The t-distribution and t-test

“ In probability and statistics, Student’s t-distribution (or simply the t-distribution) is a continuous probability distribution that arises when estimating the mean of a normally distributed population in situations where the sample size is small” (Narasimhan , 1996). Similar to the normal distribution, the t-distribution is symmetric and bell-shaped, but has heavier tails, meaning that it is more likely to produce values far from its mean. This makes the t-distribution useful for understanding statistical behaviors of random quantities.

It plays a role in a number of widely-used statistical analyses, including the Student’s t-test for assessing the statistical significance of the difference between two sample means, the construction of confidence intervals for the difference between two population means, and in linear regression analysis. Statistical significance is “ measuring the likelihood that an event occurs by chance” (Statistical Assessment Service, 2000). In statistics, a result is called statistically significant if it is unlikely to have occurred at random.

The amount of evidence required to accept that an event is unlikely to have arisen by chance is known as the significance level or p-value: “ the p-value measures consistency by calculating the probability of observing the results from your sample of data or a sample with results more extreme, assuming the null hypothesis is true” (Simon, 2007); or in simpler terms, a p-value is a measure of how much evidence there is against the null hypothesis.

The null hypothesis, traditionally represented by the symbol HO, “ represents the hypothesis of no change or no effect” (Simon, 2007). If the obtained p-value is small then it can be said either the null hypothesis is false or an

unusual event has occurred. A confidence interval “ gives an estimated range of values which is likely to include an unknown population parameter, the estimated range being calculated from a given set of sample data” (Mackowiak, 2002).

In statistics, a confidence interval is used to indicate the reliability of an estimate. A confidence interval with a particular confidence level is intended to give the assurance that, if the statistical model is correct, then taken over all the data that might have been obtained, the procedure for constructing the interval would produce a confidence interval that included the true value of the parameter the proportion of the time set by the confidence level.

More specifically, the meaning of the term “ confidence level” is that, if confidence intervals are constructed across many separate data analysis of repeated experiments, the proportion of such intervals that contain the true value of the parameter will approximately match the confidence level. “ The t-test assesses whether the means of two groups are statistically different from each other” (Trochim, 2006).

Student’s t-test deals with the problems associated with inference based on “ small” samples. A t-test is any statistical hypothesis test, (or a method of making decisions using data) in which the test statistic follows a Student’s t distribution, if the null hypothesis is supported. It is most commonly applied when the test statistic would follow a normal distribution if the value of a scaling term in the test statistic were known.

When the scaling term is unknown and is replaced by an estimate based on the data, the test statistic follows a Student’s t distribution. In statistical

hypothesis testing, a test statistic, “ is considered as a numerical summary calculated from a sample of data that reduces it to one or a small number of values that can be used to perform a hypothesis test” (Easton, ; McColl, 2003). Its value is used to decide whether or not the null hypothesis should be rejected in the hypothesis test.

Given a null hypothesis and a test statistic T, we can specify a “ null value” T0 such that values of T close to T0 present the strongest evidence in favor of the null hypothesis, whereas values of T far from T0 present the strongest evidence against the null hypothesis. An important property of a test statistic is that we must be able to determine its sampling distribution under the null hypothesis, which allows us to calculate p-values. For example, suppose we wish to test whether a coin is fair.

If we flip the coin 100 times and record the results, the raw data can be represented as a sequence of 100 Heads and Tails. If our interest is in the probability of obtaining a head, we only need to record the number T out of the 100 flips that produced a head, and use T0 = 50 as our null value. In this case, the exact sampling distribution of T is the binomial distribution, but for larger sample sizes the normal approximation can be used. Using one of these sampling distributions, it is possible to compute either a one-tailed or two-tailed p-value for the null hypothesis that the coin is fair.

The two-tailed test is when a given statistical hypothesis, H0 (the null hypothesis), will be rejected when the value of the test statistic is either sufficiently small or sufficiently large. This contrasts with a one-tailed test, in which only one of the rejection regions “ sufficiently small” or “ sufficiently

large” is preselected according to the alternative hypothesis being selected, and the hypothesis is rejected only if the test statistic satisfies that criterion.

Among the most frequently used t-tests are: a one-sample location test of whether the mean of a normally distributed population has a value specified in a null hypothesis, a two sample location test of the null hypothesis that the means of two normally distributed populations are equal, a test of the null hypothesis that the difference between two responses measured on the same statistical unit has a mean value of zero: for example, suppose we measure the size of a cancer patient’s tumor before and after a treatment; if the treatment is effective, we expect the tumor size for many of the patients to be smaller following the treatment. This is often referred to as the “ paired” or “ repeated measures” t-test. The last is a test of whether the slope of a regression line differs significantly from 0.

References

http://statweb. stanford. edu/~naras/jsm/TDensity/TDensity. html

http://www. stat. yale. edu/Courses/1997-98/101/confint. htm

http://www. socialresearchmethods. net/kb/stat_t. php

http://changingminds. org/explanations/research/analysis/t-test. htm

http://www. stats. gla. ac. uk/steps/glossary/hypothesis_testing. html#ts

http://www. ats. ucla. edu/stat/mult_pkg/faq/general/tail_tests. htm