Worksheet 17 (Solutions)

1. Consider a prior distribution $p \sim Beta(\alpha, \beta)$ for some fixed α and β for a likelihood given by $X|p \sim Bin(n, p)$. Derive the posterior distribution $p|X$.

Solution: The posterior comes from the formula in the notes, with the one change that we now have a different density of the prior distribution $f_p(p)$:

$$\begin{aligned} f_{p|X}(p|x) &\propto f_{X|p}(x|p) \times f_p(p) \\ &\propto \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} \cdot p^{\alpha-1} \cdot (1-p)^{\beta-1} \\ &\propto p^{x+\alpha-1} \cdot (1-p)^{n-x+\beta-1} \end{aligned}$$

And so, therefore, we have:

$$p|x \sim Beta(x+\alpha, n-x+\beta).$$

And that's it!

2. For reasons that we will explore in more next time, the Bayesian point estimator (the best single-number estimator) is the expected value of the posterior distribution. Under the set up from the previous question, what is the Bayesian point estimator $\widehat{p}$ in terms of X, α, and β?

Solution: From the distributions table, we see that the expected value is given by:

$$\frac{x+\alpha}{x+\alpha+n-x+\beta} = \frac{x+\alpha}{n+\alpha+\beta}.$$

3. Consider observing $X \sim Bin(n, p)$. We know that the MLE estimator of p is given by $\hat{p}_{MLE} = X/n$. The Binomial comes from doing n Bernoulli trials and adding the number of 1s. Consider creating a new Y in which we artificially augment the data X by adding (in effect) an extra 0 and an extra 1. In other words, we create a $Y = X + 1$ with the assumption that $Y \sim Bin(n+2, p)$. What is the MLE of p using the data from the augmented data Y? Where have you seen this before?

Solution: The MLE will be given by Y divided by $n+2$ (the analog of the case for X), which becomes: $(X+1)/(n+2)$. This is just the previous solution with $\alpha = \beta = 1$.

4. Consider your solution to the previous set of questions. If α and β are non-negative integers, how could you describe the Bayesian estimator based on adding data to X?

Solution: In general, if we add α 1s and β 0s to the data, we would get the MLE equal to the Bayesian estimator.

5. The standard uniform distribution is equivalent to $Beta(1, 1)$. In the notes I started by implicitly assuming that this was a fairly neutral starting assumption for indicating that we do not have any strong prior knowledge of p. Based on your results above, what would actually seem to be the best netural position if we do not want the prior to have a strong influence on the posterior mean?

Solution: In order to have an estimator where the Bayesian estimator is equal to the MLE, we would need to have $\alpha = \beta = 0$. However, this is not a proper distribution (the Beta requires both parameters to be positive).