1. (a) Use Matlab to draw graphs of Beta\((a, b)\) distributions for different \(a\) and \(b\), to get a sense of how the distributions depend on the parameters.

(b) Consider tossing a drawing pin. Define the result of a toss to be “heads” if the point lands downwards, and “tails” otherwise.
Write \(p\) for the probability that a toss will land point downwards. Think about \(p\), and choose \(a, b\), so that a Beta\((a, b)\) prior distribution approximates your subjective prior distribution for \(p\).

(c) Now collect data. Toss a drawing pin 100 times and keep track of the number of heads after 10, 50, and 100 tosses. You may find the result depends on the surface you use.

(d) Use Matlab to draw graphs of your prior distribution and your posterior distributions for \(p\) after 10, 50, and 100 tosses.
Compare these plots with each other, and with those of your fellow drawing-pin-throwers.

2. Example adapted from Bustamante et al. (2003), doi:10.1016/S0040-5809(02)00050-3. The site-frequency spectrum \(Y = (4, 3, 1, 1, 1)\) for mutations at \(n = 5\) segregating sites from a single genomic region is approximately Poisson with mean \(\Lambda\). These five site frequencies are the observations. They tell us something about the unknown true underlying Poisson mean \(\Lambda\). We assume each one of these site frequencies is an independent observation. The model is
\[
(Y_i|\Lambda) \sim \text{Poisson}(\Lambda) \quad i = 1, 2, ..., 5.
\]
We would like to estimate \(\Lambda\).

(a) Show that a prior \(\pi(\lambda)\) given by a Gamma density
\[
\pi(\lambda) = \lambda^{\alpha-1}e^{-\beta\lambda} \frac{\beta^\alpha}{\Gamma(\alpha)}
\]
with shape \(\alpha\) and scale \(\beta\) is conjugate to the Poisson Likelihood, that is, show that if the prior is Gamma\((a, b)\), and the likelihood is Poisson, then the posterior is Gamma\((a', b')\). Work out \(a'\) and \(b'\) as a function of \(a, b\) and the data.

(b) For this region \(\alpha = 2, \beta = 2\). Compute the posterior mean for \(\Lambda\), and give a 95% equal-tailed credible region for the mean number of mutations per site. Can you compute an HPD region?

(c) Consider the scale invariant prior \(\pi(\lambda)\). Show the posterior is proper. Does using this prior effect results?

3. Return to the drawing pin problem in Q1.

(a) Is tails more probable than heads \((p < 1/2)\)? Test this hypothesis by (a) comparing the posterior probability for \(p < 1/2\) to that for \(p > 1/2\) and (b) by computing the Bayes factor for this comparison. What are the relative merits of the two approaches to measuring the evidence for \(p < 1/2\)?

(b) Choose someone else for a prior show-down. Think of you prior as a hypothesis about \(p\). Who had the better prior? Compute a Bayes factor comparing your priors.