FS Exercises 01 (10 pts)

Conjugate priors

"In Bayesian probability theory, if the posterior distribution $p(\theta|y)$ is in the same probability distribution family as the prior probability distribution $p(\theta)$, the prior and posterior are then called conjugate distributions, and the prior is called a conjugate prior for the likelihood function $p(y|\theta)$."
(`https://en.wikipedia.org/wiki/Conjugate_prior`)

Beta prior, binomial likelihood, beta posterior

Exercise 1.1 (2 pts)

Consider an experiment, where it is the goal to learn about the binomial parameter θ. In order to quantify the available prior knowledge about θ, a beta distribution is chosen; the specific prior is of the form

$$p(\theta) = \text{beta}(2, 5) = \frac{\theta^{2-1}(1-\theta)^{5-1}}{\text{B}(2, 5)},$$

where the beta function is

$$\begin{aligned}\text{B}(2, 5) &= \int \theta^{2-1}(1-\theta)^{5-1} d\theta \\ &= \frac{\Gamma(2)\Gamma(5)}{\Gamma(2+5)} = \frac{1!4!}{6!} = 1/30.\end{aligned}$$

(Note the difference between beta distribution and beta function.) Assume the experiment had $n = 30$ trials, and $y = 20$ successes were observed, so the binomial likelihood is

$$p(y|\theta) = \binom{30}{20}\theta^{20}(1-\theta)^{30-20}.$$

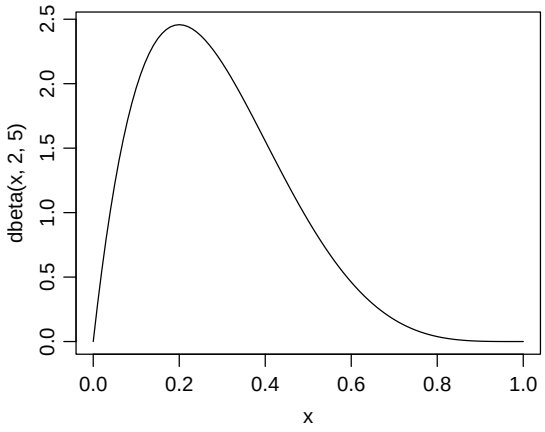
Find the posterior distribution

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{\int p(\theta)p(y|\theta)d\theta}.$$

Exercise 1.2 (4 pts)

In R, the beta prior from the previous exercise can be plotted using:

```
> par(mai = c(.6, .6, .1, .1), mgp = c(2, .7, 0))
> curve(dbeta(x, 2, 5))
```



The likelihood function is:

```
> like1 <- function(x) dbinom(20, 30, prob = x)       # unnormalized
> Z <- integrate(like1, lower = 0, upper = 1)$value   # normalization constant
> like2 <- function(x) like1(x)/Z                     # normalized
```

The posterior is:

```
> post1 <- function(x) dbeta(x, 2, 5) * like1(x)      # post = prior * like
> ZZ <- integrate(post1, lower = 0, upper = 1)$value
> post2 <- function(x) post1(x)/ZZ
```

Make a plot of the prior, the normalized likelihood, and the normalized posterior drawing them into a single display. Add a plot of the beta$(22, 15)$ distribution. What do you notice?

Exercise 1.3 (2 pts)

Show that for any y and n, the prior predictive distribution is

$$p(y) = \int p(\theta)p(y|\theta)d\theta = \frac{1}{1+n}$$

when the prior distribution is the uninformative beta$(1, 1)$. This means that, with an uninformative prior, every number of successes is a priori equally likely.

Exercise 1.4 (2 pts)

Therefore, with the knowledge from the previous exercise, the posterior distribution is

$$\begin{aligned} p(\theta|y) &= \frac{p(\theta)p(y|\theta)}{\int p(\theta)p(y|\theta)d\theta} \\ &= (n+1)\cdot 1 \cdot \binom{n}{y}\theta^y(1-\theta)^{n-y} \\ &= \frac{(n+1)n!}{(n-y)!y!}\theta^y(1-\theta)^{n-y} \\ &= \frac{(n+1)!}{(n-y)!y!}\theta^y(1-\theta)^{n-y} \end{aligned}$$

when the prior is uninformative. (This is the theorem originally given by Thomas Bayes.) Show that this is a beta$(y+1, n-y+1)$ distribution.

Exercise 1.5 (2 bonus pts)

What is the posterior predictive probability of observing one new success $\tilde{y} = 1$ given the data $y = 20$ and the prior as in Exercise 1.1? Remember that the posterior predictive distribution is $p(\tilde{y}|y) = \int p(\theta|y)p(\tilde{y}|\theta)d\theta$.

Exercise 1.6 (2 bonus pts)

What is the posterior predictive distribution of observing $\tilde{y} = 0, 1, 2, \ldots$ new successes in a new $\tilde{n}$-trial experiment given the data $y = 20$ and the prior as in Exercise 1.1?