Making a choice of statistical test…

The steps in a Hypothesis Test on a population parameter:

1. Framing the Hypothesis: Identify the situation where the proposed value is correct. The **Null hypothesis** is a statement that the value is correct. The alternative hypothesis is a statement that it is incorrect.
2. Decide, from the alternative hypothesis, if the test is one- or two-**tailed**. This is based on whether the question you are investigating has a direction.
3. Decide on a **level of significance**; this means choosing a probability which is regarded as an appropriately low. If the test is two-tailed, the level of significance is divided by 2.

The difference between the estimate of the parameter obtained from the data and the proposed value under the null hypothesis is measured with a **statistic** with a known distribution.

4. Find a **critical value** of the distribution; this will often depend on a **degrees of freedom** parameter usually based on the size of the sample. For a two-tailed test, it will be two (linked) values.
5. The **sample** value found for the test statistic, in other words calculated from the data, is compared to the critical values; if it is found to be an unlikely value, the Null hypothesis is rejected and the alternative is accepted.

1. Testing a mean

Given a claim about a true mean μ, for example:

> 'A manufacturer is producing lengths of steel pipe, and claims the mean length is $\mu = 20$m' OR 'A food supplier is supplying breakfast cereal, and claims the mean weight of the packets is $\mu = 0.8$kg'.

Null Hypothesis:

The Null hypothesis is that the claim of the mean is true.

Test to use:

The equation and statistic change slightly as follows.

- If n > 50, then the z-test can be used with the *sample* standard deviation, denoted s, as shown. In practice, once n > 30, a statistician will actually use the z-test. The equation for the sample value is now:

$$z = \frac{\overline{X} - \mu}{s/\sqrt{n}}.$$

- If n < 30, and the *sample* standard deviation (denoted s) has to be calculated, then use a t-test, with n – 1 degrees of freedom. The equation is the same; it just follows a different distribution.

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}}.$$

- If the true standard deviation, denoted σ, is available, then the z-test is used, for any n, with σ in place of s. This is a purely theoretical possibility.

2. Testing a variance

Given a claim about variance, for example:

> 'A manufacturer is producing lengths of steel pipe, and claims the variance of the length is $\sigma^2 = 20\text{mm}^2$' OR 'A food supplier is supplying breakfast cereal, and claims the variance of the weight of the packets is $\sigma^2 = 8\text{kg}^2$'.

Null Hypothesis:

In this case, the Null hypothesis is that the claim of the variance is true.

Like the test for a mean, the test can be one or two tailed.

Test to use: The test uses the 'chi-squared' variable:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}, \text{ with } n - 1 \text{ degrees of freedom.}$$

Here the symbol σ^2 refers to the claim for the true variance and symbol s^2 refers to the sample variance. If you have calculated s^2, the top line of this term will be the same as the top line in the calculation for s^2.

The distribution for χ^2 is ***not symmetric***. The critical values found in the tables for a given probability α are for 'greater than'; in other words, if a is the critical value and α is the level of significance, then

$$P[\chi^2 > a\,] = \alpha.$$

This means the critical value is an upper bound. For the lower bound, we need a 'less than' event, in other words the value b such that:

$$P[\chi^2 < b\,] \text{ so use } P[\chi^2 > b\,] = 1 - \alpha.$$

These are the figures on the right-hand side of the tables, 0.95, 0.975 etc.

So for a two tailed test at 0.05, the values for 0.025 and 0.975 are needed.

3. Testing a difference

Given a claim about a before/after or with/without analysis, for example:

- 'A food producer claims that drinking a type of yogurt will lower cholesterol.'
- 'A fuel supplier claims that a fuel additive will increase the efficiency of an engine'.

In this case, there will be two sets of data for each person/unit/subject, 'with and without' or 'before and after'. Therefore, calculate the differences.

The test is then a test on the mean of the differences, and invariably the Null hypothesis is that there is no change, in other words the mean of the differences is 0.

The test is a *t*-test if the sample size is small (the same considerations apply as before), and the equation is:

$$t = \frac{\overline{D} - \mu_D}{s_D / \sqrt{n}} .$$

Here the variable D refers to the difference, and μ_D and s_D are the claimed mean and sample standard deviation of this variable.

4. Testing for a Correlation

Given corresponding values of two variables, (X_i, Y_i) investigate whether they are correlated; for example:
'Do people smoke fewer cigarettes (or more cigarettes) depending on age?' OR 'Does the resistance of a wire decrease with increasing temperature?'

Test to use:

The correlation coefficient r is given by two equations, the first for the raw data itself, the second if means, standard deviations and the cross sum are already available:

$$r = \frac{\sum X_i Y_i - n\overline{X}\,\overline{Y}}{\sqrt{\sum X_i^{\,2} - n\overline{X}^2}\sqrt{\sum Y_i^{\,2} - n\overline{Y}^2}} \quad \text{or} \quad r = \frac{\sum_i X_i Y_i - n\overline{X}\,\overline{Y}}{(n-1)s_X s_Y}.$$

Null Hypothesis:

In this case, the Null hypothesis is almost always that there is no correlation; in other words, the true correlation (denoted ρ) is 0; $\rho = 0$.

To carry out the test, a t-value is calculated from r, according to the equation

$$t = r\sqrt{\frac{n-2}{1-r^2}} \quad \text{or} \quad t = \sqrt{n-2}\,\frac{r}{\sqrt{1-r^2}}.$$

The test becomes a t-test, with n – 2 degrees of freedom. As before, there is the following caveat:

- If $n < 30$, then use a t-test.
- If $n > 50$, then we use the z-test.
- In practice, once $n > 30$, a statistician will actually use the z-test.

5. Test for Independence

In this instance, a link is being investigated between two categories or qualitative variables. There is **no parameter** to be tested. For example:

- ‘A lecture suspects mature students are performing better in exams.’
- ‘A Nursery plant grower suspects different combinations of fertilizer give different plant heights.’

These cases are quite like correlations, but concern categories rather than actual values of variables. The test invariably has a table of counts, such as:

Age grp.	*Merit*	*Pass*	*Fail*
A	21	26	21
B	15	16	11
C	17	12	10

Null Hypothesis:

The Null Hypothesis is always that there is no link between the two categories, in other words everything is happening in proportion.

Test to use:

The statistic used is χ^2. The sample value is given by the equation:

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$, where i covers all the cells in the table.

The O_i are the observed values and E_i are the calculated expected values. This variable follows the χ^2 distribution with $(r - 1)(c - 1)$ degrees of freedom.

6. Tests on a distribution

In this instance, a frequency distribution is being tested to see if the random variable generating the numbers has a particular distribution. For example:

> 'The random variable H of Height in Men has a Normal distribution.' OR 'The numbers on a dice come up equally'.

The test invariably has a frequency distribution:

Height	*Frequency*
1.55m to 1.60m	4
1.60m to 1.65m	12
Etc…	Etc…
1.80m to 1.85m	9
1.85m to 1.90m	1

Null Hypothesis:

The Null Hypothesis is always that the frequency distribution **does have** the proposed distribution. The alternative is that it does not. Again, there is no parameter to test.

Test to use: The statistic used is χ^2. It is calculated from the data: call each of the observed values O_i and call the calculated expected values E_i, where i covers all the cells in the table. The statistic is given by the equation:

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}.$$

This variable follows the χ^2 distribution with $k - p - 1$ degrees of freedom, where k is the number of groups and p is the number of parameters for the distribution. The E_i are calculated from the distribution. Note that for the uniform distribution, $p = 0$.