Notes for Assignment 1

Haviland Wright

September 15, 2018

1 Calculate p_k

Given the stated mean value of 2 misprints per page, assume the distribution of M, the number of misprints on a page, is $Poisson(\lambda = 2)$

Conditioned on k, a limit to be determined, the probability p_k of more than k misprints on a page is:

$p_k = P(M > k) = 1 - P(M \leq k) = 1$ - ppois(k, lambda $= 2$).

2 Calculate $P(T \leq n|k)$

Probability of T no more than n pages with more than k misprints is:

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

3 Visualize $P(n \leq k)$

For a given k, the probability mass and cumulateive distribution functions of errors per page can be visualized as follows:

```
options(digits = 3)
options(scipen = 999)

produce a dataframe containing probabilities
prob <- as.data.frame(dpois(x = 0:13, lambda = 2))
prob <- cbind(prob, cumsum(prob)) ## these probabilities = ppois(0:10,lambda=2)
prob <- cbind(prob, 1 - prob[,2])

colnames(prob) <- c("prob", "cprob", "tail_prob")

barplot with mass function and CDF
df.bar <- barplot(prob$prob,
                  main="Poisson mass function, lambda = 2",
                  names.arg = c("0","1","2","3","4","5","6","7","8","9","10","11","12","13"),
                  xlab = "X: Number of misprints on a page",
                  ylab = "Probabilty: P(x)",ylim = c(0,1.1), col="red")

lines(x=df.bar, y=prob$cprob, col="black")
points(x=df.bar, y=prob$cprob, pch=16, col="blue")
abline(h=1, lty=2, col = "cyan")
```

## Poisson mass function, lambda = 2



X: Number of misprints on a page
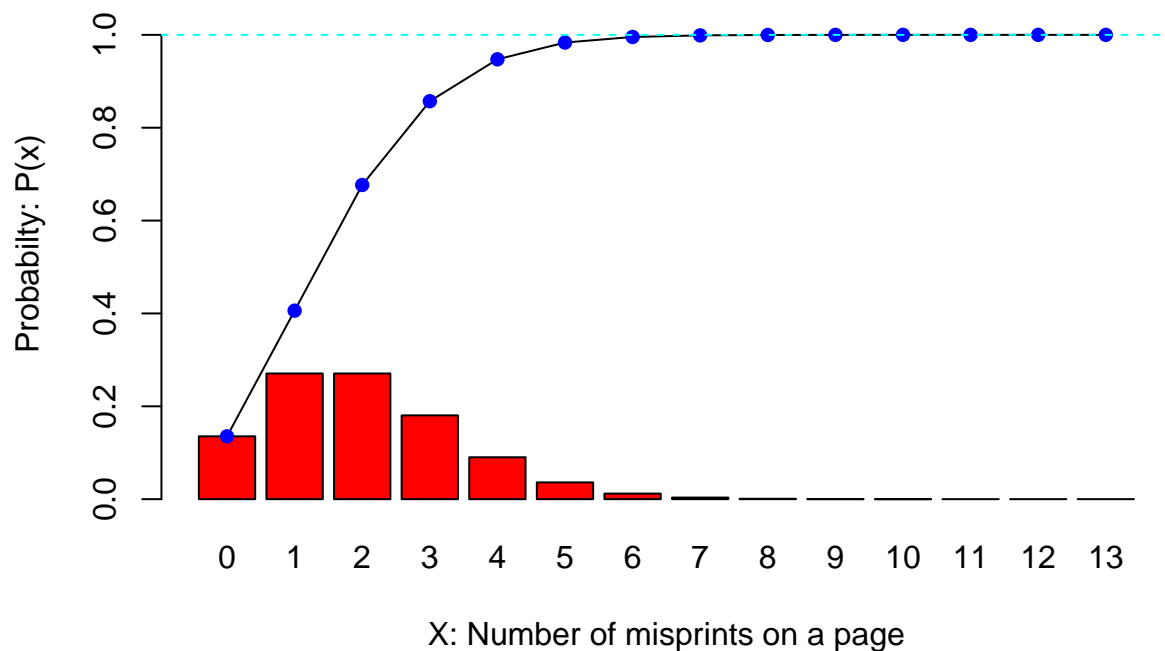
## 4 Key Values Table

A table of the first 11 probabilities provides the tail area probabilities, $p_k = P(M > k)$.

```
k <- 0:13

prob <- cbind(k, prob)

## format data frame for output as a table
p.tab <- kable(prob)
kable_styling(p.tab, bootstrap_options = "striped", full_width = FALSE, position = "left")
```

| k	prob	cprob	tail_prob
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

5 Visualize $P(T = n | k = 1, 3, 5)$

In an n rows by k columns table, each column will reflect different parameters

```r
# Setup for multiple plots
par(mfrow=c(3,1))

# calculate probability functions for display, make barplots.
f1 <- dbinom(x = 0:50, size = 50, prob = .594)
b1 <- barplot(f1, main="k = 1", axes = TRUE)


f3 <- dbinom(x = 0:50, size = 50, prob = .143)
barplot(f3, main="k = 3")

f5 <- dbinom(x = 0:50, size = 50, prob = .017)
barplot(f5, main="k = 5")
```

k = 1

k = 3

k = 5
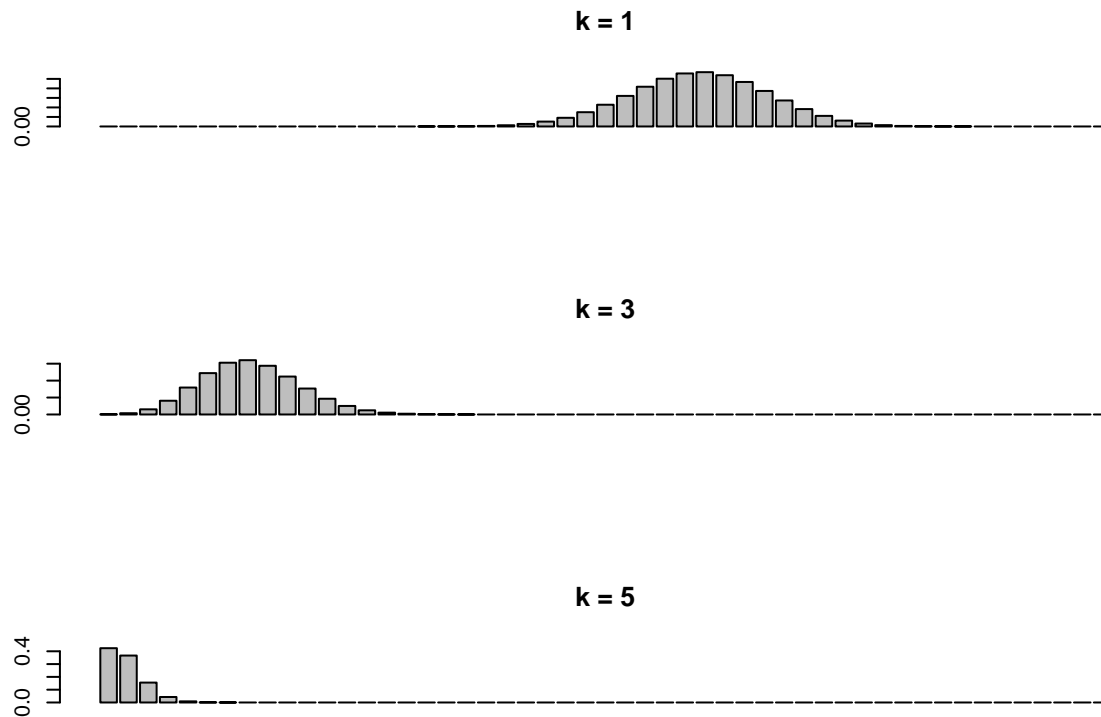
6 Example: Column from n x k table

```r
# explanition for the model

#when n pages are selected from 50 pages book, it follows binomial distribution.
#for calculation of possibility for one pages that has k misprints,
#it follows the poisson distribution.

table <- as.data.frame(matrix(ncol = 5,nrow=50, dimnames = list(c(1:50),c(1:5))))
for ( i in 1:50) {  #loop 1 for the whole 50 pages variable
  for (j in 1:5){ #loop 2 for the k = 1,2,3,4,5
```

```
    p <- ppois(j,lambda = 2,lower.tail = FALSE)
    table[i,j] = pbinom(i, size=50, prob=p)
  }
}

# to adjust the table contents
kableExtra::kable(digits = 3,x =table,row.names = TRUE)
```

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

```r
knitr::include_graphics("book.jpeg")
```