

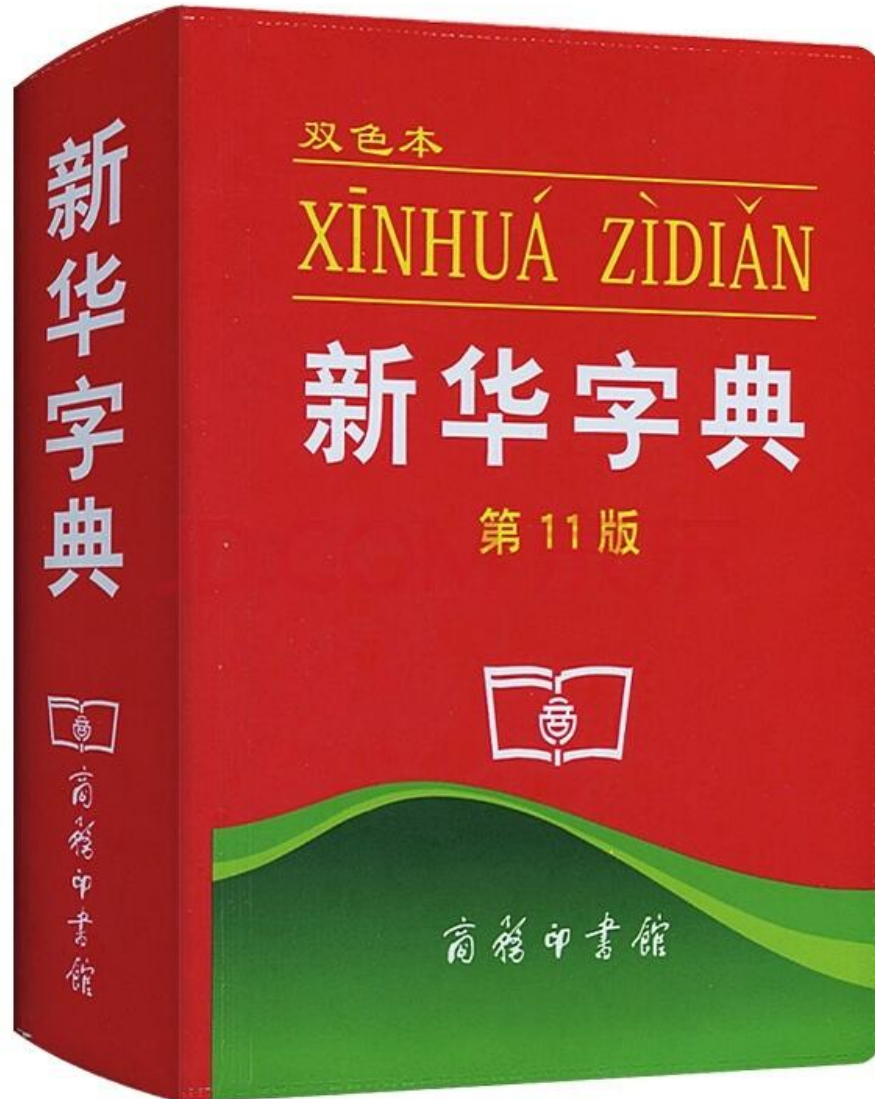
615__HW1

Yifeng Luo

9/17/2018

A sloppy printer produces books with an average of 2 misprints per page. You want to know how many pages in a 50 page book will have more than k misprints. Make an $n \times k$ table that shows the probability that n or fewer pages in a 50 page book will have more than k misprints.

```
cover_url= 'https://github.com/924935743/BU615/blob/master/.gitignore/cover.jpg'
if(!file.exists(cover_file <- 'cover.jpg'))
  download.file(cover_url, cover_file, mode = 'wb')
knitr::include_graphics(if (identical(knitr::pandoc_to(), 'html')) cover_url else cover_file)
```



Note: the book picture from website

Given the condition that a book has an average of 2 misprints on each page, the distribution of number of misprints (denoted by k) on this book is assumed by Poisson distribution:

Poisson($\lambda = 2$).

poisson distribution formula: $P(k \text{ events in interval}) = e^{-\lambda} \frac{\lambda^k}{k!}$

The probability p_k of more than k misprints on a page is:

$p_k = P(W > k) = 1 - P(W \leq k) = 1 - \text{ppois}(k, \lambda=2)$.

The probability of pages that is less than n and with more than k misprints (denoted as T) is:

$P(T \leq n) = \text{pbinom}(n, 50, p_k)$.

Our goal is to produce a table that shows that $P(T = n | k = 1, 2, \dots, 49)$. Firstly we need to know p_k for each k .

In the table below, the second column refers to $P(W = k)$, the third column refers to $1 - P(W \leq k)$

Note: $P(W \leq k)$ is the accumulative prob.

```
options(digits = 3)
options(scipen = 999)
#produce a dataframe containing $P(T = n | k = 1, 2, ..., 49)$
k <- 0:49
#P(W=k)
prob <- as.data.frame(dpois(x = 0:49, lambda = 2))
#pk = 1 - cumulative probability = P(W>k)
prob <- cbind(k, prob, cumsum(prob), 1 - cumsum(prob))
colnames(prob) <- c("k_value", "P(W=k)", "P(W>k)", "pk")
p1.table <- kable(prob)
kable_styling(p1.table, bootstrap_options = "striped", full_width = FALSE, position = "left")
```

k_value	P(W=k)	P(W>k)	pk
0	0.135	0.135	0.865
1	0.271	0.406	0.594
2	0.271	0.677	0.323
3	0.180	0.857	0.143
4	0.090	0.947	0.053
5	0.036	0.983	0.017
6	0.012	0.995	0.005
7	0.003	0.999	0.001
8	0.001	1.000	0.000
9	0.000	1.000	0.000
10	0.000	1.000	0.000
11	0.000	1.000	0.000
12	0.000	1.000	0.000
13	0.000	1.000	0.000
14	0.000	1.000	0.000
15	0.000	1.000	0.000
16	0.000	1.000	0.000
17	0.000	1.000	0.000
18	0.000	1.000	0.000
19	0.000	1.000	0.000
20	0.000	1.000	0.000
21	0.000	1.000	0.000
22	0.000	1.000	0.000
23	0.000	1.000	0.000
24	0.000	1.000	0.000
25	0.000	1.000	0.000
26	0.000	1.000	0.000
27	0.000	1.000	0.000
28	0.000	1.000	0.000
29	0.000	1.000	0.000
30	0.000	1.000	0.000
31	0.000	1.000	0.000
32	0.000	1.000	0.000
33	0.000	1.000	0.000
34	0.000	1.000	0.000
35	0.000	1.000	0.000
36	0.000	1.000	0.000
37	0.000	1.000	0.000
38	0.000	1.000	0.000
39	0.000	1.000	0.000
40	0.000	1.000	0.000
41	0.000	1.000	0.000
42	0.000	1.000	0.000
43	0.000	1.000	0.000
44	0.000	1.000	0.000
45	0.000	1.000	0.000
46	0.000	1.000	0.000
47	0.000	1.000	0.000
48	0.000	1.000	0.000
49	0.000	1.000	0.000

Next we need to calculate $P(T \leq n)$ for $k = 0, 1, \dots, 49$

```
options(digits = 3)
options(scipen = 999)
p2.table <- 0:50
p2.table <- as.data.frame(p2.table)
colnames(p2.table) <- c("n")
cp <- prob[, 4]
```

producing columns of $P(T \leq n)$ for each k value between 0 and 49

```
for (i in 1:50) {
  fc <- pbinom(q = 0:50, size = 50, prob = cp[i])
  fc <- as.data.frame(fc)
  colnames(fc) <- i - 1
  p2.table <- cbind(p2.table, fc)
}
kable(p2.table[,0:12])
```

n	0	1	2	3	4	5	6	7	8	9	10
0	0.000	0.000	0.000	0.000	0.067	0.434	0.797	0.947	0.988	0.998	1
1	0.000	0.000	0.000	0.004	0.253	0.799	0.978	0.999	1.000	1.000	1
2	0.000	0.000	0.000	0.019	0.506	0.950	0.998	1.000	1.000	1.000	1
3	0.000	0.000	0.000	0.060	0.731	0.991	1.000	1.000	1.000	1.000	1
4	0.000	0.000	0.000	0.140	0.878	0.999	1.000	1.000	1.000	1.000	1
5	0.000	0.000	0.000	0.262	0.953	1.000	1.000	1.000	1.000	1.000	1
6	0.000	0.000	0.001	0.416	0.985	1.000	1.000	1.000	1.000	1.000	1
7	0.000	0.000	0.003	0.576	0.996	1.000	1.000	1.000	1.000	1.000	1
8	0.000	0.000	0.008	0.720	0.999	1.000	1.000	1.000	1.000	1.000	1
9	0.000	0.000	0.018	0.831	1.000	1.000	1.000	1.000	1.000	1.000	1
10	0.000	0.000	0.039	0.908	1.000	1.000	1.000	1.000	1.000	1.000	1
11	0.000	0.000	0.076	0.954	1.000	1.000	1.000	1.000	1.000	1.000	1
12	0.000	0.000	0.133	0.979	1.000	1.000	1.000	1.000	1.000	1.000	1
13	0.000	0.000	0.212	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1
14	0.000	0.000	0.312	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1
15	0.000	0.000	0.427	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1
16	0.000	0.000	0.547	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
17	0.000	0.000	0.662	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
18	0.000	0.001	0.762	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
19	0.000	0.002	0.843	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
20	0.000	0.004	0.903	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
21	0.000	0.010	0.944	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
22	0.000	0.020	0.970	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
23	0.000	0.038	0.985	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
24	0.000	0.068	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
25	0.000	0.114	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
26	0.000	0.178	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
27	0.000	0.262	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
28	0.000	0.362	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
29	0.000	0.473	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
30	0.000	0.588	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
31	0.000	0.695	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
32	0.000	0.789	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
33	0.000	0.864	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
34	0.001	0.918	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
35	0.002	0.955	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
36	0.005	0.977	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
37	0.014	0.989	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
38	0.032	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
39	0.068	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
40	0.131	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
41	0.229	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
42	0.364	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
43	0.524	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
44	0.686	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
45	0.825	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
46	0.921	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
47	0.973	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
48	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
49	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1