Biostatistics – Spring 2026
Lecture 04: Yates Continuity Correction and Chi-square Approximation

Dr. Zaid T. Al-Khaledi
Department of Statistics and Informatics
University of Mosul

March 4, 2026

Introduction

In the previous lecture we learned Fisher's Exact Test, which is recommended when sample sizes are small or when expected counts are very low.

In practice, many studies fall into an intermediate situation: the sample size is not extremely small, but it is not very large either.

In such cases, the usual Chi-square test may slightly overestimate statistical significance.

To reduce this problem, a modification called the **Yates Continuity Correction** is sometimes used for 2×2 tables.

This lecture explains why the correction is used, how it is calculated, and how it compares to Fisher's Exact Test.

By the end of this lecture, you should be able to:

- Why a Continuity Correction Is Needed.
- Yates Correction Formula.
- Step-by-Step Example.
- Comparison: Chi-square vs Yates vs Fisher.
- Practical Guidelines for Choosing the Test.

1. Continuity Correction (Yates Correction) for 2×2

When the sample size is not very large (but not extremely tiny), some books suggest a correction to the Chi-square test for 2×2 tables:

$$\chi^2_{\text{Yates}} = \sum \frac{(|O - E| - 0.5)^2}{E}.$$

Why subtract 0.5?

Because the Chi-square distribution is continuous, but table counts are discrete (whole numbers). The correction tries to reduce the approximation error and makes the test more conservative.

Example

Suppose we observe:

	B_1	B_2	Row Total
A_1	8	2	10
A_2	4	6	10
Column Total	12	8	20

So:

$$r_1 = 10, \quad r_2 = 10, \quad c_1 = 12, \quad c_2 = 8, \quad n = 20.$$

Step 1: Expected counts

$$E_{11} = \frac{r_1 c_1}{n} = \frac{10 \cdot 12}{20} = 6, \qquad E_{12} = \frac{r_1 c_2}{n} = \frac{10 \cdot 8}{20} = 4,$$
$$E_{21} = \frac{r_2 c_1}{n} = \frac{10 \cdot 12}{20} = 6, \qquad E_{22} = \frac{r_2 c_2}{n} = \frac{10 \cdot 8}{20} = 4.$$

Step 2: Usual Chi-square (no correction)

Compute each contribution $\frac{(O-E)^2}{E}$:

$$\frac{(8-6)^2}{6} = \frac{4}{6} = 0.6667, \qquad \frac{(2-4)^2}{4} = \frac{4}{4} = 1.0000,$$
$$\frac{(4-6)^2}{6} = \frac{4}{6} = 0.6667, \qquad \frac{(6-4)^2}{4} = \frac{4}{4} = 1.0000.$$

Add them:

$$\chi^2 = 0.6667 + 1.0000 + 0.6667 + 1.0000 = 3.3333.$$

Degrees of freedom for a 2×2 table:

$$df = (2-1)(2-1) = 1.$$

So the critical value is:

$$\chi^2_{(1)} = 3.841 > \chi^2 = 3.3333 \quad \Rightarrow \quad \text{Not significant (without correction).}$$

Step 3: Chi-square with Yates correction

Now compute each corrected contribution $\frac{(|O-E|-0.5)^2}{E}$.

For cell $(1, 1)$:

$$\frac{(|8-6|-0.5)^2}{6} = \frac{(2-0.5)^2}{6} = \frac{2.25}{6} = 0.3750.$$

For cell $(1, 2)$:

$$\frac{(|2-4|-0.5)^2}{4} = \frac{(2-0.5)^2}{4} = \frac{2.25}{4} = 0.5625.$$

For cell $(2, 1)$:

$$\frac{(|4-6|-0.5)^2}{6} = \frac{(2-0.5)^2}{6} = \frac{2.25}{6} = 0.3750.$$

For cell $(2,2)$:

$$\frac{(|6-4|-0.5)^2}{4} = \frac{(2-0.5)^2}{4} = \frac{2.25}{4} = 0.5625.$$

Add them:

$$\chi^2_{\text{Yates}} = \ 0.3750 + 0.5625 + 0.3750 + 0.5625 = \ 1.8750.$$

With $df = 1$, the critical value is:

$$\chi^2_{(1)} = 3.841 > \chi^2_{\text{Yates}} = 1.8750 \quad \Rightarrow \quad \text{Not significant (with Yates correction).}$$

Both tests do not give enough evidence of an association at $\alpha = 0.05$. Yates correction makes the Chi-square smaller and the p-value larger, so it is more conservative.

Homework (HW)

HW1 (Yates Continuity Correction)

For the following table, use $\alpha = 0.05$:

	Disease	*NoDisease*	*Total*
Exposed	12	8	20
NotExposed	5	15	20
Total	17	23	40

Required:

1. Write H_0 and H_1.
2. Compute all expected counts.
3. Compute the usual Chi-square statistic (without correction).
4. Compute the Chi-square statistic using Yates correction.
5. Compare the two Chi-square values and make your decisions about H_0.
6. State which test you would report and explain why.