Handout 18: Bayesian Statistics II

Bayesian Statistics Let's start by defining the setup from last in a formal, generic way. Bayesian statistics first considers a random variable $\vec{\theta}$, called the **prior**, with a pdf/pmf equal to $f_{\vec{\theta}}$. Then, we have another random variable $\vec{X}$, the **likelihood**, defined through the conditional pdf/pmf given by $f_{\vec{X}|\vec{\theta}}$. Finally, we can compute the **posterior** distribution of $\vec{\theta}|\vec{X}$ by using a form of Bayes rule.

Point Estimation While the entire posterior distribution is the clearest picture of our knowledge of the parameter $\vec{\theta}$, sometimes we need to convert our knowledge into a single best guess point estimator. The Bayesian point estimator is the expected value of the posterior distribution. In general, this will be somewhere between the expected value of the prior distribution and the mean of the data $\vec{X}$.

Credible Intervals We can represent a version of a confidence interval for Bayesian statistics. For a univariate θ, a **credible interval** with credibility $1 - \alpha$ is simply a set of bounds l and u such that:

$$\mathbb{P}[l \leq \theta \leq u|\vec{X}] \geq 1 - \alpha.$$

A similar interval can also be constructed for any component of a multivariate parameter $\vec{\theta}$. We can compute these values using R.

Conjugate Priors In many cases, as we will see on today's worksheet, we can choose a prior distribution that aligns with the likelihood function such that the prior and the posterior come from the same family. For example, the Beta and the Binomial, as we saw on the last worksheet. The prior for a likelihood is called the **conjugate prior**.

Random Sample The most common case, as we have seen throughout the semester, is where $\vec{X}$ is a sequence of n i.i.d. observations. An easy way to get the posterior distribution is to first consider the distribution $\vec{theta}|X_1$, then condition this on X_2, then on X_3, and so forth. When we have conjugate priors, each of these distributions will be from the same family, thus making it easy to determine the distribution of each sub-step.