Some Data on the Frequency of Errors in Mathematics Papers

Leslie Lamport

14 September 2022
minor changes 18 February 2023 and 9 December 2023

Abstract

A report on the frequency of errors in a non-random selection of 84 papers in one small field of mathematics. The data are meager, but they seem to be the only data on the subject.

While there is plenty of anecdotal evidence of errors in papers published in mathematical journals, I know of no published data on the frequency of those errors. A Web search on the topic revealed just one paper with any relevant data [1]; but they were on the fraction of papers for which errata or corrigenda were published, and the author states that few errors are reported in that way.

I have compiled a very small amount of data suggesting that about a third of published math papers contain an incorrect theorem or proof. While the data are of quite limited significance, they seem to be the only data there are and therefore worth reporting. Here is a description of the data, beginning with an explanation of their source.

George Bergman is a mathematician at the University of California at Berkeley. He wrote many reviews for *Mathematical Reviews*, a publication that contains reviews of published mathematical articles. Bergman was one of the few *Math Reviews* reviewers who read the papers carefully, and he found errors in many of them.

At one time, he showed me the 51 reviews he had by then written in *Math Reviews*. Based on those reviews, I judged that exactly 1/3 of the papers he reviewed contained an error, which I defined to mean an incorrect statement in a proof or result that the author believed to be correct. These were what I thought to be serious errors, not easily corrected mistakes. I mentioned this result in a 1995 paper on how to write proofs [2].

All the reviews in *Math Reviews* are now available on-line (by subscription) from the American Mathematical Society's service *MathSciNet* (`https://mathscinet.ams.org/mathscinet/`). Bergman wrote a total of 84 reviews. In what I believe were the original 51 reviews, I again judged that 17 of them reported errors. (I found one review to be borderline and couldn't decide if a mistake it reported was a serious error. I classified it as not an error in order to obtain the same number of errors as before.) Eight of those 17 reviews reported incorrect results. In two of those eight, theorems were incorrect because they were based on results from two of the other papers Bergman had found to be incorrect. I also found that two of the error-free papers should not have been counted, since they just corrected errors reported in Bergman's earlier reviews.

The statistics for all 84 of the reviews were quite similar to those for the original 51. The 28 reviews reporting errors are an identical fraction of the total number of reviews. The 11 reviews reporting incorrect results is a reduction in the ratio of incorrect results to total errors from 47% to 39%.

These numbers should not be considered accurate for the entire mathematics literature. They were all in Bergman's area of expertise. (He is an algebraist specializing in ring theory.) However, I know of no reason to expect the number of errors in other fields of math to be significantly different.

These statistics are questionable for another reason. Whether a mistake in a paper meets my definition of an error, rather than just being a typo, depends on what the author believed. These statistics therefore rest on a subjective judgement. My belief that they are significant is enhanced by the consistency in the classifications of the first 51 reviews I made about 25 years apart. It would be useful to have an independent classification of the reviews, and I encourage others to provide one. The reviews can be obtained by searching *Math Reviews* for reviewer "Bergman,G" on MathSciNet.

References

[1] Joseph F. Grcar. Errors and corrections in mathematics literature. *Notices of the AMS*, 60(4):418–425, April 2013.

[2] Leslie Lamport. How to write a proof. *American Mathematical Monthly*, 102(7):600–608, August-September 1995.