

Does an academic copy-editor require knowledge of statistics?

Marisha Fonseca

Cactus Communications, Mumbai, India; marishaf@cactusglobal.com

When I started out as an academic copy-editor, I was concerned about everything except statistics. Back then, it seemed more important to identify and correct a comma splice than a negative *P* value! But as I grew more comfortable with the ‘language’ aspects of editing, I increasingly realised that fine language alone would not help my authors get their papers published.

Currently, many journal articles contain statistical flaws.^{1,2} Journal editors and researchers are becoming increasingly concerned with incorrect reporting of statistical data^{3,4} (to the point of ensuring that manuscripts are reviewed by at least one expert in statistics).⁴ With knowledge in statistics, copy-editors can help speed up the peer review and avoid some mistakes.

As an editor, I frequently come across reviewer comments and author responses for papers, which I previously edited, and I am surprised at how ‘elementary-looking’ issues keep popping up. I usually delve deeper into statistics textbooks and websites during my free time, and all this extra reading pays off. This knowledge helps me not only edit better, but also guide my clients on how to maximise their chances of publication.

Some of the most glaring errors, which I have managed to catch, concern the use of Greek symbols. Here is my favourite example: ‘We found a strong correlation between X and Y ($p = 0.761$).’ The parenthetical text implies that the correlation was *not* statistically significant. But the author actually meant to use the Greek letter ‘ ρ ’ (to denote Spearman’s rank-order correlation coefficient) instead of the English letter ‘*p*’!

I have also come across ‘*r*’ being used for Spearman’s correlation coefficient. It is typically used for Pearson’s product-moment correlation coefficient. If it is used in place of Spearman’s rho, it should be accompanied by a subscripted ‘*s*’. Another frequent mistake is the incorrect use of ‘Cronbach’s *a*’ and ‘Cohen’s *k*’ instead of correct ‘Cronbach’s α ’ and ‘Cohen’s κ ’.

Most often, I deal with the problems of *P* values. Many authors report ‘ $P = 0.000$ ’, though *P* cannot be equal to zero. The ‘0.000’ occurs when statistical programs automatically truncate or round off extremely small *P* values. Actually, ‘ $P < 0.001$ ’ reads well and has the same meaning. Another tip concerns the use of ‘ $P < 0.000$ ’. Technically, this means that *P* is negative, which is again impossible.

Knowledge of statistics can also be useful when copy-editors try to meet a word limit. For example:

Original: When X increased, Y decreased and vice versa.

Edited: X and Y were inversely correlated.

Original: Because we did not have a large number of participants, it may not be possible to state that our findings hold true for all such patients.

Edited: The small sample size limits the generalisability of our findings.

Original: A t test showed that X was higher than Y and this difference was significant ($t[38] = 5.2, P < .001$).

Edit: X was higher than Y ($t[38] = 5.2, P < .001$).

Knowledge of statistics helps format papers as per certain style guides. For example, the Publication Manual of the American Psychological Association, Sixth Edition, recommends that the zero before the decimal point be omitted for values that cannot exceed one. Such zeroes can be removed for *P* values, correlation coefficients, coefficients of determination, but not for odds ratios, standard deviations, and so on.

Finally, understanding of statistics helps copy-editors avoid incorrect editing. For example, even if the author uses ‘significantly increased’ several times in a paragraph, we should not replace ‘significantly’ with ‘substantially’ for some variation in word choice. In a statistical sense, ‘significantly’ is more to do with probability and does not mean ‘considerably’ or ‘notably’.

One may think that it is more important to correct a split infinitive than a correlation coefficient of ‘38’ (which came about because the author missed typing in the decimal point). While language is the primary focus for copy-editors, the authors will be delighted if wordsmiths prevent them from appearing silly to the scientific community. Knowledge of statistics is crucial for any specialist in biomedicine, psychology, sociology or education. It seems like a lot to learn but, in the end, our authors and journals benefit!

References

- 1 Bakker M, Wicherts JM. The (mis)reporting of statistical results in psychology journals. *Behavior Research Methods* 2011;43(3):666–678. doi: 10.3758/s13428-011-0089-5
- 2 Strasak AM, Zaman Q, Pfeiffer KP, Göbel G, Ulmer H. Statistical errors in medical research—a review of common pitfalls. *Swiss Medical Weekly* 2007;137(3–4):44–49.
- 3 Lang T. Twenty statistical errors even you can find in biomedical research articles. *Croatian Medical Journal* 2004;45(4):361–370.
- 4 PLoS Medicine Editors. Better reporting, better research: Guidelines and guidance in PLoS Medicine. *PLoS Medicine* 2008;5(4):e99. DOI: 10.1371/journal.pmed.0050099