

Assignment 1

Angela Zhai

September 14, 2018

Document assignment

Consider the following situation:

A sloppy printer produces books with an average of 2 misprints per page. You want to know how many pages have more than k misprints in a book of n pages. Make an $n \times k$ table that shows the relationship between the total number of pages in a book and the number of pages with k misprints.

Show and explain your work. Include equations and calculations to teach the reader how to solve the problem. Include an image of a book.

Push your solution to a github repository and submit the url for repository on blackboard. Be sure your repo includes your document as a pdf file and as an RMD file. Include other files needed to recompile your document.

Result

Table 1: Probability of pages with k misprints

| n | Least Misprints | | | | | |
|----|-----------------|-----------|-----------|-----------|-----------|-----------|
| | 0 | 1 | 2 | 3 | 4 | 5 |
| 5 | 0.0e+00 | 0.0000000 | 0.0002221 | 0.2624659 | 0.9533530 | 0.9998243 |
| 10 | 0.0e+00 | 0.0000000 | 0.0393399 | 0.9076959 | 0.9999525 | 1.0000000 |
| 15 | 0.0e+00 | 0.0000250 | 0.4271040 | 0.9988426 | 1.0000000 | 1.0000000 |
| 20 | 0.0e+00 | 0.0043677 | 0.9032285 | 0.9999982 | 1.0000000 | 1.0000000 |
| 25 | 0.0e+00 | 0.1138421 | 0.9969101 | 1.0000000 | 1.0000000 | 1.0000000 |
| 30 | 3.3e-06 | 0.5876265 | 0.9999842 | 1.0000000 | 1.0000000 | 1.0000000 |

For this question, the result should be the conditional probability. This situation is to find out n pages have more than k misprints. We can assume there is a 50 pages book which needs to be printed.

First, we should calculate a cumulative probability for more than k misprints which follows poisson distribution.

The equation to get CDF of poisson distribution is: $e^{-\lambda} \sum_{i=0}^k \frac{\lambda^i}{i!}$

Like if we want to know the probability of more than 2 misprints, the equation should look like: $e^{-2} \sum_{i=2}^n \frac{2^i}{i!}$

Then, we need to calculate the conditional probability of pages which have more than k misprints. This is a binomial distribution situation because we will get pages which have no more than k misprints and pages have more than k misprints.

Through these process, we can get the final table as above.

Image

