Tidsskriftet Tidsskrift for Den norske legeforening

Paired categorical data

MEDICINE AND NUMBERS

STIAN LYDERSEN

stian.lydersen@ntnu.no
Stian Lydersen, PhD, professor of medical statistics at the Regional Centre for Child and Youth Mental Health and Child Welfare, Department of Mental Health, Norwegian University of Science and Technology.
The author has completed the ICMJE form and declares no conflicts of interest.

To compare proportions in two independent groups, one can use, for example, Pearson's chi-squared test. However, if the data originate from matched pairs, we must use methods that account for the dependence within the pairs.

Matched pairs can originate from various situations and trial designs. The pairs may represent the status of a person before and after a treatment; a randomised controlled trial with a contralateral design where, for example, the left and right eye are randomised to eye drops A and B respectively; a matched case control study; or a crossover trial, as in the example below.

As previously described in this column, a crossover trial is a randomised controlled trial where the participants receive both a treatment A and a treatment B in randomised order (1). As an example, we used a trial that compared the effect of a probiotic with that of placebo in participants with irritable bowel syndrome (2), and we described relevant methods of analysis for a continuous outcome variable. We will now look at a categorical outcome variable, more specifically whether the participant experienced success in terms of satisfactory symptom reduction.

Concordant and discordant pairs

Here, there are four possible results for each participant: the participant experiences success with both A and B, neither A nor B, only with A or only with B. The results from this trial (2) are shown in Table 1, where A is placebo and B is active treatment. We see that eleven participants experienced no success with any of the treatments, while one participant experienced success with both treatments. These twelve paired observations are called concordant pairs, because the two observations in the pairs match. If the participant has success only with A or only with B, this is called a discordant pair. In this trial there were four discordant pairs, all four of which experienced success with A and none with B.

Table 1

Results for 16 participants, of whom 4 + 1 = 5 experienced success with treatment A and 0 + 1 = 1 experienced success with treatment B (2).

	Success with B (active treatment)		
Success with A (placebo)	No	Yes	Total
No	11	0	11
Yes	4	1	5
Total	15	1	16

Confidence interval and *p*-value

In this example, the proportions of successes with treatment A and B were 5/16 = 0.31 and 1/16 = 0.06 respectively, with a difference of 0.25 in favour of treatment A, which was the placebo. It is often relevant to estimate a confidence interval for the difference between the probability of success with the two treatments, for the ratio between the probabilities, or for the odds ratio. The latter will be relevant for a matched case-control study. Various methods can be used to estimate the confidence interval and *p*-value for matched pairs (3). Several of these methods are based on the discordant pairs. If there is no difference between the treatments, we could expect that one-half of these experience success with A and one-half the opposite. The more the success rate for A deviates from one-half, the stronger the evidence that the treatments have different effects.

McNemar's test is based exactly on this. We have previously compared a total of five different versions of McNemar's test (3) and found that one of the versions that is the simplest to estimate, called McNemar's mid-*p* test, can always be

recommended. Surprisingly, it has better properties than a so-called exact test, including for small samples. In this example, we obtain $p = 0.0625$. We also studied various methods for confidence intervals (3) and found that among others, Newcombe's method for the difference can always be recommended. In this example, a 95 % confidence interval for the difference is -0.494 to 0.015. An overview of different methods and recommendations is found in (4). Many of these methods were unavailable in statistical software before 2022, but are now included in SPSS version 28 and in the software package 'contingencytables' for R (5).

More than two categories

In this example, the outcome variable had only two categories, and the paired observations could be shown in a 2×2-table, such as in Table 1. If the outcome variable has three categories, the results can be shown in a 3×3 table etc. An overview of methods and recommendations for paired categorical data with more than two categories is found in Fagerland et al. (4). These methods are also implemented in 'contingencytables' for R (5).

REFERENCES

1. Lydersen S. Overkrysningsforsøk. Tidsskr Nor Legeforen 2022; 142. doi: 10.4045/tidsskr.22.0013. [CrossRef]

2. Ligaarden SC, Axelsson L, Naterstad K et al. A candidate probiotic with unfavourable effects in subjects with irritable bowel syndrome: a randomised controlled trial. BMC Gastroenterol 2010; 10: 16. [PubMed][CrossRef]

3. Fagerland MW, Lydersen S, Laake P. Recommended tests and confidence intervals for paired binomial proportions. Stat Med 2014; 33: 2850–75. [PubMed][CrossRef]

4. Fagerland MW, Lydersen S, Laake P. Statistical analysis of contingency tables. Philadelphia, PA: CRC Press, 2017.

5. contingencytables: Statistical Analysis of Contingency Tables. https://cran.r-project.org/web/packages/contingencytables/index.html Accessed 1.2.2022.

Publisert: 19 May 2022. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.22.0014