ChatGPT 5.2 Share

Would like a Pearson test (probably with Yates correction) for the following data. There are two classification variables the first with values "PG&E" and "non-PG&E" and the second with values "RFW" and "non-RFW".
PG&E
11
RFW 9
Non-RFW 2

Non-PG&E
20
RFW 4
Non-RFW 16

Here's the **2×2 contingency table** written explicitly:

	RFW	Non-RFW	Total
PG&E	9	2	11
Non-PG&E	4	16	20
Total	13	18	31

Pearson χ² test (with Yates continuity correction)

Expected counts

$$E_{ij} = \frac{(\text{row total})(\text{column total})}{N}$$

- PG&E, RFW: $11 \times 13/31 = 4.61$
- PG&E, Non-RFW: $11 \times 18/31 = 6.39$
- Non-PG&E, RFW: $20 \times 13/31 = 8.39$
- Non-PG&E, Non-RFW: $20 \times 18/31 = 11.61$

All expected counts are > 1, so χ² is acceptable.

Yates-corrected χ² statistic

For each cell:

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

Contributions:

- PG&E, RFW: $(|9 - 4.61| - 0.5)^2 / 4.61 = 3.34$
- PG&E, Non-RFW: $(|2 - 6.39| - 0.5)^2 / 6.39 = 2.41$
- Non-PG&E, RFW: $(|4 - 8.39| - 0.5)^2 / 8.39 = 1.84$
- Non-PG&E, Non-RFW: $(|16 - 11.61| - 0.5)^2 / 11.61 = 1.33$

Total χ² (Yates):

$$\chi^2 = 8.92$$

Degrees of freedom

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

---

p-value

For χ² = 8.92 with 1 df:

$$\boxed{p \approx 0.0028}$$

Interpretation

There is a **statistically significant association** between utility type (PG&E vs non-PG&E) and RFW classification at conventional levels (p < 0.01).

PG&E observations are **much more likely** to be RFW than non-PG&E observations.

Notes

- Without Yates correction: χ² ≈ 11.3, p ≈ 0.0008
- Fisher's exact test (two-sided) would also be appropriate here and will give a

similarly small p-value

If you want, I can:

- Run Fisher's exact test explicitly
- Compute odds ratio + confidence interval
- Provide R / Python / Excel formulas for reproducibility