Bayesian inference

Example: Two different machines can stamp cloth. Machine 1 makes an average of $\lambda_1 = 5$ errors per yard and Machine 2 makes an average of $\lambda_2 = 2$ errors per yard.

A sample of 1 yard is taken from a roll and a total of $X = 3$ errors are found. What machine produced the roll.

$X | \lambda \sim \text{Poisson}(\lambda)$

$P(X | \lambda)$

What we would like is to have

$p(\lambda | X) \propto p(X | \lambda) p(\lambda)$

Posterior

$p(\lambda | X) = \frac{p(X | \lambda) p(\lambda)}{p(X)}$

Prior $p(\lambda) = \begin{cases} 
\frac{1}{2} & \lambda = 2 \\
\frac{1}{2} & \lambda = 5 
\end{cases}$

Assuming that both machines produce about the same number of rolls
\[ p(\lambda | x) = \frac{p(x|\lambda) p(\lambda)}{p(x)} \quad x = 3 \]

\[ p(x) = p(x=3 | \lambda = 2) p(\lambda = 2) + p(x=3 | \lambda = 5) p(\lambda = 5) \]

\[ \Rightarrow \text{total probability law} \]

\[ \Rightarrow \text{"marginal likelihood"} \]

\[ \Rightarrow \text{"prior predictive"} \]

\[ p(x=3 | \lambda) = \frac{e^{-\lambda} \lambda^3}{3!} \]

\[ p(x = 3) = \frac{e^{-2} 2^3}{3!} \frac{1}{2} + \frac{e^{-5} 5^3}{3!} \frac{1}{2} \]

\[ = \frac{1}{12} \left[ e^{-2} 2^3 + e^{-5} 5^3 \right] \]

\[ p(\lambda | x = 3) = \begin{cases} 
  p(\lambda = 2 | x = 3) \\
  p(\lambda = 5 | x = 3) 
\end{cases} \]

\[ p(\lambda = 2 | x = 3) = \frac{e^{-2} 2^3 / 12}{e^{-2} 2^3 / 12 + e^{-5} 5^3 / 12} = 0.5624 \]
If instead you had observed $x=0$ errors

$$p(x=0) = \frac{e^{-2} \cdot 2^{0 \cdot \frac{1}{2}}}{e^{-2} \cdot 2^{0 \cdot \frac{1}{2}} + e^{-5} \cdot 5^{0 \cdot \frac{1}{2}}}$$

$$= \frac{e^{-2}}{e^{-2} + e^{-5}} = 0.9525$$

More generally, you could have $p(\lambda)$ to be a continuous distribution.

For example $p(\lambda) = \frac{1}{5} \exp(-\frac{\lambda}{5}) \quad \lambda \geq 0$

$\lambda \sim \text{Exp}(5)$

$$p(x|\lambda) = \frac{e^{-\lambda} \lambda^x}{x!}$$

$$p(\lambda|x) \propto \int_0^\infty \frac{e^{-\lambda} \lambda^x}{5} \exp(-\frac{\lambda}{5}) \, d\lambda$$

$$= \left( \frac{x e^{-\lambda (1 + \frac{1}{5})}}{\int_0^\infty \lambda^x e^{-\lambda (1 + \frac{1}{5})} \, d\lambda} \right) \quad \text{looks like the Gamma}(x+1, (1+\frac{1}{5}))$$
\[ X = 3 \]
\[ \lambda | x = 3 \sim \text{Gamma}(4, 5/6) \]
\[ E(\lambda | x = 3) = 4 \times \frac{5}{6} = \frac{10}{3} \approx 3.33 \]

Where is the prior coming from?
- Previous knowledge/studies
- "Non-informative" priors
- "Flat" priors
- Who cares?

Example: You want to infer the probability that a coin comes up heads. You toss it 10 times and it comes up heads 6 times.

\[ X \sim \text{Bin}(10, \theta) \] what you are interested in