Stats for DS HW 9

Matthew DeSantis

2022-11-09

6.2.6

(a)

E(phat1 - phat2) = E(phat1) - E(phat2) = (from law of large numbers) p1-p2

(b)

se = sqrt(var(phat1)+var(phat2)) = sqrt((p1(1-p1)/m)+(p2(1-p2)/n))

For the estimated se, just replace p1 and p2 with phat1 and phat2 respecitvely.

(c)

```r
X = 70
Y = 160
m = 100
n = 200
phat1 = X/m
phat2 = Y/n
phat1 - phat2
```

```
## [1] -0.1
```

```r
sehat = sqrt((phat1*(1-phat1)/m)+(phat2*(1-phat2)/n))
sehat
```

```
## [1] 0.05385165
```

6.3.4

(a)

MoM estimator for theta would be xbar/(sqrt(pi/2)) or xbar*(sqrt(2/pi)). It is unbiased, because E(thetahat) = E(xbar*sqrt(2/pi)) = sqrt(2/pi) * E(xbar) = sqrt(2/pi) * sqrt(pi/2) * theta = theta

6.3.6

(a)

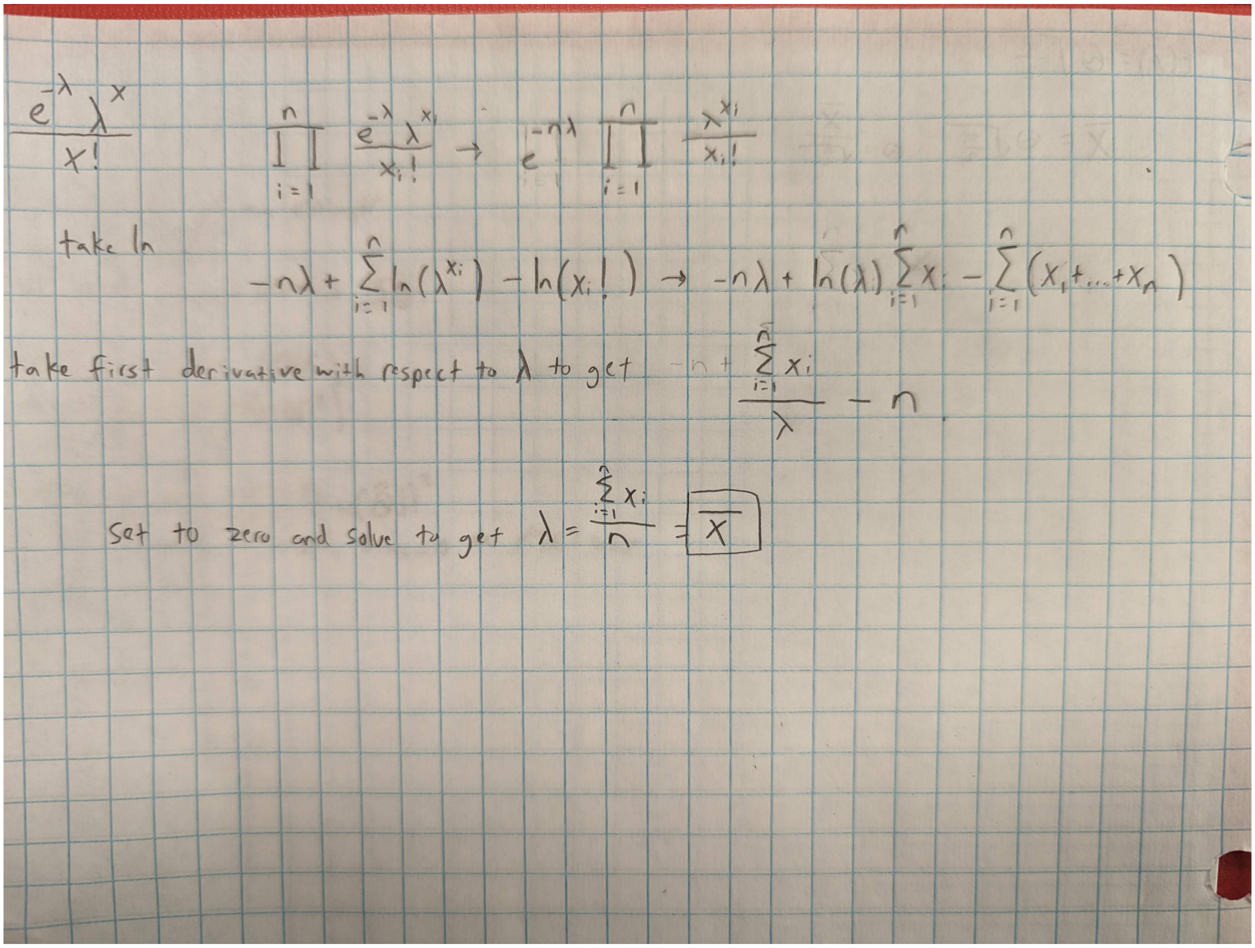


Figure 1: Solution for part a

(b)

```
x = 12+(2*11)+(3*14)+(4*9)
x/(4+12+11+14+9)
```

```
## [1] 2.24
```

(c)

Model-based population variance is xbar (poisson model has the same mean and variance), which is 2.24. The sample variance is

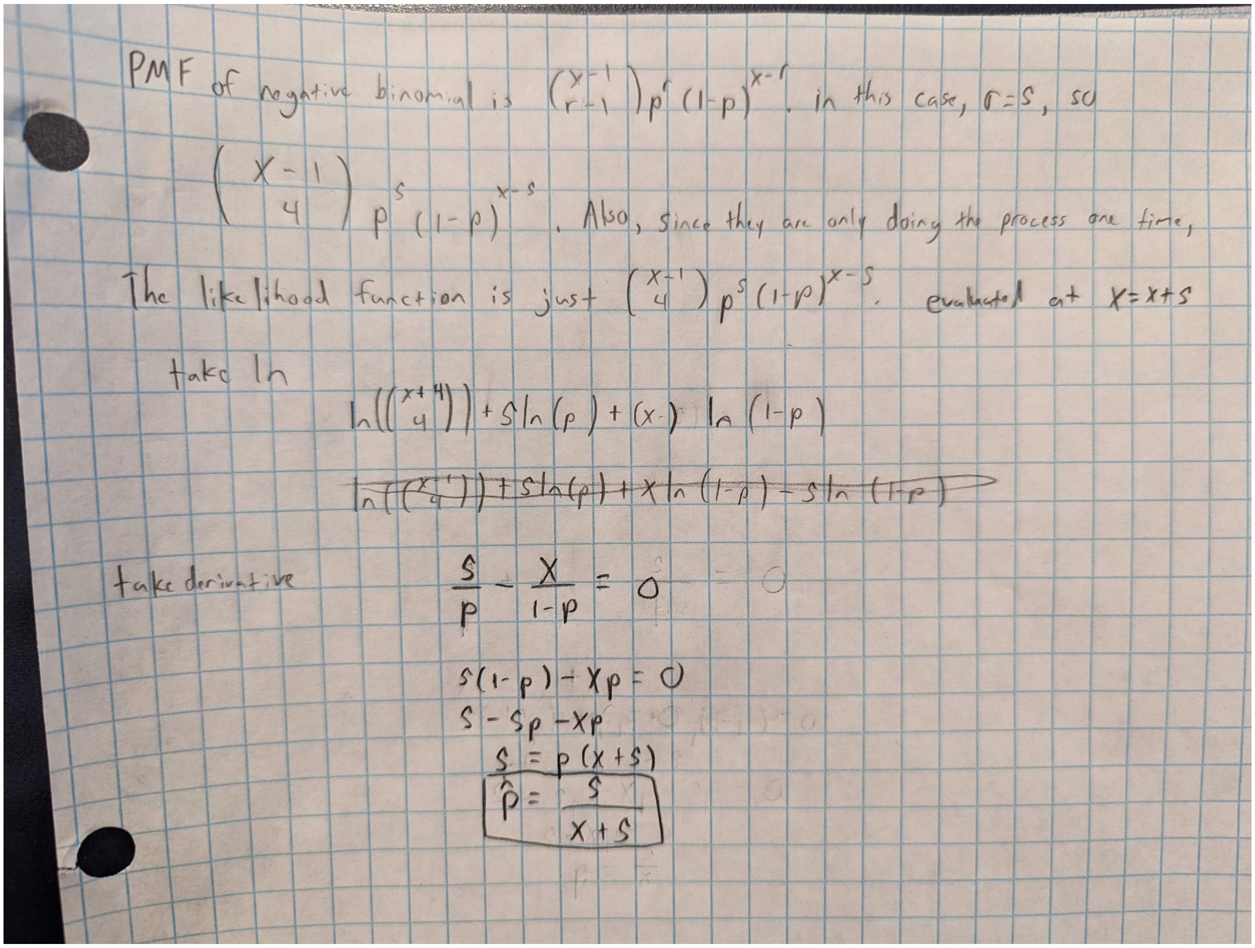
```
var(c(0,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,2,2,2,3,3,3,3,3,3,3,3,3,3,3,3,3,3,4,4,4,4,4,4,4,4
```

```
## [1] 1.533061
```

If we are assuming that Poisson accurately describes the population, we should prefer the model-based estimate.

6.3.7

(a)



PMF of negative binomial is $\binom{x-1}{r-1} p^r (1-p)^{x-r}$. in this case, $r=5$, so

$\binom{x-1}{4} p^5 (1-p)^{x-5}$. Also, since they are only doing the process one time,

The likelihood function is just $\binom{x-1}{4} p^5 (1-p)^{x-5}$. evaluated at $X=x+5$

take ln

$$\ln\left(\binom{x+4}{4}\right) + 5\ln(p) + (x)\ln(1-p)$$

$$\cancel{\ln\left(\binom{x-1}{4}\right) + 5\ln(p) + x\ln(1-p) - 5\ln(1-p)}$$

take derivative

$$\frac{5}{p} - \frac{x}{1-p} = 0$$

$$5(1-p) - xp = 0$$

$$5 - 5p - xp$$

$$5 = p(x+5)$$

$$\boxed{\hat{p} = \frac{5}{x+5}}$$

Figure 2: Solution for part a

(c)

```r
5/(47+5)
```

```
## [1] 0.09615385
```