What is the definition and example of the Bonferroni Correction?

Authored by
stats writer

April 26, 2024

RECOMMENDED CITATION

stats writer (2024). *What is the definition and example of the Bonferroni Correction?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=139457

The Bonferroni Correction is a statistical method used to adjust for multiple comparisons in a set of data. It is used to reduce the likelihood of false positive results when conducting multiple statistical tests. The correction involves dividing the desired level of significance (usually 0.05) by the number of comparisons being made.

For example, if a researcher is conducting 10 statistical tests with a desired level of significance of 0.05, the Bonferroni Correction would adjust the significance level to 0.005 (0.05/10) for each test. This ensures that the overall probability of a false positive result remains at 0.05 for the entire set of tests. This correction is commonly used in fields such as medical research and social sciences to account for the potential of Type I errors.

The Bonferroni Correction: Definition & Example

Whenever you perform a hypothesis test, there is always a chance of committing a type I error. This is when you reject the null hypothesis when it is actually true.

We sometimes call this a "false positive" - when we claim there is a statistically significant effect, but there actually isn't.

When we perform one hypothesis test, the type I error rate is equal to the significance level (α), which is commonly chosen to be 0.01, 0.05, or 0.10. However, when we conduct multiple hypothesis tests at once, the probability of getting a false positive increases.

When we conduct multiple hypothesis tests at once, we

have to deal with something known as a family-wise error rate, which is the probability that at least one of the tests produces a false positive. This can be calculated as:

Family-wise error rate = $1 - (1-\alpha)n$

where:

α: The significance level for a single hypothesis testn: The total number of tests

If we conduct just one hypothesis test using $\alpha = .05$, the probability that we commit a type I error is just .05.

Family-wise error rate = $1 - (1-\alpha)c = 1 - (1-.05)1 = 0.05$

If we conduct two hypothesis tests at once and use $\alpha = .05$ for each test, the probability that we commit a type I error increases to 0.0975.

Family-wise error rate = $1 - (1-\alpha)c = 1 - (1-.05)2 = 0.0975$

And if we conduct five hypothesis tests at once using $\alpha = .05$ for each test, the probability that we commit a type I error increases to 0.2262.

Family-wise error rate = 1 - (1-α)c = 1 - (1-.05)5 = 0.2262

It's easy to see that as we increase the number of statistical tests, the probability of commiting a type I error with at least one of the tests quickly increases.

One way to deal with this is by using a Bonferroni Correction.

What is a Bonferroni Correction?

A Bonferroni Correction refers to the process of adjusting the alpha (α) level for a family of statistical tests so that we control for the probability of committing a type I error.

αnew = αoriginal / n

where:

αoriginal: The original α leveln: The total number of comparisons or tests being performed

For example, if we perform three statistical tests at once and wish to use α = .05 for each test, the Bonferroni Correction tell us that we should use αnew = .01667.

αnew = αoriginal / n = .05 / 3 = .01667

Thus, we should only reject the null hypothesis of each individual test if the p-value of the test is less than .01667.

Bonferroni Correction: An Example

Suppose a professor wants to know whether or not three different studying techniques lead to different exam scores among students.

To test this, she randomly assigns 30 students to use each studying technique. After one week of using their assigned study technique, each student takes the same exam.

She then performs a and finds that the overall p-value is 0.0476. Since this is less than .05, she rejects the null hypothesis of the one-way ANOVA and concludes that not each studying technique produces the same mean exam score.

To find out *which* studying techniques produce statistically significant scores, she performs the following pairwise t-tests:

Technique 1 vs. Technique 2Technique 1 vs. Technique 3Technique 2 vs. Technique 3

She wants to control the probability of committing a type I error at α = .05. Since she's performing multiple tests at once, she decides to apply a Bonferroni Correction and use αnew = .01667.

αnew = αoriginal / n = .05 / 3 = .01667

She then proceeds to perform t-tests for each group and finds the following:

Technique 1 vs. Technique 2 | p-value = .0463Technique 1 vs. Technique 3 | p-value = .3785Technique 2 vs. Technique 3 | p-value = .0114

Since the p-value for Technique 2 vs. Technique 3 is the only p-value less than .01667, she concludes that there is only a statistically significant difference between technique 2 and technique 3.