These aren't exhaustive!! You will be provided a copy of the distribution sheet available on the website. See the last page of this document for a copy of the conjugate priors we've derived thus far in the course, along with posterior predictive distributions.

1. There are many students on campus who bike to get to class. Suppose that the college wants to install an appropriate number of bike racks around campus, so they need to know the proportion of students θ_1 who are regular bike riders during the Fall semester. The staff believes that an average number of students who bike regularly is 1 in 6, with a mode of $\frac{1}{8}$.

   (a) Specify a prior that reflects the staff's prior ideas about θ. Check Wikipedia for the mode of interest...

   (b) Among 60 surveyed students, 17 are regular bike riders. What is the posterior distribution for θ_1? What are the posterior mean, mode, and standard deviation?

   (c) Does the posterior model more closely reflect the prior information or the data? Explain your reasoning.

   (d) Now suppose the staff wants to learn about the proportion of students θ_2 who ride bikes in the Spring semester. They have reason to believe that the prior for θ_2 should be different from the prior for θ_1, due to snow and stolen bikes! With a picture or words or mathematically, try to describe a kind of *joint* prior distribution for θ_1 and θ_2 that represents the belief that θ_1 and θ_2 are close to each other, but with high uncertainty.

2. Let $Y_i|\theta \overset{\text{iid}}{\sim} \text{Geometric}(\theta)$ for $i = 1, \dots, n$. Let θ be the unknown probability of success. The interpretation of our Geometric distribution is the number of failures before first success, so $P(Y_i = y|\theta) = (1-\theta)^y\theta$, where $y = 0, 1, \dots$. Since θ is a probability, maybe we'd like to put a Beta prior on it! Let $\theta \sim \text{Beta}(a, b)$.

   (a) Obtain the posterior for θ given data $y_1, y_2, \dots, y_n$.

   (b) Based on (a), provide some rough interpretations of the a and b values in our prior.

   (c) Find the Jeffreys' prior (either named distribution, or up to proportionality) for a Geometric sampling model.

3. The Weibull distribution with scale parameter $\theta > 0$ and shape parameter $\beta > 0$ is given by the following parameterization:

$$f(y|\theta, \beta) = \frac{\beta}{\theta} y^{\beta-1} \exp\left\{-\frac{y^\beta}{\theta}\right\} \qquad y \geq 0$$

   Suppose we observe $Y_1, \dots, Y_n|\theta, \beta \overset{\text{iid}}{\sim} \text{Weibull}(\theta, \beta)$. In this problem, assume that β is known.

   (a) Find the MLE of θ based on the n observations.

   (b) Derive the conjugate family of priors for this sampling model.

(c) Using the prior in (b), write the posterior mean as a weighted sum of the prior mean and the MLE. Provide an interpretation to the parameters in the prior.

4. Haldane (1931) suggested the following "non-informative prior" for Benoulli(θ) data: $p(\theta) \propto \frac{1}{\theta(1-\theta)}$ for $\theta \in (0,1)$, which is a Beta$(0,0)$ distribution. We write $\propto$ here instead of $=$ because this prior isn't proper (doesn't integrate to 1).

   (a) Suppose $Y_1, \ldots, Y_n \overset{\text{iid}}{\sim}$ Bernoulli(θ). Under this choice of Haldane's prior for these data, obtain the posterior of θ.

   (b) Is the posterior obtained in (a) proper (i.e. does it integrate to 1)? If it depends, state the scenarios where the posterior is proper/improper.

5. Let $Y|\theta, \sigma^2$ have a normal distribution with mean θ and **known** variance σ^2. Your goal is to study the mean without injecting too much prior information.

   (a) Derive Jeffrey's prior for θ for this Normal sampling model, and state if the prior is a proper distribution.

   (b) Derive the posterior for θ using Jeffrey's prior and state if it is proper.

6. Suppose one is interested in monitoring the popularity of a particular blog. In this setting, the event of interest is number of visits to the blog website and the time interval is a single day. The blog author is particularly interested in learning about the average number of visitors per day during weekdays (days Monday through Friday) and predicting the number of visits for a future weekday in Summer 2019.

	Fri	Sat	Sun	Mon	Tue	Wed	Thu
Week 1	95	81	85	100	111	130	113
Week 2	92	65	78	96	118	120	104
Week 3	91	91	79	106	91	114	110
Week 4	98	61	84	96	126	119	90

Count of visitors to blog during 28 days during June 2019.

   (a) What sampling model is appropriate for these data? Once you've identified the sampling model, identify the conjugate prior with a mean and variance of 80 visits per day. Under these data and your selected prior, identify the exact posterior distribution of interest for this problem. *You might want a calculator!*

   (b) Provide pseudocode (i.e. a mix of words and `R` code) that would help this person obtain a 95% credible interval for predicting the number of blog visits we would see at a future summer weekday.

   (c) Provide pseudocode to perform a posterior predictive check where our test function of interest is the ratio of sample variance to sample mean. Without actually implementing the code, do you think that our data will pass this PPC? Why or why not?

7. Suppose one observes the outcomes of four fair coin flips $W_1, \ldots, W_4$ where $W_i = 1$ is Heads, and 0 if Tails. Define $X = W_1 + W_2 + W_3$ as the number of Heads in first three flips, and $Y = W_2 + W_3 + W_4$ as number of Heads in last three flips.

   (a) Create a table that displays the joint PMF of X and Y.

   (b) Describe in words or use pseudocode to demonstrate how Gibbs sampling can be used to simulate from the joint distribution of X and Y.

8. Review Bayes estimators.

Conjugate priors

- $\theta \sim \text{Beta}(a, b)$ for $Y_i|\theta \overset{\text{iid}}{\sim} \text{Bernoulli}(\theta)$
  - Posterior: $\theta|\mathbf{y} \sim \text{Beta}(a + \sum y_i, b + n - \sum y_i)$
  - Posterior predictive:
- $\theta \sim \text{Gamma}(a, b)$ for $Y_i|\theta \sim \text{Poisson}(\theta)$
  - Posterior: $\theta|\mathbf{y} \sim \text{Gamma}(a + \sum y_i, b + n)$
  - Posterior predictive:
- $\boldsymbol{\theta} \sim \text{Dirichlet}(a_1, \ldots, a_K)$ for $Y_i|\boldsymbol{\theta} \sim \text{Categorical}(\boldsymbol{\theta})$
  - Posterior:$\boldsymbol{\theta}|\mathbf{y} \sim \text{Dirichlet}(a_1 + n_1, \ldots, a_K + n_K)$ where $n_k = \sum_{i=1}^{n} \mathbf{1}_{\{Y_i = k\}}$
  - Posterior predictive:
- $\theta \sim N(\mu_0, \sigma_0^2)$ for $Y_i|\theta, \sigma^2 \sim N(\theta, \sigma^2)$ where σ^2 known
  - Posterior: $\theta|\mathbf{y}, \sigma^2 \sim N\left((\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2})^{-1}\left(\frac{n\bar{y}}{\sigma^2} + \frac{\mu_0}{\sigma_0^2}\right), (\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2})^{-1}\right)$
  - Posterior predictive: