Discrete Response Model Lecture 5

Models for Count Response, Discrete Response Model Evaluation, and Model Selection

datascience@berkeley

An Example (continue)

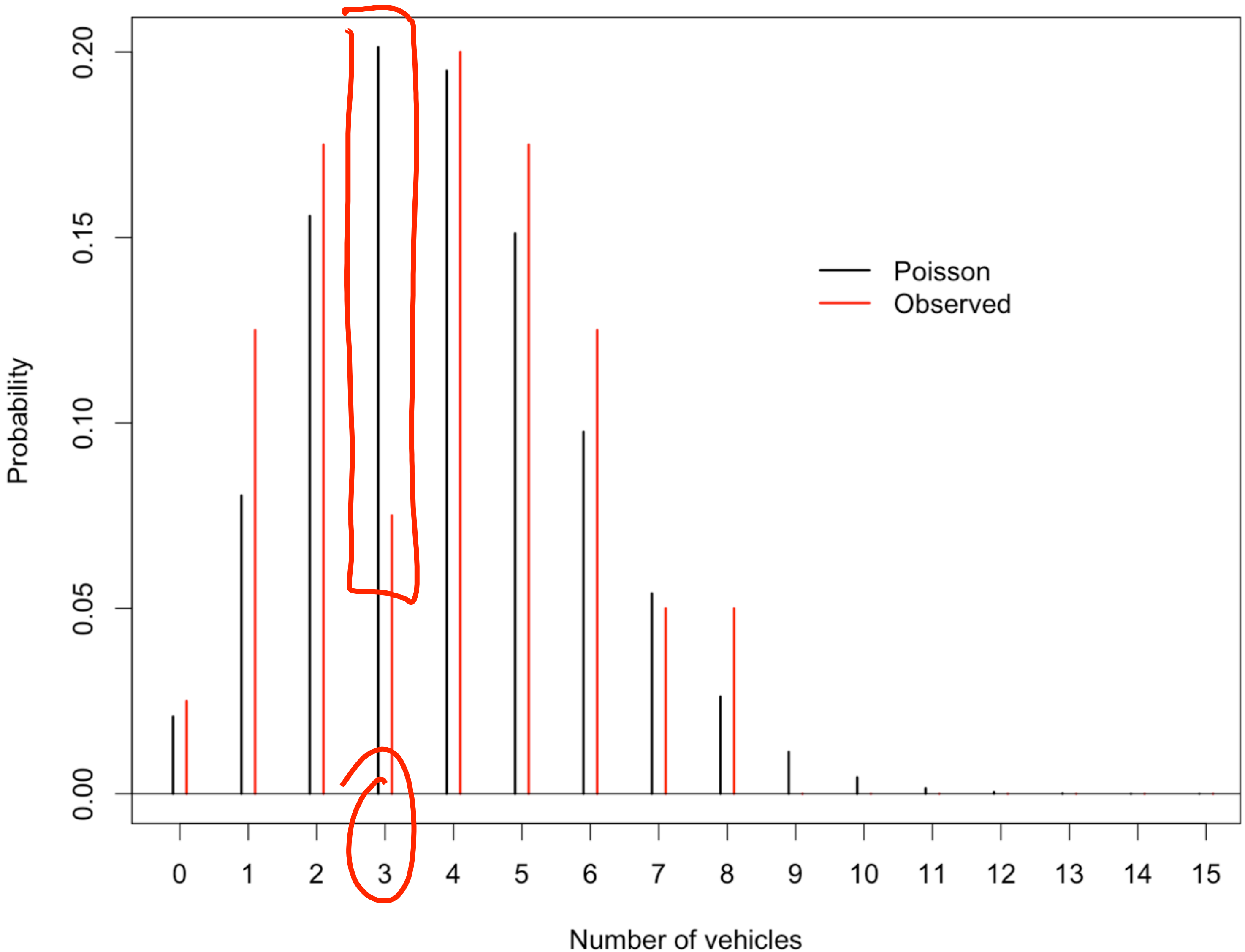
Example

```
> mean(stoplight$vehicles)
[1] 3.875
> var(stoplight$vehicles)
[1] 4.317308
> table(stoplight$vehicles) #Note that y = 0, 1, ..., 8 all have positive counts

0 1 2 3 4 5 6 7 8
1 5 7 3 8 7 5 2 2
```

```
> rel.freq <- table(stoplight$vehicles)/length(stoplight$vehicles)
> rel.freq2 <- c(rel.freq, rep(0, times = 7))
> y <- 0:15
> prob <- round(dpois(x = y, lambda = mean(stoplight$vehicles)), 4)
> data.frame(y, prob, rel.freq = rel.freq2)
    y   prob rel.freq
1   0 0.0208    0.025
2   1 0.0804    0.125
3   2 0.1558    0.175
4   3 0.2013    0.075
5   4 0.1950    0.200
6   5 0.1511    0.175
7   6 0.0976    0.125
8   7 0.0540    0.050
9   8 0.0262    0.050
10  9 0.0113    0.000
11 10 0.0044    0.000
12 11 0.0015    0.000
13 12 0.0005    0.000
14 13 0.0001    0.000
15 14 0.0000    0.000
16 15 0.0000    0.000
```

Example



Probability
0.00
0.05
0.10
0.15
0.20
Poisson
Observed
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Number of vehicles

Example

Wald confidence interval

```
> mu.hat <- mean(stoplight$vehicles)
> mu.hat + qnorm(p = c(alpha/2, 1 - alpha/2))*sqrt(mu.hat/n)
[1] 3.264966 4.485034
```

Note that the Wald interval using the $\log(\mu)$ transformation is

$$e^{\log(\hat{\mu}) \pm Z_{1-\alpha/2}\sqrt{1/(\hat{\mu}n)}}$$

Exponentiate the log() transformation.

```
> exp(log(mu.hat) + qnorm(p = c(alpha/2, 1-alpha/2)) * sqrt(1/(mu.hat*n)))
[1] 3.310561 4.535674
```

Score C.I.

```
> (mu.hat + qnorm(p = c(alpha/2, 1 - alpha/2))/(2*n)) + qnorm(p = c(alpha/2, 1 - alpha/2
)) * sqrt((mu.hat + qnorm(p = 1 - alpha/2)/(4*n))/n)
[1] 3.239503 4.510497
```

Berkeley SCHOOL OF INFORMATION