Poisson probability formula

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

How likely were we to see the exact sample we saw?

$$\Pr\left(Y_1 = y_1, Y_2 = y_2, \ldots Y_n = y_n | \lambda\right)$$

$$\Pr\left(Y_1 = y_1 | \lambda\right) \cdot \Pr\left(Y_2 = y_2 | \lambda\right) \cdot \cdots \cdot \Pr\left(Y_n = y_n | \lambda\right)$$

$$\frac{e^{-\lambda} \cdot \lambda^{y_1}}{y_1!} \cdot \frac{e^{-\lambda} \cdot \lambda^{y_2}}{y_2!} \cdot \cdots \cdot \frac{e^{-\lambda} \cdot \lambda^{y_n}}{y_n!}$$

Reminder the little y_i are data observed in the world.

$$\mathcal{L} = \frac{e^{-\lambda} \cdot \lambda^{y_1}}{y_1!} \cdot \frac{e^{-\lambda} \cdot \lambda^{y_2}}{y_2!} \cdot \cdots \cdot \frac{e^{-\lambda} \cdot \lambda^{y_n}}{y_n!}$$

Taking log converts multiplication problems into additional problems

$$\ln \mathcal{L} = \ln \frac{e^{-\lambda} \cdot \lambda^{y_1}}{y_1!} + \ln \frac{e^{-\lambda} \cdot \lambda^{y_2}}{y_2!} + \cdots \ln \frac{e^{-\lambda} \cdot \lambda^{y_n}}{y_n!}$$

$$\ln \mathcal{L} = \ln \frac{e^{-\lambda} \cdot \lambda^{y_1}}{y_1!} + \ln \frac{e^{-\lambda} \cdot \lambda^{y_2}}{y_2!} + \cdots \ln \frac{e^{-\lambda} \cdot \lambda^{y_n}}{y_n!}$$

Zoom in on one of the elements:

$$\ln \mathcal{L}_1 = \ln \frac{e^{-\lambda} \cdot \lambda^{y_1}}{y_1!} = \ln e^{-\lambda} + \ln \lambda^{y_1} - \ln y_1! =$$

$$-\lambda + y_1 \ln \lambda - \ln(y_1) - \ln(y_1 - 1) - \ln(y_1 - 2) - \cdots - 1$$

Zoom back out (there are n of these expressions to add up):

$$\ln \mathcal{L} = -n\lambda + (y_1 + y_2 + \cdots + y_n) \ln \lambda$$

$$\begin{aligned}
&- \ln(y_1) - \ln(y_1 - 1) - \ln(y_1 - 2) - \cdots - 1 \\
&- \ln(y_2) - \ln(y_2 - 1) - \ln(y_2 - 2) - \cdots - 1 \\
&- \ln(y_n) - \ln(y_n - 1) - \ln(y_n - 2) - \cdots - 1
\end{aligned}$$

The second batch of stuff changes the level of $\ln \mathcal{L}$ but does not depend on λ. Only the first batch does:

$$\ln \mathcal{L} = -n\lambda + (y_1 + y_2 + \cdots + y_n) \ln \lambda$$

$$\begin{aligned} &- \ln(y_1) - \ln(y_1 - 1) - \ln(y_1 - 2) - \cdots - 1 \\ &- \ln(y_2) - \ln(y_2 - 1) - \ln(y_2 - 2) - \cdots - 1 \\ &- \ln(y_n) - \ln(y_n - 1) - \ln(y_n - 2) - \cdots - 1 \end{aligned}$$

$$\max_{\lambda} \ln \mathcal{L} = -n\lambda + (y_1 + y_2 + \cdots + y_n) \ln \lambda$$

$$-n + \frac{1}{\hat{\lambda}}(y_1 + y_2 + \cdots + y_n) \equiv 0$$

$$\hat{\lambda} = (y_1 + y_2 + \cdots + y_n)/n$$