Homework # 5

Elliott Pryor

9/25/2020

Problem 1

Statement

3.121 - (clearly show steps to achieve a written probability statement, then write the code you would use to find the probability in R/find probability using R)

Let Y denote a random variable that has a Poisson distribution with mean $\lambda = 2$. Find

1. P(Y = 4)
2. P(Y $\geq$ 4)
3. P(Y < 4)
4. P(Y $\geq$ 4|Y $\geq$ 2)

Solution

$f_Y(y) = \frac{\lambda^y e^{-\lambda}}{y!}$

```
poisson_pmf <-function(y) {# generate a sequence of 100 independent Bernoulli trials
  p <- ((2^y) * exp(-2))/(factorial(y))
  return(p)}
```

1. $f_Y(4) = \frac{2^4 e^{-2}}{4!} = \frac{16e^{-2}}{24} = 2/3e^{-2}$

```
2/3 * exp(-2)
```

```
## [1] 0.09022352
```

2. This is the probability of $P(Y \geq 4) = 1 - P(Y < 4) = 1 - (P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3))$

```
1 - (poisson_pmf(0)  + poisson_pmf(1) + poisson_pmf(2) + poisson_pmf(3))
```

```
## [1] 0.1428765
```

3. $P(Y < 4) = P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3)$

```
(poisson_pmf(0)  + poisson_pmf(1) + poisson_pmf(2) + poisson_pmf(3))
```

```
## [1] 0.8571235
```

4. This is very similar to problem 2, but since we know that $Y \geq 2$ we can ignore the $P(Y = 0)$ and $P(Y = 1)$ terms. So we get $P(Y \geq 4|Y \geq 2) = 1 - (P(Y = 2) + P(Y = 3))$

```
1 - (poisson_pmf(2) + poisson_pmf(3))
```

```
## [1] 0.5488824
```

Problem 2

Statement

3.128

Cars arrive at a toll both according to a Poisson process with mean 80 cars per hour. If the attendant makes a one-minute phone call, what is the probability that at least 1 car arrives during the call?

Solution

This can be represented as $1 - P(\text{no cars arrive at the booth}) = 1 - f_Y(0)$. We must scale the unit effort from hour to minute, so $\lambda' = \lambda/60 = 4/3$ Following the Poisson pmf distribution $f_Y(y) = \frac{\lambda^y e^{-\lambda}}{y!}$ results in:

$$1 - \frac{\lambda^0 e^{-\lambda}}{0!} = 1 - e^{-\lambda} = 1 - e^{-4/3} = 0.7364$$

Problem 3

Statement

First, run the example R code provided, then answer the following question. On the same plotting region, plot the distribution of Y ~ Binomial(n, p) and W ~ Poisson(np) for p = 0.1 and varying values of n. Create a maximum of 3 plots and explain what you notice about Y and W as n $\rightarrow \infty$

Solution

```r
# set the largest y for which you want to plot P(Y = y)
max_y <- 50
y <- 0:max_y
```

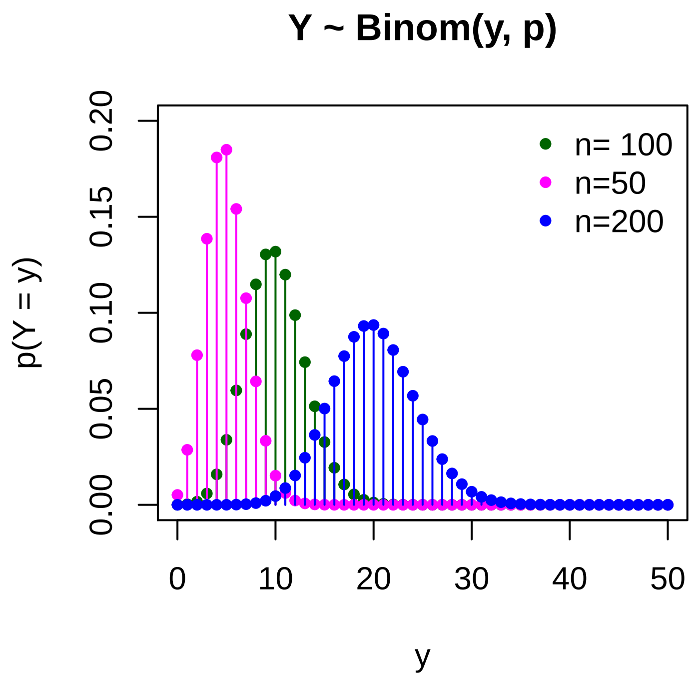
1. Binomial distribution

```r
v1 <- 100
v2 <- 50
v3 <- 200

plot(y , dbinom(y, v1, 0.1), type="h",
     xlim=c(0, max_y), ylim=c(0,0.2),
     xlab = "y", ylab = "p(Y = y)",
     main = "Y ~ Binom(y, p)",
     col = "darkgreen")
# add points and lines for the distribution of Y if Y ~ Geom(p) for different p
points(y, dbinom(y, v1, 0.1), pch = 20, col = "darkgreen")

lines(y, dbinom(y, v2, 0.1), type = "h", col = "magenta")
points(y, dbinom(y, v2, 0.1), pch = 20, col = "magenta")

lines(y, dbinom(y, v3, 0.1), type = "h", col = "blue")
points(y, dbinom(y, v3, 0.1), pch = 20, col = "blue")
# add a legend
legend("topright", legend = c(paste0("n= ", v1), paste0("n=", v2), paste0("n=", v3)),
       col = c("darkgreen", "magenta", "blue"),
       pch = c(20,20,20), bty = "n")
```




```r
# set the largest y for which you want to plot P(Y = y)
max_y <- 50
y <- 0:max_y
```

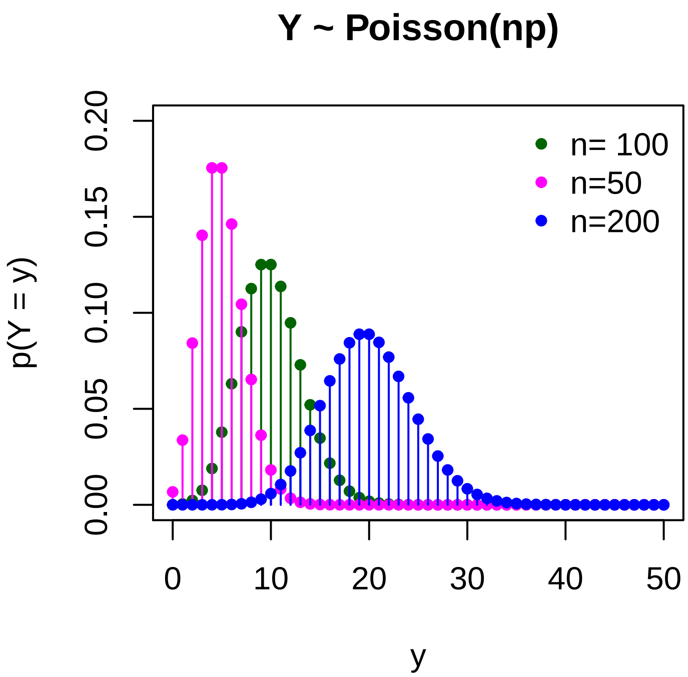
2. Poisson distribution

```r
v1 <- 100
v2 <- 50
v3 <- 200

plot(y, dpois(y, v1 * 0.1), type="h",
     xlim=c(0, max_y), ylim=c(0,0.2),
     xlab = "y", ylab = "p(Y = y)",
     main = "Y ~ Poisson(np)",
     col = "darkgreen")
# add points and lines for the distribution of Y if Y ~ Geom(p) for different p
points(y, dpois(y, v1 * 0.1), pch = 20, col = "darkgreen")

lines(y, dpois(y, v2 * 0.1), type = "h", col = "magenta")
points(y, dpois(y, v2 * 0.1), pch = 20, col = "magenta")

lines(y, dpois(y, v3 * 0.1), type = "h", col = "blue")
points(y, dpois(y, v3 * 0.1), pch = 20, col = "blue")
# add a legend
legend("topright", legend = c(paste0("n= ", v1), paste0("n=", v2), paste0("n=", v3)),
       col = c("darkgreen", "magenta", "blue"),
       pch = c(20,20,20), bty = "n")
```



As $n \to \infty$ we see that the distribution spreads out and shifts further to the right. This is easiest to analyze with the Poisson distribution. Since $\lambda = np$ then as $n \to \infty$ λ is increasing (hence the right shift as the mean moves to greater and greater values). Also since the variance $Var(Y) = \lambda = np$ as $n \to \infty$ the variance increases, hence why the distribution spreads out.

Extra Credit

Statement

3.85

Find E(Y(Y-1)) for a geometric random variable Y by finding $\frac{d^2}{dq^2}(\sum_{y=1}^{\infty} q^y)$. Use this result to find the variance of Y.

Solution

To start, we first give that $Var(Y) = E(Y(Y-1)) + E(Y) - E(Y)^2$.

We first need to find $E(Y(Y-1))$. We first start a couple math relations

$$f(x) = x^a \tag{1}$$

$$\frac{d}{dx}f(x) = ax^{a-1} \tag{2}$$

$$\frac{d^2}{dx^2}f(x) = a(a-1)x^{a-2} \tag{3}$$

$$
\begin{aligned}
E(Y(Y-1)) &= \sum_{y=1}^{\infty}(y(y-1))P(y) \\
&= \sum_{y=1}^{\infty}(y(y-1))q^{y-1}p \\
&= p\sum_{y=1}^{\infty}(y(y-1))q * q^{y-2} \\
&= qp\sum_{y=1}^{\infty}(y(y-1))q^{y-2} \qquad \text{By (3)} \\
&= qp\sum_{y=1}^{\infty}\frac{d^2}{dq^2}q^y \qquad \text{can factor } \frac{d^2}{dq^2} \text{ because 'certain conditons' hold} \\
&= qp\frac{d^2}{dq^2}(\sum_{y=1}^{\infty}q^y) \qquad \text{By geometric series} \\
&= qp\frac{d^2}{dq^2}(\frac{q}{1-q}) \\
&= qp(\frac{2q}{(1-q)^3} + \frac{2}{(1-q)^2}) \\
&= qp(\frac{2q+2(1-q)}{(1-q)^3}) \\
&= qp(\frac{2q+2(p)}{(p)^3}) \\
&= 2q\frac{q+p}{p^2} \\
&= \frac{2q}{p^2}
\end{aligned}
$$

Now we can solve for $Var(Y) = E(Y(Y-1)) + E(Y) - E(Y)^2 = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q+(p-1)}{p^2} = \frac{2q+(-q)}{p^2} = \frac{q}{p^2}$