Test statistics explained

Published on July 17, 2020 by Rebecca Bevans. Revised on August 2, 2021.

The test statistic is a number calculated from a statistical test of a hypothesis. It shows how closely your observed data match the distribution expected under the null hypothesis of that statistical test.

The test statistic is used to calculate the *p*-value of your results, helping to decide whether to reject your null hypothesis.

Table of contents

What exactly is a test statistic?

A test statistic describes how closely the distribution of your data matches the distribution predicted under the null hypothesis of the statistical test you are using.

The **distribution** of data is how often each observation occurs, and can be described by its central tendency and variation around that central tendency. Different statistical tests predict different types of distributions, so it's important to choose the right statistical test for your hypothesis.

The test statistic summarizes your observed data into a single number using the central tendency, variation, sample size, and number of predictor variables in your statistical model.

Generally, the test statistic is calculated as the pattern in your data (i.e. the correlation between variables or difference between groups) divided by the variance in the data (i.e. the standard deviation).

Example

You are testing the relationship between temperature and flowering date for a certain type of apple tree. You use a long-term data set that tracks temperature and flowering dates from the past 25 years by randomly sampling 100 trees every year in an experimental field.

- **Null hypothesis:** There is no correlation between temperature and flowering date.
- **Alternate hypothesis:** There is a correlation between temperature and flowering date.

To test this hypothesis you perform a regression test, which generates a *t*-value as its test statistic. The *t*-value compares the observed correlation between these variables to the null hypothesis of zero correlation.

Types of test statistics

Below is a summary of the most common test statistics, their hypotheses, and the types of statistical tests that use them.

Different statistical tests will have slightly different ways of calculating these test statistics, but the underlying hypotheses and interpretations of the test statistic stay the same.

***t*-value**
Null and alternative hypotheses: Null: The means of two groups are equal **Alternative:** The means of two groups are not equal
Statistical tests that use it: • *T*-test • Regression tests

***z*-value**
Null and alternative hypotheses: Null: The means of two groups are equal **Alternative:**The means of two groups are not equal
Statistical tests that use it: • *Z*-test

***F*-value**
Null and alternative hypotheses: Null: The variation among two or more groups is greater than or equal to the variation between the groups **Alternative:** The variation among two or more groups is smaller than the variation between the groups
Statistical tests that use it: • ANOVA • ANCOVA • MANOVA

X^2-value
Null and alternative hypotheses: Null: Two samples are independent **Alternative:** Two samples are not independent (i.e. they are correlated)
Statistical tests that use it: • Chi-squared test • Non-parametric correlation tests

In practice, you will almost always calculate your test statistic using a statistical program (R, SPSS, Excel, etc.), which will also calculate the *p*-value of the test statistic. However, formulas to calculate these statistics by hand can be found online.

Example

To test your hypothesis about temperature and flowering dates, you perform a regression test. The regression test generates:

- a regression coefficient of 0.36
- a *t*-value comparing that coefficient to the predicted range of regression coefficients under the null hypothesis of no relationship

The *t*-value of the regression test is 2.36 – this is your test statistic.

Vague sentences

Grammar

Style consistency

See an example

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.

Example

Your calculated *t*-value of 2.36 is far from the expected range of *t*-values under the null hypothesis, and the *p*-value is < 0.01. This means that you would expect to see a *t*-value as large or larger than 2.36 less than 1% of the time if the true relationship between temperature and flowering dates was 0.

Therefore, it is statistically unlikely that your observed data could have occurred under the null hypothesis. Using a significance threshold of 0.05, you can say that the result is **statistically significant**.

Reporting test statistics

Test statistics can be reported in the results section of your research paper along with the sample size, *p*-value of the test, and any characteristics of your data that will help to put these results into context.

Whether or not you need to report the test statistic depends on the type of test you are reporting.

Type of test: Correlation and regression tests
Which statistics to report: • Correlation coefficient or regression coefficient for each predictor variable • *p*-value for each predictor

Type of test: Tests of difference between groups
Which statistics to report: • Test statistic • Degrees of freedom • *p*-value for the test statistic

Example: Reporting the results of a regression test

In your survey of apple tree flowering dates, it is not necessary to report the test statistic – the regression coefficient and the *p*-value are sufficient:

> By surveying a random subset of 100 trees over 25 years we found a statistically significant (p < 0.01)

> positive correlation between temperature and flowering dates ($R2 = 0.36$, $sd = 0.057$).

Example: Reporting the results of a t-test

In a t-test of the difference between two groups, it is necessary to report the test statistic as well as the degrees of freedom and the *p*-value:

> In our comparison of mouse diet A and mouse diet B, we found that the lifespan on diet A (mean = 2.1 years; sd = 0.12) was significantly shorter than the lifespan on diet B (mean = 2.6 years; sd = 0.1), with an average difference of 6 months ($t(80) = -12.75$; $p < 0.01$).

Frequently asked questions about test statistics

- What is a test statistic?
- How do you calculate a test statistic?
- How do I know which test statistic to use?
- What factors affect the test statistic?
- What is statistical significance?

Other students also liked

Rebecca Bevans

e

cs

The p-value explained

The p-value shows the likelihood of your data occurring under the null hypothesis. P-values help determine statistical significance.

324

Statistical tests: which one should you use?

Your choice of statistical test depends on the types of variables you're dealing with and whether your data meets certain assumptions.

1453

Effect size in statistics

Effect size tells you how meaningful the relationship between variables or the difference between groups is.

87