Conditional Example

Prof. Eric A. Suess

August 31, 2022

Particle Counter

A particle counter is imperfect and independently detects each incoming particle with probability p. If the distribution of the number of incoming particles in a unit of time is a Poisson distribution with parameter λ, what is the distribution of the number of counted particles?

Let $N = \#$ of incoming particles and $X = \#$ counted.

1. What is $P(X = k|N = n) =$?
2. What is $P(N = n)$?
3. Compute $P(X = k)$.

Is the conditional probability a regression?

Simulation of the conditional distribution of $X|N = n$.

The imperfect particle counter

```
lam <- 10
p <- 0.9
```

The conditional probability distributions for $n = 0, 1, 2, ..., 5$.

```
len <- 5
x <- matrix(0, 1+len, 1+len) # distributions down the columns of the matrix

for(i in 0:len){
       for(j in 0:i){
               x[1+j, 1+i] = dbinom(j, size=i, prob=p)
       }
}
```

```r
n <- c(0,1,2,3,4,5)

dimnames(x) <- list(n, n)

x
```

```
  0   1    2     3      4       5
0 1 0.1 0.01 0.001 0.0001 0.00001
1 0 0.9 0.18 0.027 0.0036 0.00045
2 0 0.0 0.81 0.243 0.0486 0.00810
3 0 0.0 0.00 0.729 0.2916 0.07290
4 0 0.0 0.00 0.000 0.6561 0.32805
5 0 0.0 0.00 0.000 0.0000 0.59049
```

The sums of each conditional distribution should be 1.

```r
x.sum <- apply(x, 2, sum) # column sum
x.sum
```

```
0 1 2 3 4 5
1 1 1 1 1 1
```

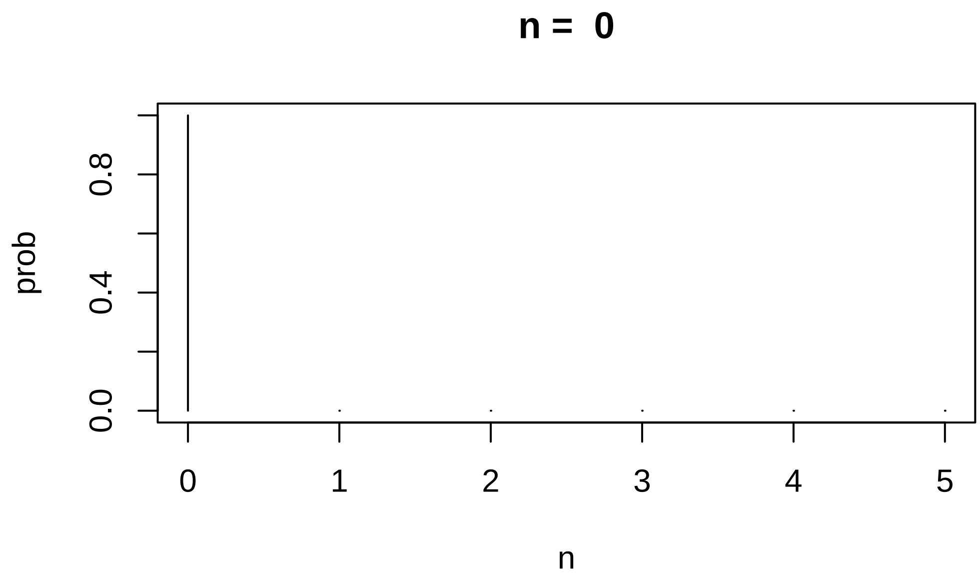
Determine the conditional mean $E[X|N = n]$ Column means.

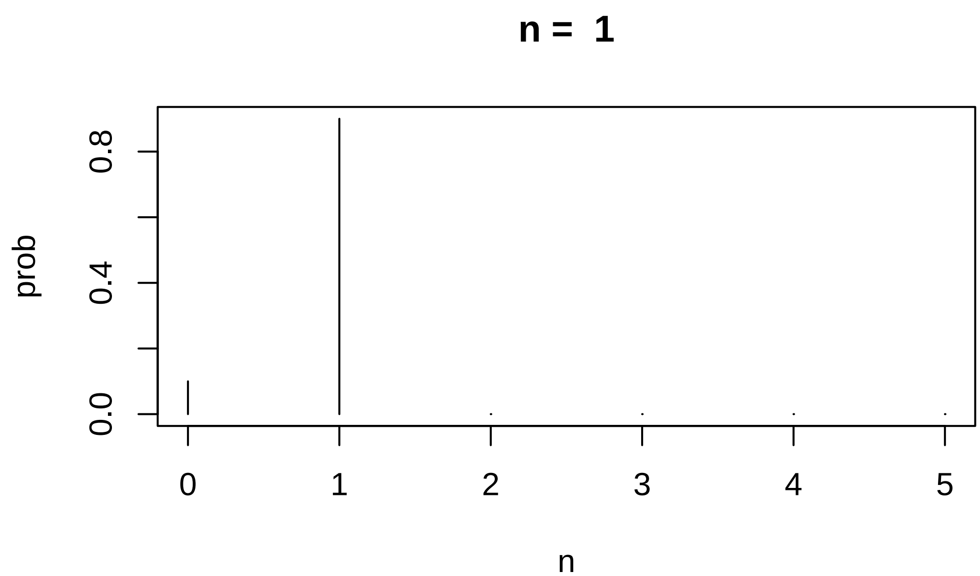
```r
x.mean <- 0

for(i in 1:len){
       x.mean <- c(x.mean, i*p)
}
x.mean
```

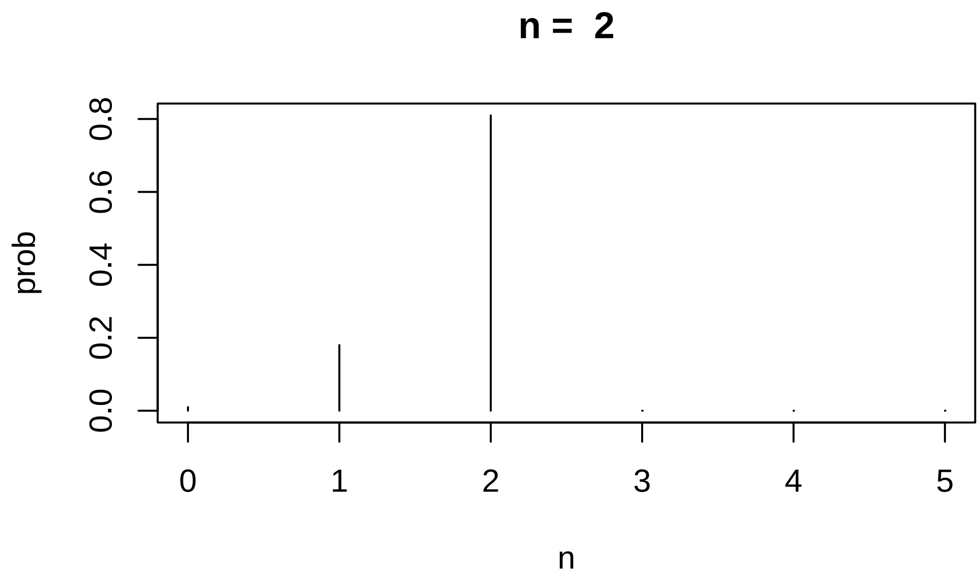
```
[1] 0.0 0.9 1.8 2.7 3.6 4.5
```

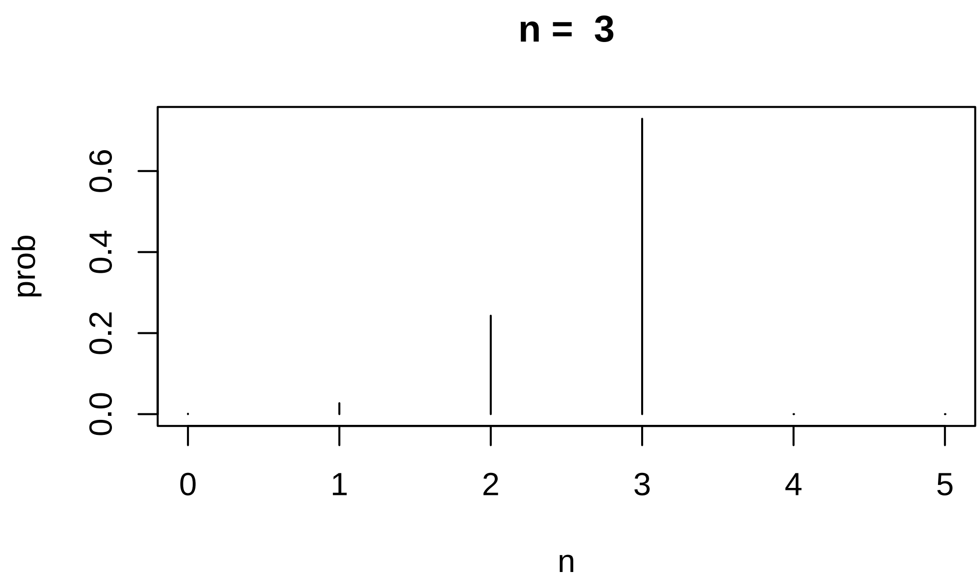
Plots of each probability distribution.

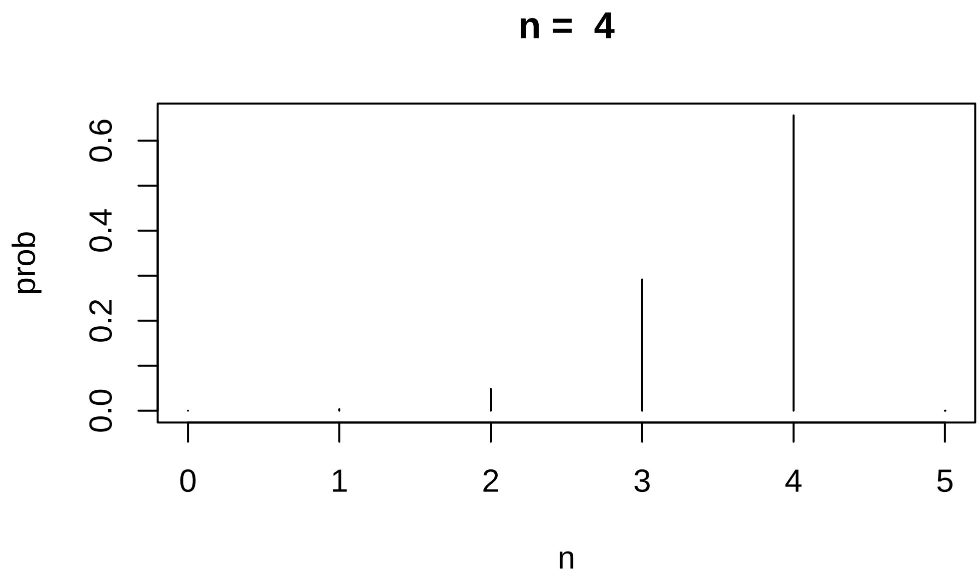
```r
for(i in 0:len){
       plot(n, x[,1+i], type="h", main = paste("n = ", i), ylab = "prob")
}
```

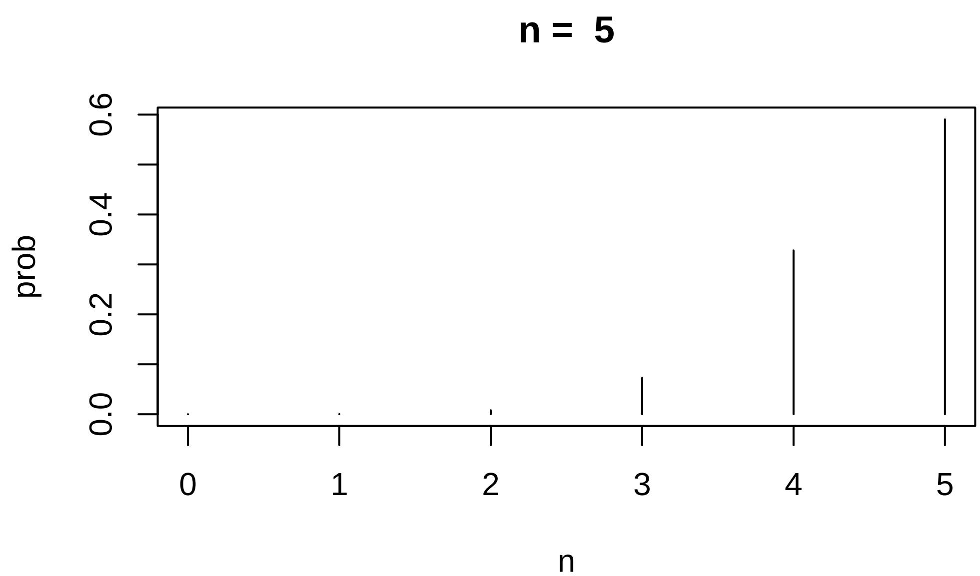

n = 0
prob
0.0
0.4
0.8
0
1
2
3
4
5
n


n = 1
prob
0.0
0.4
0.8
0
1
2
3
4
5
n


n = 2
prob
0.0
0.2
0.4
0.6
0.8
0
1
2
3
4
5
n


n = 3
prob
0.0
0.2
0.4
0.6
0
1
2
3
4
5
n



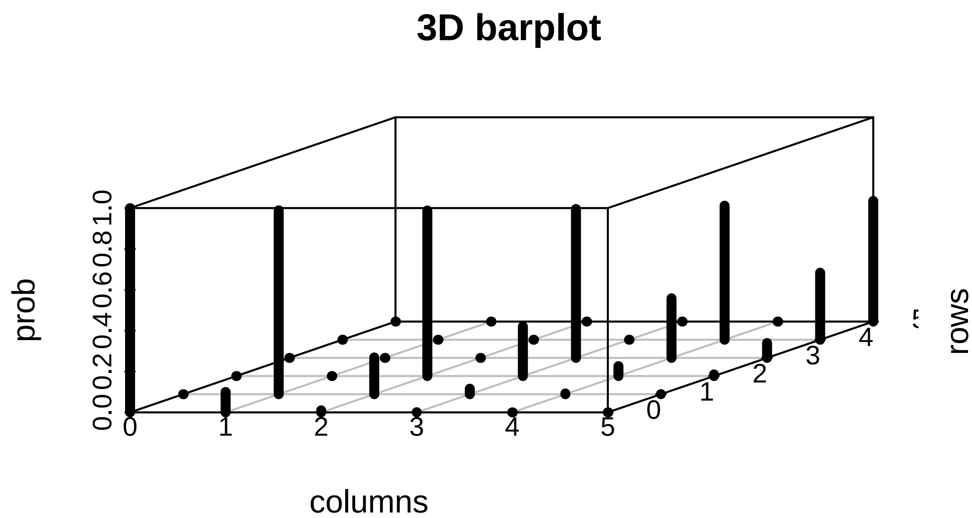



```
library(scatterplot3d)

x
```

```
  0   1    2     3      4       5
0 1 0.1 0.01 0.001 0.0001 0.00001
1 0 0.9 0.18 0.027 0.0036 0.00045
2 0 0.0 0.81 0.243 0.0486 0.00810
```

```
3 0 0.0 0.00 0.729 0.2916 0.07290
4 0 0.0 0.00 0.000 0.6561 0.32805
5 0 0.0 0.00 0.000 0.0000 0.59049
```

```
s3d.dat <- data.frame(columns = c(col(x)), rows = c(row(x)), prob = c(x))

scatterplot3d(s3d.dat, type = "h", lwd = 5, pch = " ",
              x.ticklabs = colnames(x), y.ticklabs = rownames(x),
              main = "3D barplot")
```



Simulation, how to estimate p using linear regression through the origin.

Simulate the data.

```
B <- 20000

n <- rpois(B,lam)

mean(n)
```

```
[1] 10.02725
```

```
x <- rbinom(n=B, size=n, prob=p)

mean(x)
```

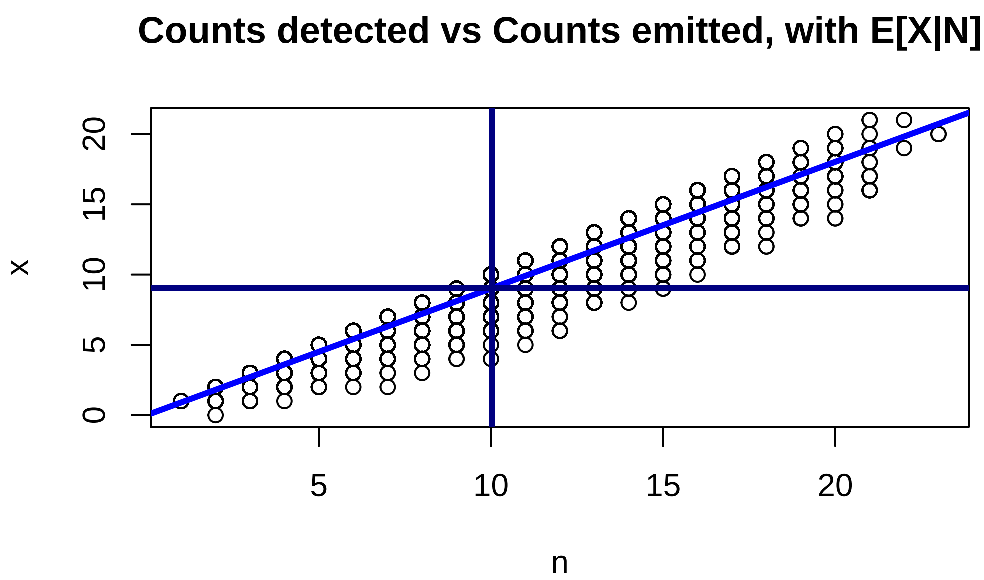
```
[1] 9.03505
```

Fit the model through the origin.

```r
plot(n,x,main="Counts detected vs Counts emitted, with E[X|N]")

x.fit = lm(x ~ 0+n)
summary(x.fit)
```

```

Call:
lm(formula = x ~ 0 + n)

Residuals:
    Min      1Q  Median      3Q     Max 
-5.0110 -0.5164  0.0879  0.7912  2.0770 

Coefficients:
   Estimate Std. Error t value Pr(>|t|)    
n 0.9010962  0.0006384    1412   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9488 on 19999 degrees of freedom
Multiple R-squared:  0.9901,	Adjusted R-squared:  0.9901 
F-statistic: 1.993e+06 on 1 and 19999 DF,  p-value: < 2.2e-16
```

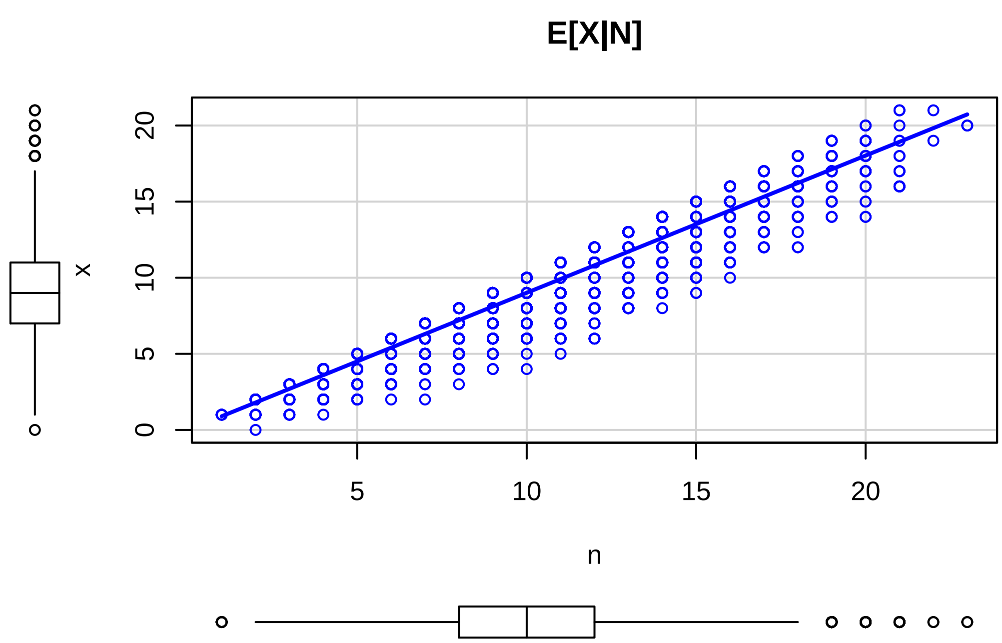
```r
abline(x.fit, col="blue", lwd = 3)         # E[X|N]
abline(v = mean(n), col="navy", lwd = 3)  # E[N]
abline(h = mean(x), col="navy", lwd = 3)  # E[X]
```




```r
library(car)
```

```
Loading required package: carData
```

```r
scatterplot(n, x, xlab="n", ylab="x", main="E[X|N]", smooth=FALSE)
```



Note that the estimated regression slope is very close to the true p.