples are same i.e. all sample means are equal
   Ha: At least one pair of samples is significantly different

   In anova test we calculate F value and compare it with critical value
   $F= ((SSE1 — SSE2)/m)/ SSE2/n-k$, where
   SSE = residual sum of squares
   m = number of restrictions
   k = number of independent variables

Non parametric statistical test: Non parametric tests are used when data is not normally distributed. Non parametric tests include chi-square test.

Chi-square test($\chi2$ test): chi-square test is used to compare two categorical variables. Calculating the Chi-Square statistic value and comparing it against a critical value from the Chi-Square distribution allows to assess whether the observed frequency are significantly different from the expected frequency. The hypothesis being tested for chi-square is-

Ho: Variable x and Variable y are independent
Ha: Variable x and Variable y are not independent.

$$\chi^2 = \sum \frac{(o-e)^2}{e}$$

where o=observed , e=expected.

Example 1:

Blood glucose levels for obese patients have a mean of 100 with a standard deviation of 15. A researcher thinks that a diet high in raw cornstarch will have a positive effect on blood glucose levels. A sample of 36 patients who have tried the raw cornstarch diet have a mean glucose level of 108. Test the hypothesis that the raw cornstarch had an effect or not.

Solution:- Follow the above discussed steps to test this hypothesis:

Step-1: State the hypotheses. The population mean is 100.

$$H0: \mu = 100$$
$$H1: \mu > 100$$

Step-2: Set up the significance level. It is not given in the problem so let's assume it as 5% (0.05).

Step-3: Compute the random chance probability using z score and z-table.

$$z = \frac{x - \mu}{\sigma}$$

$\mu =$ Mean
$\sigma =$ Standard Deviation

For this set of data: z= (108-100) / (15/√36)=3.20

You can look at the probability by looking at z- table and p-value associated with 3.20 is 0.9993 i.e. probability of having value less than 108 is 0.9993 and more than or equals to 108 is (1-0.9993)=0.0007.

Step-4: It is less than 0.05 so we will reject the Null hypothesis i.e. there is raw cornstarch effect.

Note: Setting significance level can also be done using z-value known as critical value. Find out the z- value of 5% probability and it is 1.65 (positive or negative, in any direction). Now we can compare calculated z-value with critical value to make a decision.

Example2

Templer and Tomeo (2002) reported that the population mean score on the quantitative portion of the Graduate Record Examination (GRE) General Test for students taking the exam between 1994 and 1997 was 558 ± 139 ($\mu \pm \sigma$). Suppose we select a sample of 100 participants (n = 100). We record a sample mean equal to 585 (M = 585). Compute the p-value t0 check whether or not we will retain the null hypothesis (μ = 558) at 0.05 level of significance (α = .05).

Solution:

Step-1: State the hypotheses. The population mean is 558.

H0: $\mu = 558$
H1: $\mu \neq 558$ (two tail test)

Step-2: Set up the significance level. As stated in the question, it as 5% (0.05). In a non-directional two-tailed test, we divide the alpha value in half so that an equal proportion of area is placed in the upper and lower tail. So, the significance level on either side is calculated as: $\alpha/2 = 0.025$. and z score associated with this (1-0.025=0.975) is 1.96. As this is a two-tailed test, z-score(observed) which is less than -1.96 or greater than 1.96 is a evidence to reject the Null hypothesis.

Step-3: Compute the random chance probability or z score

$$z = \frac{x - \mu}{\sigma}$$

$\mu =$ Mean
$\sigma =$ Standard Deviation

For this set of data: $z = (585-558) / (139/\sqrt{100}) = 1.94$

You can look at the probability by looking at z- table and p-value associated with 1.94 is 0.9738 i.e. probability of having value less than 585 is 0.9738 and more than or equals to 585 is (1-0.9738)=0.03

Step-4: Here, to make a decision, we compare the obtained z value to the critical values (+/- 1.96). We reject the null hypothesis if the obtained value exceeds a critical values. Here obtained value ($Z_{obt} = 1.94$) is less than the critical value. It does not fall in the rejection region. The decision is to retain the null hypothesis.