

What is Yate's Continuity Correction and can you provide an example?

Authored by
stats writer

April 24, 2024

RECOMMENDED CITATION

stats writer (2024). *What is Yate's Continuity Correction and can you provide an example?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=138895>

Yate's Continuity Correction is a mathematical adjustment made to the calculation of a chi-square statistic, to account for the fact that the chi-square distribution is discrete while the expected values used in the calculation are continuous. This correction helps to improve the accuracy of the chi-square test and is commonly used in statistical analysis.

For example, let's say we want to compare the observed frequencies of two categories with the expected frequencies. Without Yate's Continuity Correction, the chi-square statistic may overestimate the significance of the differences between the observed and expected frequencies due to the discreteness of the distribution. By applying Yate's Continuity Correction, we are able to adjust the chi-square statistic and obtain a more accurate p-value for the test. This correction is particularly useful when the sample size is small.

Yate's Continuity Correction: Definition & Example

A is used to determine whether or not there is a significant association between two categorical variables.

This test uses the following null and alternative hypotheses:

H0: (null hypothesis) The two variables are independent.
H1: (alternative hypothesis) The two variables are *not* independent. (i.e. they are associated)

We use the following formula to calculate the Chi-Square test statistic X^2 for this test:

$$X^2 = \sum (O_i - E_i)^2 / E_i$$

where:

Σ : is a fancy symbol that means "sum"
O: observed value
E: expected value

This test assumes that the discrete probabilities of the frequencies in a contingency table can be approximated by the Chi-Square distribution, which is a continuous distribution.

However, this assumption tends to be slightly incorrect and the resulting test statistic tends to be biased upwards.

To correct for this bias we can apply Yate's continuity correction, which applies the following correction to the χ^2 formula:

$$\chi^2 = \Sigma (|O_i - E_i| - 0.5)^2 / E_i$$

We typically only use this correction when at least one cell in the contingency table has an expected frequency less than 5.

Example: Applying Yate's Continuity Correction

Suppose we want to know whether or not gender is

associated with political party preference. We take a simple random sample of 40 voters and survey them on their political party preference. The following table shows the results of the survey:

	Republican	Democrat	Independent	Total
Male	8	9	4	21
Female	11	3	5	19
Total	19	12	9	40

Here is how to perform a Chi-Square Test of Independence with Yate's continuity correction:

Observed Values:

	Republican	Democrat	Independent	Total
Male	8	9	4	21
Female	11	3	5	19
Total	19	12	9	40

Expected Values:

	Republican	Democrat	Independent
Male	9.975	6.3	4.725
Female	9.025	5.7	4.275

Note: We calculate the expected value in each cell by multiplying the row total by the column total, then dividing by the grand total. For example, the expected number of male republicans is $(21 \cdot 19) / 40 = 9.975$.

Chi-Square Test Statistic: $X^2 = \sum (|O_i - E_i| - 0.5)^2 / E_i$

$$\begin{aligned} & (|8 - 9.975| - 0.5)^2 / 9.975 = .218 (|9 - 6.3| - 0.5)^2 / 6.3 = \\ & .768 (|4 - 4.725| - 0.5)^2 / 4.725 = .011 (|11 - 9.025| - 0.5)^2 / \\ & 9.025 = .241 (|3 - 5.7| - 0.5)^2 / 5.7 = .849 (|5 - 4.275| - 0.5)^2 / \\ & 4.275 = .012 \end{aligned}$$

Thus, $X^2 = .218 + .768 + .011 + .241 + .849 + .012 = 2.099$

P-Value: According to the , the p-value that corresponds to a Chi-Square test statistic with 2 degrees of freedom is 0.3501.

Since this p-value is not less than .05, we would fail to reject the null hypothesis. This means we do not have sufficient evidence to say that there is an association between gender and political party preference.