1. The Beta$(1, 2)$ prior implies Kimya believes it's somewhat unlikely the store is open. The Beta$(0.5, 1)$ prior implies Fernando believes STRONGLY that the store is NOT open. The Beta$(3, 10)$ prior implies Ciara believes that the store is not open, but is not as certain as Fernando. The Beta$(2, 0.1)$ prior implies Taylor believes STRONGLY that the store IS open.

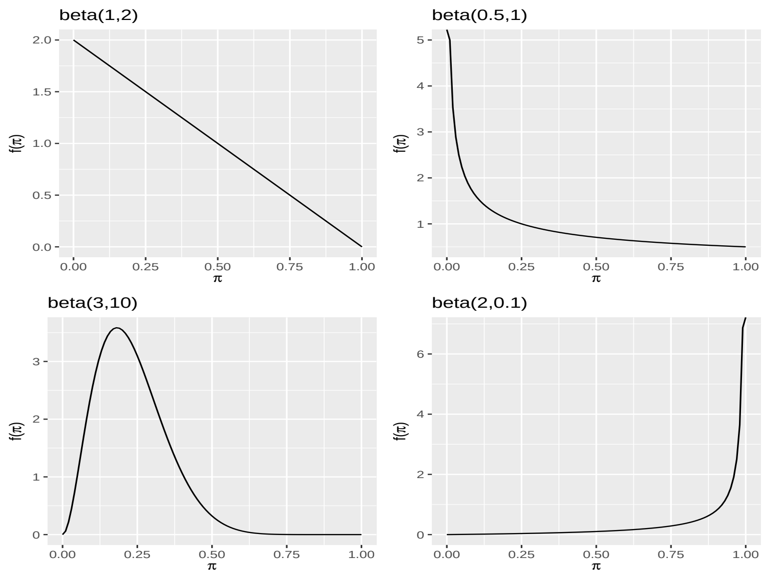

Figure 1: Problem 4.4: Beta priors.

2. (a) $y = 8, n = 10$ (b) $y = 3, n = 13$ (c) $y = 2, n = 16$ (d)$y = 7, n = 10$ (e) $y = 3, n = 6$ (f) $y = 29, n = 31$

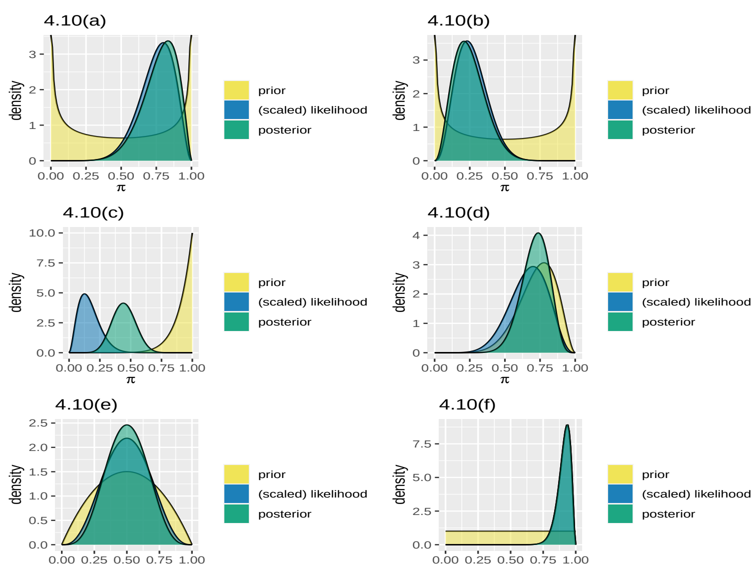

Figure 2: Problem 4.10: Beta priors, scaled likelihoods, posteriors.

3. (a) Beta$(2, 3)$ prior with $y = 3, n = 5 \Rightarrow$ Beta$(2 + 3, 3 - 5 + 3) \Rightarrow$ Beta$(5, 5)$ posterior.

(b) Beta$(5,5)$ prior with $y=1, n=5 \Rightarrow$ Beta$(6,9)$ posterior.

(c) Beta$(6,9)$ prior with $y=1, n=5 \Rightarrow$ Beta$(7,13)$ posterior.

(d) Beta$(7,13)$ prior with $y=2, n=5 \Rightarrow$ Beta$(9,16)$ posterior.

4.
```
> bechdel %>% filter(year==1980) %>% tabyl(binary)
 binary  n   percent
   FAIL 10 0.7142857
   PASS  4 0.2857143
> summarize_beta_binomial(alpha = 1, beta = 1, y = 4, n = 14)
      model alpha beta   mean      mode        var        sd
1     prior     1    1 0.5000       NaN 0.08333333 0.2886751
2 posterior     5   11 0.3125 0.2857143 0.01263787 0.1124183
> bechdel %>% filter(year==1990) %>% tabyl(binary)
 binary n percent
   FAIL 9     0.6
   PASS 6     0.4
> summarize_beta_binomial(alpha = 5, beta = 11, y = 6, n = 15)
      model alpha beta      mean      mode         var         sd
1     prior     5   11 0.3125000 0.2857143 0.012637868 0.11241827
2 posterior    11   20 0.3548387 0.3448276 0.007154006 0.08458136
>
> bechdel %>% filter(year==2000) %>% tabyl(binary)
 binary  n   percent
   FAIL 34 0.5396825
   PASS 29 0.4603175
> # Next, 2000 data:
> summarize_beta_binomial(alpha = 11, beta = 20, y = 29, n = 63)
      model alpha beta      mean      mode         var         sd
1     prior    11   20 0.3548387 0.3448276 0.007154006 0.08458136
2 posterior    40   54 0.4255319 0.4239130 0.002573205 0.05072677

> # All data at once:
> bechdel %>% filter(year==1980|year==1990|year==2000) %>% tabyl(binary)
 binary  n  percent
   FAIL 53 0.576087
   PASS 39 0.423913
>
> summarize_beta_binomial(alpha = 1, beta = 1, y = 39, n = 92)
      model alpha beta      mean     mode         var         sd
1     prior     1    1 0.5000000      NaN 0.083333333 0.28867513
2 posterior    40   54 0.4255319 0.423913 0.002573205 0.05072677
```

5. (a) The R code is given on the course web page. The posterior mean and median for σ^2 is around 228.2. The posterior mean and median for μ is around 225.2. A 95% HPD credible interval for σ^2 is around $(214.95, 241.9)$. A 95% HPD credible interval for μ is around $(217.1, 233.51)$.

(b) The R code is given on the course web page. The posterior mean and median for μ is around 223.2. A 95% HPD credible interval for μ is around $(216.8, 229.7)$.

(c) The point estimate for μ is slightly different in the two cases. But the most notable difference is that the credible interval for μ is narrower when σ^2 is assumed known, which makes sense, since there is less randomness in the model.

6. (a) We need $s/r = 5$ and $\sqrt{\frac{s}{r^2}} = 0.25$. Set $s = 5r$ and substitute into the second equation, and solving these equations yields $s = 400, r = 80$.

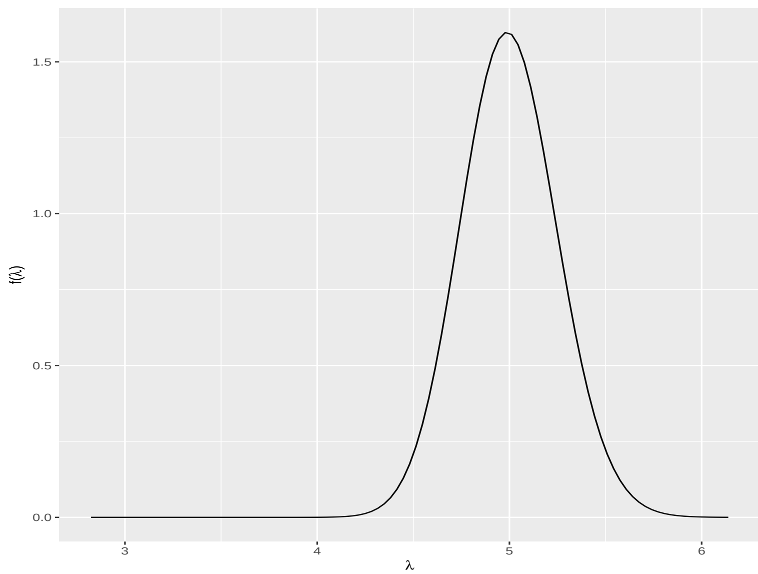

Figure 3: Problem 5.5: prior.

(b) The R code `1-pgamma(10,shape=400,rate=80)` shows that this probability is approximately 0.

7. (a), (b)

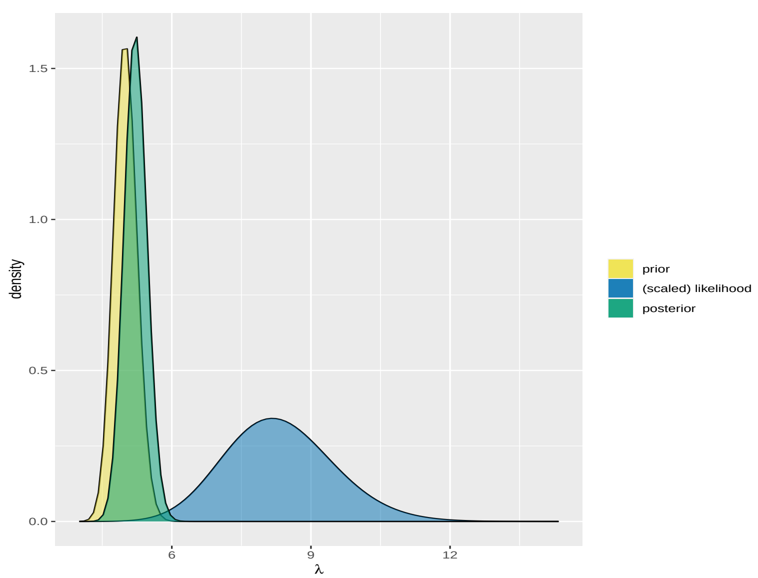

Figure 4: Problem 5.6: prior, likelihood, posterior.

(c)

```
> summarize_gamma_poisson(shape = 400, rate = 80, sum_y = 49, n = 6)
      model shape rate    mean     mode        var        sd
1     prior   400   80 5.00000 4.987500 0.06250000 0.2500000
2 posterior   449   86 5.22093 5.209302 0.06070849 0.2463909
```

(d) Seeing the data has slightly increased my estimate of the Poisson mean λ, from a prior estimate of 5 to a posterior estimate of around 5.2.

8. (a)

```
> control_subjects <- football %>%
+   filter(group == "control")
> control_subjects %>%
+   summarize(mean(volume))
  mean(volume)
1       7.6026
```

(b)

```
> summarize_normal_normal(mean = 6.5, sd = 0.4, sigma = 0.5,
+                          y_bar = 7.6026, n = 25)
      model     mean     mode         var         sd
1     prior 6.500000 6.500000 0.160000000 0.40000000
2 posterior 7.537741 7.537741 0.009411765 0.09701425
```

(c)

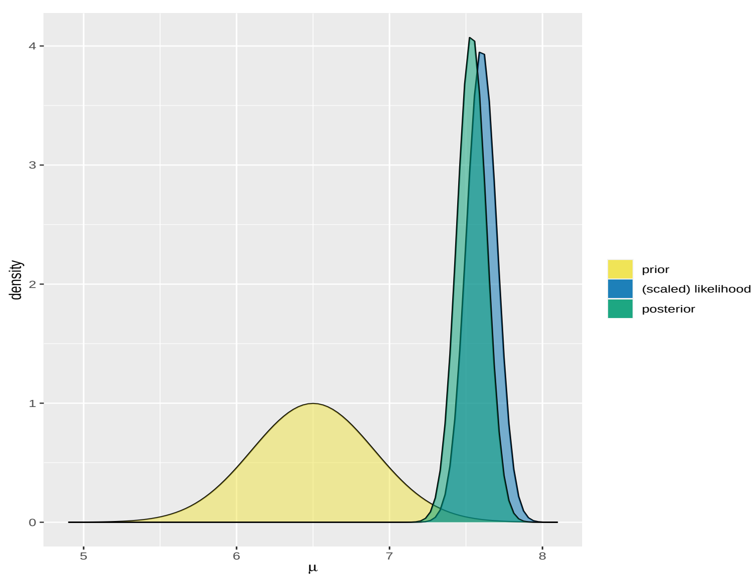

Figure 5: Problem 5.12: prior, likelihood, posterior.

Seeing the data has substantially increased my estimate of the normal mean μ, from a prior estimate of 6.5 to a posterior estimate of around 7.5.

9. (a) The posterior is proportional to

$$\begin{aligned} L(\theta|y)p(\theta) &\propto \theta(1-\theta)^{y-1}\theta^{\alpha-1}(1-\theta)^{\beta-1} \\ &= \theta^{(\alpha+1)-1}(1-\theta)^{(y+\beta-1)-1} \end{aligned}$$

This posterior is a Beta$(\alpha+1, y+\beta-1)$ distribution.

(b) Yes, since the prior and posterior are both Beta distributions, the Beta is a conjugate prior.

10. (a) The posterior is proportional to

$$\begin{aligned} L(\lambda|y_1,\ldots,|y_n)p(\lambda) &\propto [\lambda^{y_1}e^{-\lambda}]\cdots[\lambda^{y_n}e^{-\lambda}]e^{-\lambda} \\ &= [\lambda^{\sum_{i=1}^n y_i}]e^{-(n+1)\lambda} \end{aligned}$$

This posterior is a gamma$(\sum_{i=1}^n y_i+1, n+1)$ distribution.

(b) For this data set, $\sum_{i=1}^n y_i = 33$ and $n = 15$, so the posterior is gamma$(34, 16)$. R code:

```
# sum of Y's = 33, n = 15
# Posterior median:
qgamma(.5,shape=33+1,rate=15+1)
# Posterior mean:
34/16
# quantile-based interval:
qgamma(c(.025,.975),shape=33+1,rate=15+1)
# HPD interval:
hpd(qgamma,shape=33+1,rate=15+1)
```

The posterior mean is $34/16 = 2.125$. The 95% quantile-based credible interval is $(1.47, 2.90)$. The 95% HPD credible interval is $(1.44, 2.85)$.