ISYE 6420 Bayesian Statistics

Homework 2

September 28th, 2021

1.

Based on class, we know that the poisson distrubtion has the following characteristics for it probability mass function:

Probability mass function [edit]

A discrete random variable X is said to have a Poisson distribution, with parameter $\lambda > 0$, if it has a probability mass function given by:[

$$f(k;\lambda) = \Pr(X{=}k) = \frac{\lambda^k e^{-\lambda}}{k!},$$

where

- k is the number of occurrences ($k = 0, 1, 2\ldots$)
- e is Euler's number ($e = 2.71828\ldots$)
- ! is the factorial function.

The positive real number λ is equal to the expected value of X and also to its variance[3]

$$\lambda = \operatorname{E}(X) = \operatorname{Var}(X).$$

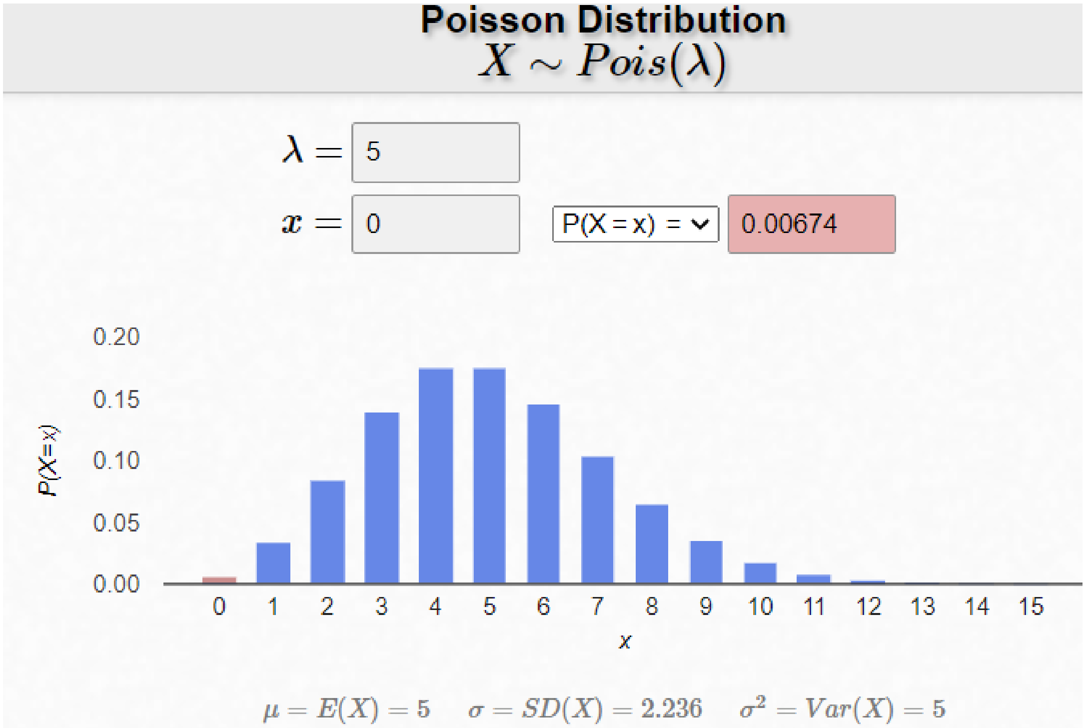
We know that our E(x) and Var(X) = 5, based on the above formula.

So we can derive the percentage of petri dishes containing the following number of clusters:

a) 0 Clusters:

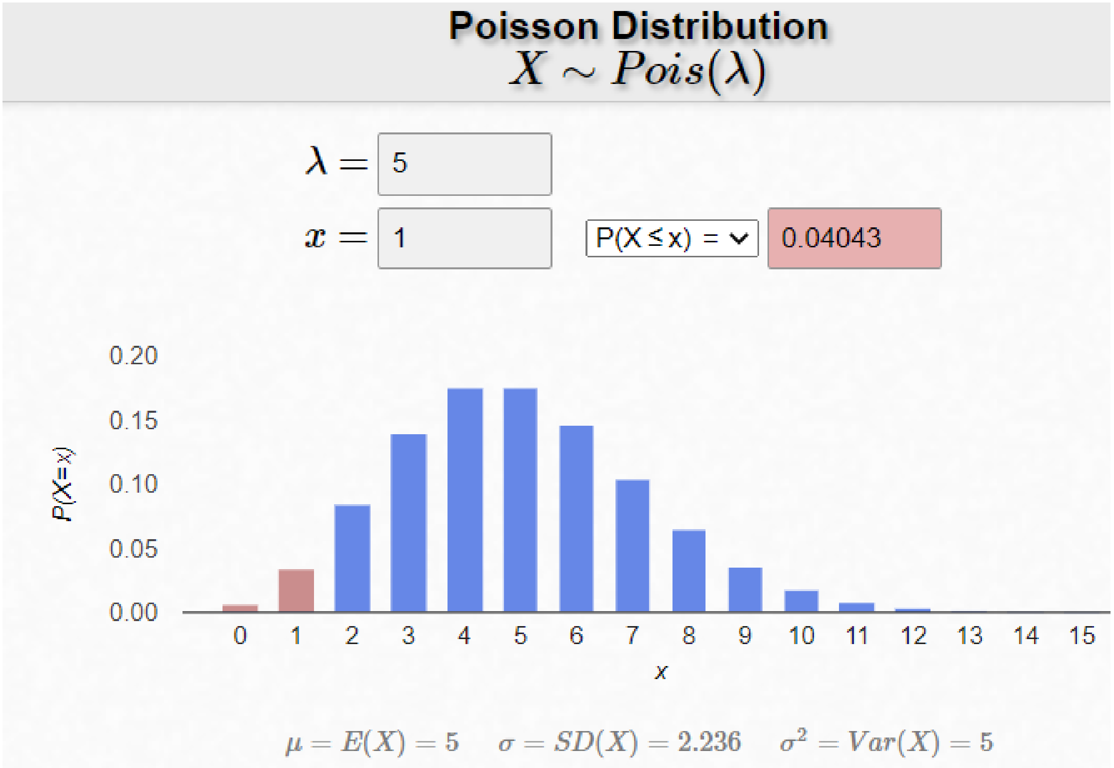
$$Pr(X = 0) \;=\; \frac{5^0 e^{-5}}{0!}$$
$$= 0.00673794699$$

Meaning that that the probablity of the petri dishes containing 0 clusters is 0.006



b) At least 1 cluster:

$$Pr(X \leq 1) = \frac{5^0 e^{-5}}{0!} + \frac{5^1 e^{-5}}{1!}$$
$$= 0.00673794699 + 0.0336897345$$
$$= 0.0404276814$$



c) More than 8 clusters

When we use our graphical representation, we can see that no probality is derived for the density function for values > 12.

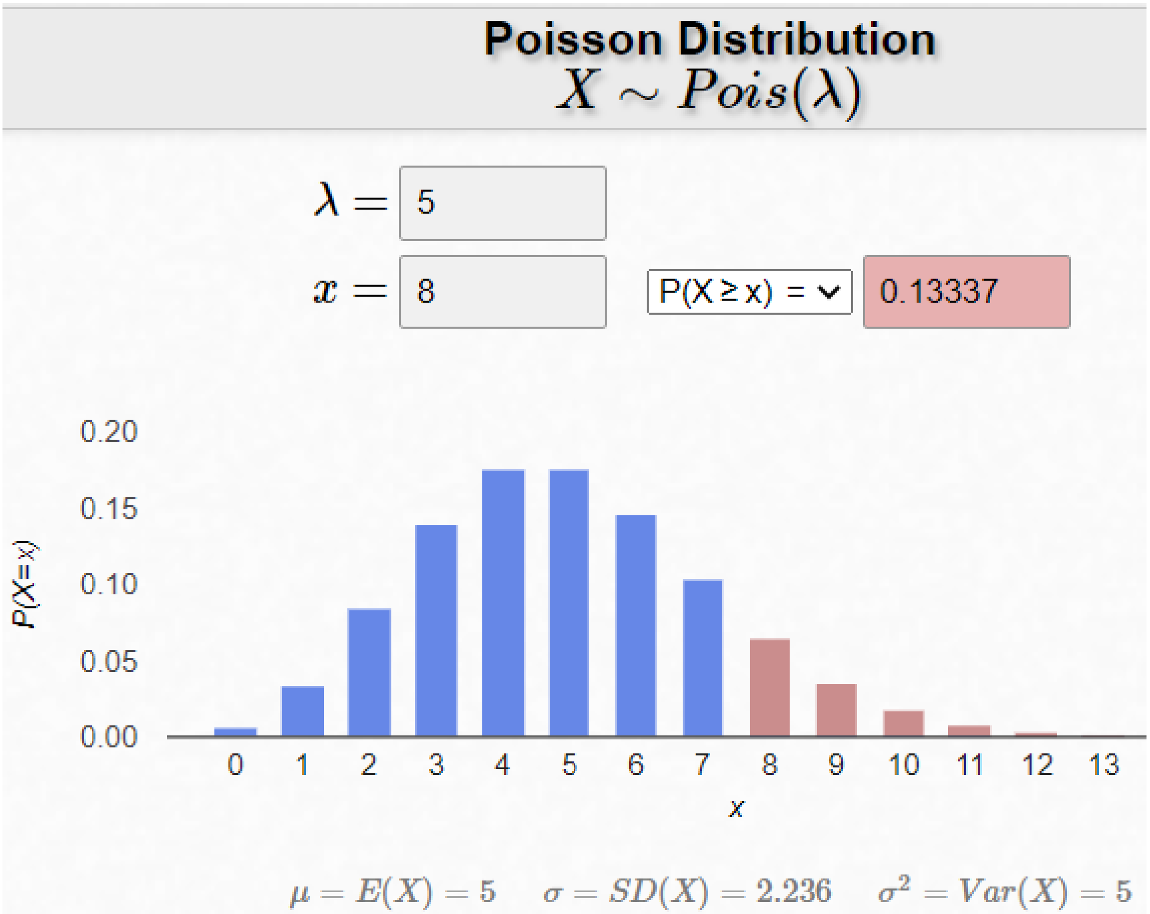


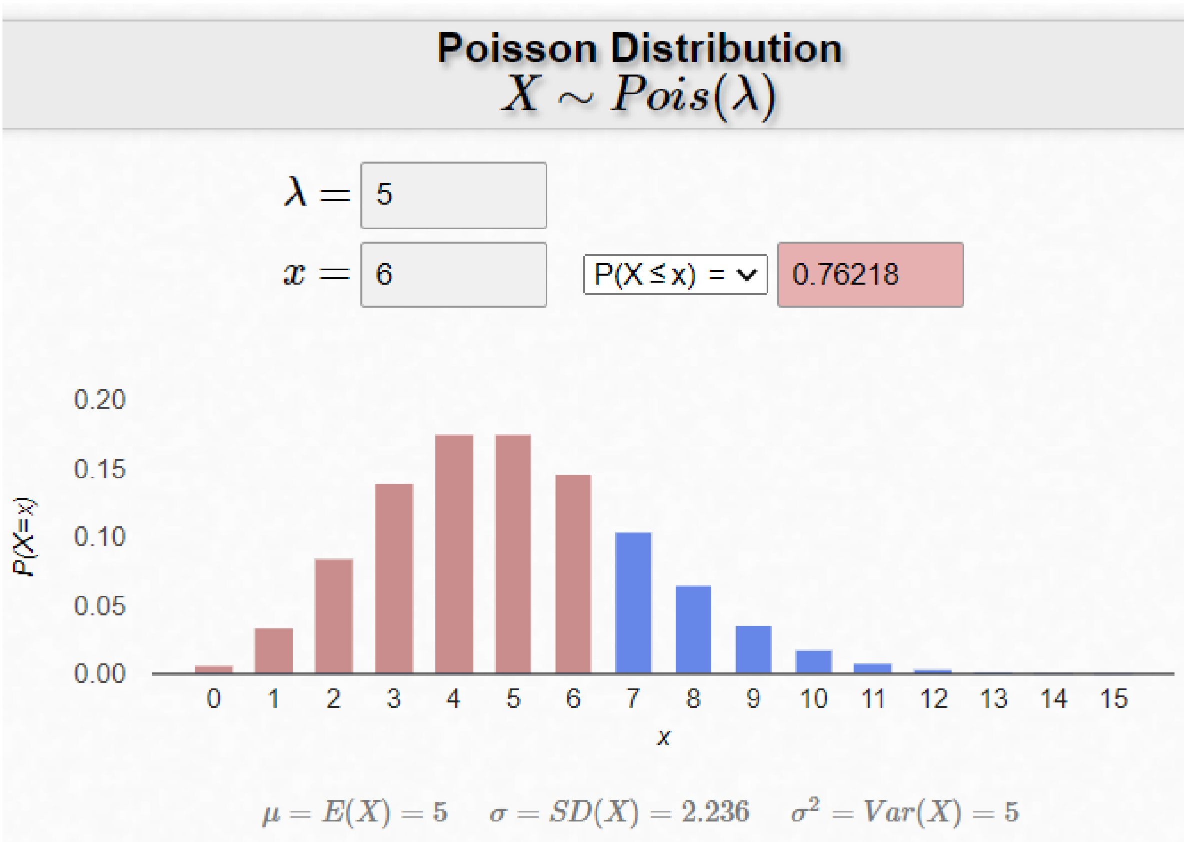
Figure 1: Probablity $\geqslant 8$

So from this we can approximate our probaility to be:

$$\begin{aligned} Pr(X \geq 8) \ - Pr(X = 8) \\ &= 0.13337 - 0.06528 \\ &= 0.06809 \end{aligned}$$

d)

Between 4 and 6 clusters:



$Pr(X \geq 4)\ AND\ Pr(X \leq 6) = Pr(X \leq 6) - Pr(X \leq 3)$

= 0.76218 - 0.26503

= 0.49715

2.

For this question can use the cumulative distribution plot for exponential in order to determine the probability of event X in our case, the probablity of a run continuing up interval an interval upper bound T.

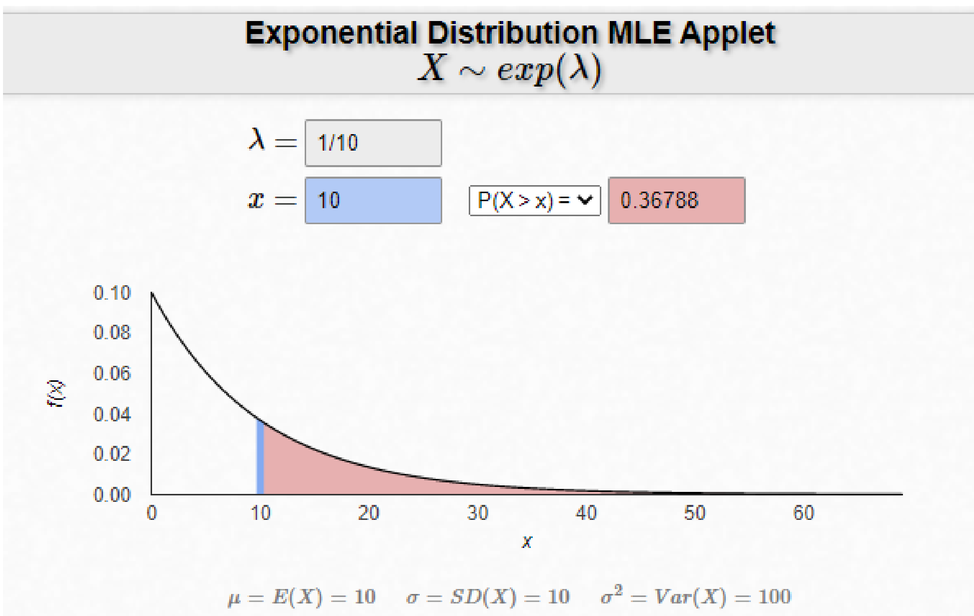
a) a run continues for at least 10 hours.

In this case for the cumulative distribution, the formula is as follows:

Cumulative distribution function [edit]

The cumulative distribution function is given by

$$F(x;\lambda) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

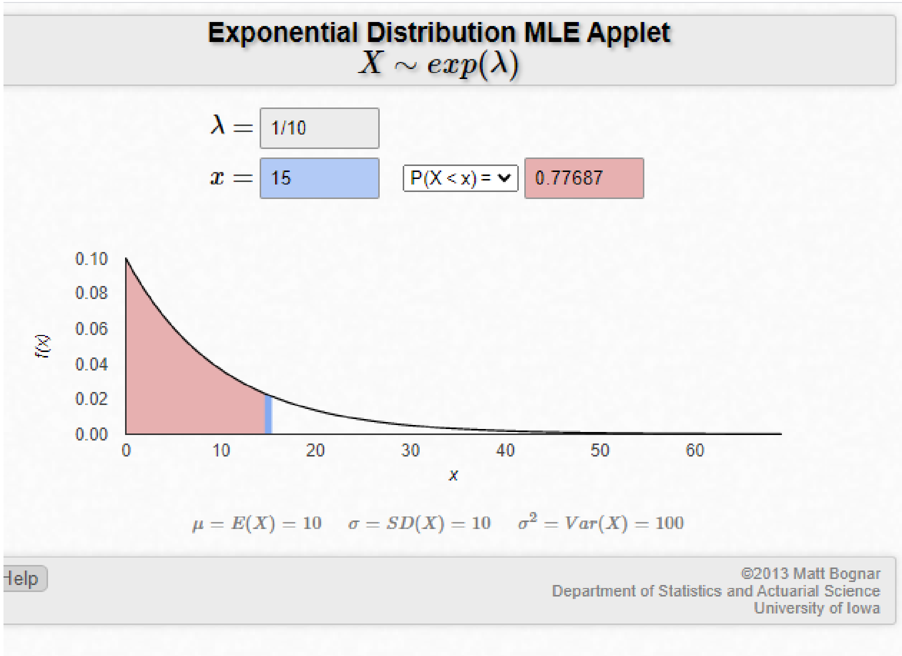


Since this is the probability densition function, we need to take the complement of this result.

$$
\begin{aligned}
Pr(T \geq 10) &= 1 - e^{-\lambda x} \\
&= 1 - e^{-1} \\
&= 1 - 0.36787944117 \\
&= 0.63212055882
\end{aligned}
$$

b)

A run lasts less than 15 hours:



Similarly to above, since we need to use the cumulative distribution we would take the complement of the above result to get:

$$
\begin{aligned}
Pr(T < 15) &= e^{-\lambda x} \\
&= e^{-1.5} \\
&= 0.22313016014
\end{aligned}
$$

c)
A run continues for at least 20 hours, given that it has lasted 10 hours:

Using the cumulative distribution we can forumulate this problem as follows:

The probability you want is $Pr(T > 20 \mid T > 10)$ where X has an exponential distribution. We're seeing whether it's greater than 20 because you want to know if it lasts 10 additional years, after it's already been functioning for 10 years, with the rate parameter $\frac{1}{10}$:

$$
\begin{aligned}
Pr(T \leq 20 \mid T > 10) &= \frac{Pr(T \leq 20 \cap T > 10)}{Pr(T > 10)} \\
&= \frac{Pr(10 < X \leq 20)}{Pr(T > 10)}
\end{aligned}
$$

$$= \frac{e^{-10\lambda} - e^{-20\lambda}}{e^{-10\lambda}}$$

$$= 1 - e^{-10\lambda}$$
$$= 0.63212055882$$

Note that this is the same as answer a due to the memoryless property.

3.

a)

$$f(x, y) = \begin{cases} \lambda^2 e^{-\lambda y}, & 0 \le x \le y,\ \lambda > 0 \\ 0, & \text{else} \end{cases}$$

From the following, we know that we'll need to calculate the following to get the marginal distribution:

$$\int_x^\infty \lambda^2 e^{-\lambda y} dy \tag{1}$$

When we factor out lambda square as a constant then we can seperate out the term with y to get the integral:

$$\lambda^2 \int_x^\infty e^{-\lambda y} dy \tag{2}$$

This result in the following:

$$\left(-\lambda e^{-\lambda \infty}\right) \ - \ \left(-\ \lambda e^{-\lambda x}\right) \tag{3}$$

From this we know any term to infinity is 0, therefore leaving the following result which is the exponential distribution:

$$\lambda e^{-\lambda x} \tag{4}$$

b)

For the next part, we're first going to need to take the integral with respect to x for the range of 0 to y:

$$\int_0^y \lambda^2 e^{-\lambda y} dx \tag{5}$$

$$\int_0^y \lambda^2 x \, e^{-\lambda y} dx \tag{6}$$

After taking the integral for this range we derive the following:

$$\lambda^2 y e^{-\lambda y} \tag{7}$$

Now to resemble the gamma function in the denomiator which is equal to $(n-1)!$

When we pair up the terms above into the following equation we can see our $\alpha = 2,$ which yields the following:

$$\Gamma(2) = 1$$

$$\frac{\beta^2 y^{2-1} e^{-\beta y}}{\Gamma(\alpha)} \tag{8}$$

= Ga(2,λ)

c)

conditional distribution $f(y|x)$ is shifted exponential, $f(y|x) = \lambda e^{-\lambda(y-x)}$, $y \geq x$.

When we take the marginal probablity found from a), over the join probability from our question in the numerator to get the following conditional probaility, then simplifying we get the following:

$$f(y|x) = \frac{\lambda e^{-\lambda y}}{\lambda e^{-\lambda x}} \tag{9}$$
$$= \lambda e^{-\lambda(y-x)}$$

d)

When we take the marginal probablity found from a), over the join probability from our question in the numerator to get the following conditional probaility, then simplifying we get the following:

$$f(x|y) = \frac{\lambda e^{-\lambda y}}{\lambda e y^{-\lambda y}} \tag{10}$$
$$= \frac{1}{y}$$

$$Since\ y > 0\ and\ x > 0,\ we\ know\ that\ this\ will\ result\ in\ a\ value\ between\ 0\ and\ 1,\ and\ therefore\ U(0,1)$$

4

a)

In order to determine λ, a common approach would be to MLE (Maximum Likelihood Estimator), in order to deterive our sample mean which would work out to 1/x. This is solveable since it's a closed-form solution compared to MLE that is used for GLM models like logistic regression.

b)

We know the posterior probability follows this form:

$$Post(\lambda) = \frac{Pr(\lambda) * \mathscr{L}(\lambda)}{Marginal\ Dist}$$

In our case since we don't need to calculate the integral for the marginal distribution, we can use the following:

$$\frac{1}{\sqrt{\lambda}} * \lambda^n e^{-\lambda \sum_i^n x_i} \tag{11}$$

Based on the question we know that the summation of xi = 24 where n = 3

$$\sum_i^3 x_i = 24$$

From this we can derive the following:

$$\lambda^{3-\frac{1}{2}} e^{-24\lambda} \tag{12}$$

$$\lambda^{3.5-1} e^{-24\lambda}$$

Therefore, the bayes estimator yields $G(3.5,\ 24)$