Questions and Exercises
for Astro Hack Week 2015

Brendon J. Brewer

October 1, 2015

1 Probability theory

Here are some equations of probability theory. Which rules are being used in each? Is each equation completely general, or has something specific been assumed (e.g. independence, mutual exclusivity)?

$$
\begin{aligned}
P(A, B|C) &= P(A|C)P(B|C) & (1)\\
P(A \vee B|C) &= P(A|C) + P(B|C) & (2)\\
P(A, B|C, D) &= P(A, B|D)P(C|A, B, D)/P(C|D) & (3)\\
P(Z|I) &= \sum_i P(A_i|I)P(Z|A_i, I) & (4)
\end{aligned}
$$

For an interesting physicsy way of getting the hang of writing correct probability expressions, see arxiv: 1205.4446 by Hogg. Be aware that his arguments about units apply to probability density functions, whereas our probabilities above are all dimensionless. Nevertheless, both satisfy rules of very similar form.

2 Bayes' Rule

The most useful form of Bayes' rule is

$$
P(H_i|D) = \frac{P(H_i)P(D|H_i)}{P(D)} \tag{5}
$$

where

$$
P(D) = \sum_{i=1}^{N} P(H_i)P(D|H_i). \tag{6}
$$

H_i can be interpreted as a hypothesis and D as data. Give an intuitive explanation for why Equation 6 is true.

3 Discrete parameter estimation

You are a customs agent. Among other things, you are supposed to prevent drugs being smuggled in packages sent into the country. A colleague finds a package containing two toys. Let the number of toys containing drugs be η. Clearly the value of η is either 0, 1, or 2.

You drill into one of the toys and find that it does not have drugs (call this proposition D for data). Calculate the posterior distribution for η given that the tested toy did not contain drugs. Assume a (discrete) uniform prior, i.e. $P(\eta = 0) = P(\eta = 1) = P(\eta = 2) = 1/3$.

Hint: Use Bayes' rule three times, in the form $P(\eta = i|D) = \frac{P(\eta=i)P(D|\eta=i)}{P(D)}$, where the denominator is the same each time.

4 Continuous parameter estimation

An X-ray source emits photons at a steady rate, but since it's so distant, we only pick up a few photons per second. A standard probabilistic model for the arrival times of the photons is the "Poisson process". A specific prediction of this model is that, if the expected number of photons in a time interval is μ, the (discrete) probability distribution for the actual number of photons x in that interval is a Poisson distribution:

$$p(x|\mu) \quad = \quad \frac{\mu^x e^{-\mu}}{x!} \tag{7}$$

Find and plot the posterior distribution for μ given $x = 5$. Use an improper log-uniform prior proportional to μ^{-1}.

5 Prediction

Use the posterior distribution obtained in the previous question to calculate the predictive distribution for x', the number of photons observed in a different one second interval, given $x = 5$. Plot the predictive distribution and compare it with what you'd get by naively assuming the best fit (maximum likelihood) value $\mu = 5$ to make the prediction.

Hint 1: Write down the probability distribution for x' given μ and x, and then marginalise out μ.

Hint 2: You may need to do some numerical integration.