STAT 535 – Homework 1 Solutions

1.

$$P(Q_2|G) = \frac{(0.19)(0.25)}{(0.11)(0.25) + (0.31)(0.25) + (0.53)(0.25)} = \frac{(0.19)(0.25)}{0.285} = 0.1667$$

$$P(Q_3|G) = \frac{(0.31)(0.25)}{(0.11)(0.25) + (0.31)(0.25) + (0.53)(0.25)} = \frac{(0.31)(0.25)}{0.285} = 0.272$$

$$P(Q_4|G) = \frac{(0.53)(0.25)}{(0.11)(0.25) + (0.31)(0.25) + (0.53)(0.25)} = \frac{(0.53)(0.25)}{0.285} = 0.465$$

Unconditional on graduation status, the probability of being in each of the income quartiles is 0.25. However, given that a 30+ male is a college graduate, it is quite unlikely for him to be in the lowest income quartile, and substantially likelier for him to be in the higher income quartiles (almost 50% chance of being in the top income quartile).

2. (a) $P(B|A) = 0.73$. (b) $P(A) = 0.20$. (c) $P(D) = 0.15$. (d) $P(D|C) = 0.91$. (e) $P(E \cap F) = 0.38$. (f) $P(E|F) = 0.95$.

3. Let $D$ = delayed flight and $M$ = morning flight. (a) By Bayes' Rule,

$$P(D|M) = \frac{P(M|D)P(D)}{P(M)} = \frac{(0.40)(0.15)}{0.30} = 0.20$$

(b) From part (a), $P(D^c|M) = 1 - P(D|M) = 0.80$. By Bayes' Rule,

$$P(M|D^c) = \frac{P(D^c|M)P(M)}{P(D^c)} = \frac{(0.80)(0.30)}{0.85} = 0.282$$

4. Let $G$ = good mood, $B$ = bad mood, $L$ = low number of texts, $M$ = medium number, and $H$ = high number. Note that $P(L \cap G) = P(L|G)P(G) = (0.05)(0.40) = 0.02$; $P(L \cap B) = P(L|B)P(B) = (0.13)(0.60) = 0.078$;
$P(M \cap G) = P(M|G)P(G) = (0.84)(0.40) = 0.336$; $P(M \cap B) = P(M|B)P(B) = (0.86)(0.60) = 0.516$;
$P(H \cap G) = P(H|G)P(G) = (0.11)(0.40) = 0.0924$; $P(H \cap B) = P(H|B)P(B) = (0.01)(0.60) = 0.006$.

| | good mood | bad mood | total |
|---|---|---|---|
| 0 texts | 0.02 | 0.078 | 0.098 |
| 1-45 texts | 0.336 | 0.516 | 0.852 |
| 46+ texts | 0.044 | 0.006 | 0.05 |
| Total | 0.4 | 0.6 | 1 |

(b) This is $P(G) = 0.40$. This is a prior probability.

(c) This is $P(H|G) = 0.11$. This is a likelihood value.

(d)

$$P(G|H) = \frac{P(H|G)P(G)}{P(H)} = \frac{(0.11)(0.40)}{(0.05)} = 0.88.$$

5. (a) Given $\pi$, $Y$ is binomial$(6, \pi)$, so the pmf of $Y|\pi$ is:

$$f(y|\pi) = \binom{6}{y}\pi^y(1-\pi)^{6-y}, y = 1, 2, \ldots, 6.$$

(b) $P[Y = 4|\pi = 0.3] = \binom{6}{4}0.3^4(0.7)^{6-4} = 0.06$

(c)

$$\begin{aligned}
&P[\pi = 0.3|Y = 4] \\
&= \frac{P[Y = 4|\pi = 0.3]P(\pi = 0.3)}{P[Y = 4|\pi = 0.3]P(\pi = 0.3) + P[Y = 4|\pi = 0.4]P(\pi = 0.4) + P[Y = 4|\pi = 0.5]P(\pi = 0.5)} \\
&= \frac{(0.06)(0.25)}{(0.06)(0.25) + (0.138)(0.60) + (0.234)(0.15)} = 0.113.
\end{aligned}$$

$$\begin{aligned}
&P[\pi = 0.4|Y = 4] \\
&= \frac{P[Y = 4|\pi = 0.4]P(\pi = 0.4)}{P[Y = 4|\pi = 0.3]P(\pi = 0.3) + P[Y = 4|\pi = 0.4]P(\pi = 0.4) + P[Y = 4|\pi = 0.5]P(\pi = 0.5)} \\
&= \frac{(0.138)(0.60)}{(0.06)(0.25) + (0.138)(0.60) + (0.234)(0.15)} = 0.623.
\end{aligned}$$

$$\begin{aligned}
&P[\pi = 0.5|Y = 4] \\
&= \frac{P[Y = 4|\pi = 0.5]P(\pi = 0.5)}{P[Y = 4|\pi = 0.3]P(\pi = 0.3) + P[Y = 4|\pi = 0.4]P(\pi = 0.4) + P[Y = 4|\pi = 0.5]P(\pi = 0.5)} \\
&= \frac{(0.234)(0.15)}{(0.06)(0.25) + (0.138)(0.60) + (0.234)(0.15)} = 0.264.
\end{aligned}$$

6. (a) Without doing any math: The value of $\pi = 0.6$ has the highest prior probability, and the sample proportion 47/80 seems roughly around 0.6, so I'd say that $\pi = 0.6$ has the highest posterior probability, maybe followed by $\pi = 0.7$ and $\pi = 0.5$.

(b)

$$\begin{aligned}
&P[\pi = 0.4|Y = 47] \\
&= \frac{P[Y = 47|\pi = 0.4]P(\pi = 0.4)}{P[Y = 47|\pi = 0.4]P(\pi = 0.4) + P[Y = 47|\pi = 0.5]P(\pi = 0.5) + P[Y = 47|\pi = 0.6]P(\pi = 0.6) + P[Y = 47|\pi = 0.7]P(\pi = 0.7)} \\
&= \frac{(0.0003)(0.1)}{(0.0003)(0.1) + (0.0264)(0.2) + (0.088)(0.44) + (0.0093)(0.26)} = 0.0006.
\end{aligned}$$

$$\begin{aligned}
&P[\pi = 0.5|Y = 47] \\
&= \frac{P[Y = 47|\pi = 0.5]P(\pi = 0.5)}{P[Y = 47|\pi = 0.4]P(\pi = 0.4) + P[Y = 47|\pi = 0.5]P(\pi = 0.5) + P[Y = 47|\pi = 0.6]P(\pi = 0.6) + P[Y = 47|\pi = 0.7]P(\pi = 0.7)} \\
&= \frac{(0.0264)(0.2)}{(0.0003)(0.1) + (0.0264)(0.2) + (0.088)(0.44) + (0.0093)(0.26)} = 0.1137.
\end{aligned}$$

$$\begin{aligned}
&P[\pi = 0.6|Y = 47] \\
&= \frac{P[Y = 47|\pi = 0.6]P(\pi = 0.6)}{P[Y = 47|\pi = 0.4]P(\pi = 0.4) + P[Y = 47|\pi = 0.5]P(\pi = 0.5) + P[Y = 47|\pi = 0.6]P(\pi = 0.6) + P[Y = 47|\pi = 0.7]P(\pi = 0.7)} \\
&= \frac{(0.088)(0.44)}{(0.0003)(0.1) + (0.0264)(0.2) + (0.088)(0.44) + (0.0093)(0.26)} = 0.8336.
\end{aligned}$$

$$
\begin{aligned}
&P[\pi = 0.7|Y = 47] \\
&= \frac{P[Y = 47|\pi = 0.7]P(\pi = 0.7)}{P[Y = 47|\pi = 0.4]P(\pi = 0.4) + P[Y = 47|\pi = 0.5]P(\pi = 0.5) + P[Y = 47|\pi = 0.6]P(\pi = 0.6) + P[Y = 47|\pi = 0.7]P(\pi = 0.7)} \\
&= \frac{(0.0093)(0.26)}{(0.0003)(0.1) + (0.0264)(0.2) + (0.088)(0.44) + (0.0093)(0.26)} = 0.0525.
\end{aligned}
$$

The probability values in the posterior distribution roughly match what I had suspected before doing the math.

(c)

$$
\begin{aligned}
&P[\pi = 0.4|Y = 470] \\
&= \frac{P[Y = 470|\pi = 0.4]P(\pi = 0.4)}{P[Y = 470|\pi = 0.4]P(\pi = 0.4) + P[Y = 470|\pi = 0.5]P(\pi = 0.5) + P[Y = 470|\pi = 0.6]P(\pi = 0.6) + P[Y = 470|\pi = 0.7]P(\pi = 0.7)} \\
&\approx \frac{(0)(0.1)}{(0)(0.1) + (0)(0.2) + (0.0221)(0.44) + (0)(0.26)} = 0.
\end{aligned}
$$

$$
\begin{aligned}
&P[\pi = 0.5|Y = 470] \\
&= \frac{P[Y = 470|\pi = 0.5]P(\pi = 0.5)}{P[Y = 470|\pi = 0.4]P(\pi = 0.4) + P[Y = 470|\pi = 0.5]P(\pi = 0.5) + P[Y = 470|\pi = 0.6]P(\pi = 0.6) + P[Y = 470|\pi = 0.7]P(\pi = 0.7)} \\
&\approx \frac{(0)(0.2)}{(0)(0.1) + (0)(0.2) + (0.0221)(0.44) + (0)(0.26)} = 0.
\end{aligned}
$$

$$
\begin{aligned}
&P[\pi = 0.6|Y = 470] \\
&= \frac{P[Y = 470|\pi = 0.6]P(\pi = 0.6)}{P[Y = 470|\pi = 0.4]P(\pi = 0.4) + P[Y = 470|\pi = 0.5]P(\pi = 0.5) + P[Y = 470|\pi = 0.6]P(\pi = 0.6) + P[Y = 470|\pi = 0.7]P(\pi = 0.7)} \\
&\approx \frac{(0.0221)(0.44)}{(0)(0.1) + (0)(0.2) + (0.0221)(0.44) + (0)(0.26)} = 1.
\end{aligned}
$$

$$
\begin{aligned}
&P[\pi = 0.7|Y = 470] \\
&= \frac{P[Y = 470|\pi = 0.7]P(\pi = 0.7)}{P[Y = 470|\pi = 0.4]P(\pi = 0.4) + P[Y = 470|\pi = 0.5]P(\pi = 0.5) + P[Y = 470|\pi = 0.6]P(\pi = 0.6) + P[Y = 470|\pi = 0.7]P(\pi = 0.7)} \\
&\approx \frac{(0)(0.26)}{(0)(0.1) + (0)(0.2) + (0.0221)(0.44) + (0)(0.26)} = 0.
\end{aligned}
$$

All the posterior probability is now on $\pi = 0.6$, since the information in the data has overwhelmed the prior information.

7. (a) Frank's test statistic has a Binomial($n = 10, p = 0.5$) distribution if the coin is fair, so the p-value is the probability under this distribution of seeing at least 8 heads (as extreme or more extreme of a test statistic value as we actually obtained). Using the code, this is 0.0547, so Frank will fail to conclude $\theta > 0.5$ using an $\alpha = 0.05$ significance level.

   (b) Jerry's test statistic has a negative binomial($r = 2, p = 0.5$) distribution if the coin is fair, so the p-value is the probability under this distribution of seeing at least 8 heads ("failures") before the second head. Using the code, this is 0.0195, so Jerry will conclude $\theta > 0.5$ using an $\alpha = 0.05$ significance level.

   (c) As explained above, the conclusions of Frank and Jerry will not agree: Frank will **fail to conclude** $\theta > 0.5$, while Jerry will conclude $\theta > 0.5$.

   (d) The likelihood based on Frank's experiment is binomial:

$$
L(\theta|y) = \binom{10}{8}\theta^8(1-\theta)^{10-8}
$$

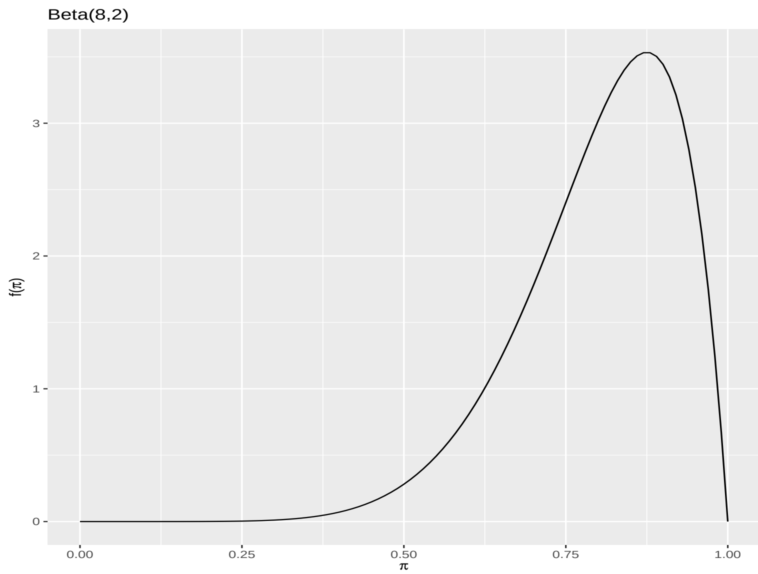


Figure 1: Beta(8,2) prior.

The likelihood based on Jerry's experiment is negative binomial:

$$L(\theta|y) = \binom{10-1}{2-1}(1-\theta)^2\theta^{10-2}$$

Using the prior $p(\theta) = 1$, in either case the posterior is

$$p(\theta|y) \propto L(\theta|y)p(\theta) \propto \theta^8(1-\theta)^2$$

Since the posteriors are Beta$(9,3)$ using either data set, Betty will make the same inference no matter which data set she uses. The Likelihood Principle implies that experiments that produce the same (or proportional) likelihoods should result in the same conclusions. The Bayesian approach assures this is true in this case, but the frequentist approach does not.

8. (a) North Dakota: Prior mean: $8/(8+2) = 0.8$. Prior mode: $(8-1)/(8+2-2) = 7/8 = 0.875$. Prior standard deviation: $[\frac{8\times 2}{(8+2)^2(8+2+1)}]^{1/2} = 0.121$.

   Louisiana: Prior mean: $1/(1+20) = 0.048$. Prior mode: $(1-1)/(1+20-2) = 0$. Prior standard deviation: $[\frac{1\times 20}{(1+20)^2(1+20+1)}]^{1/2} = 0.047$.

   (b) See plots.

   (c) The North Dakota salesperson has a fairly strong prior belief that most people say "pop". The Louisiana salesperson has a very strong prior belief that most people DO NOT say "pop".

9. (a) A beta$(6,18)$ prior distribution has mean $\frac{6}{6+18} = 1/4$ and mode $\frac{6-1}{6+18-2} = 5/22$. See plot.

   (b) From R:

```
> summarize_beta_binomial(alpha=6,beta=18,y=15,n=50)
      model alpha beta      mean      mode         var         sd
1     prior     6   18 0.2500000 0.2272727 0.007500000 0.08660254
2 posterior    21   53 0.2837838 0.2777778 0.002710007 0.05205773
```

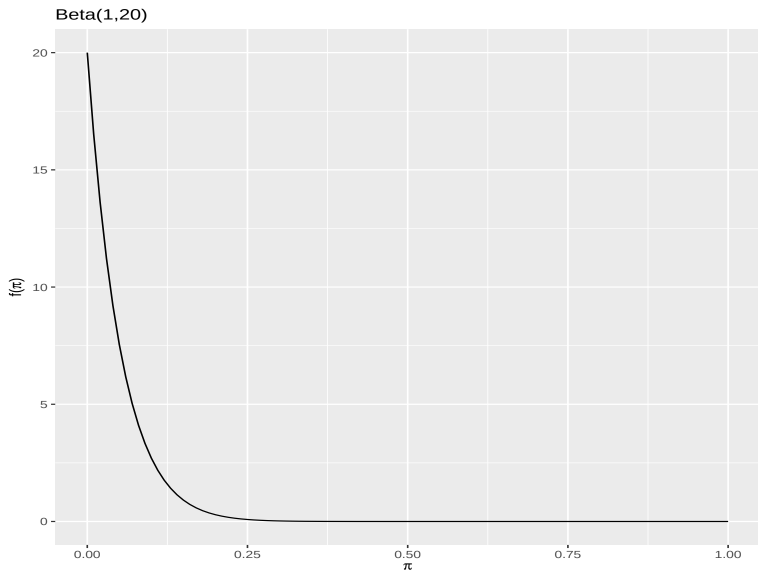


Figure 2: Beta(1,20) prior.

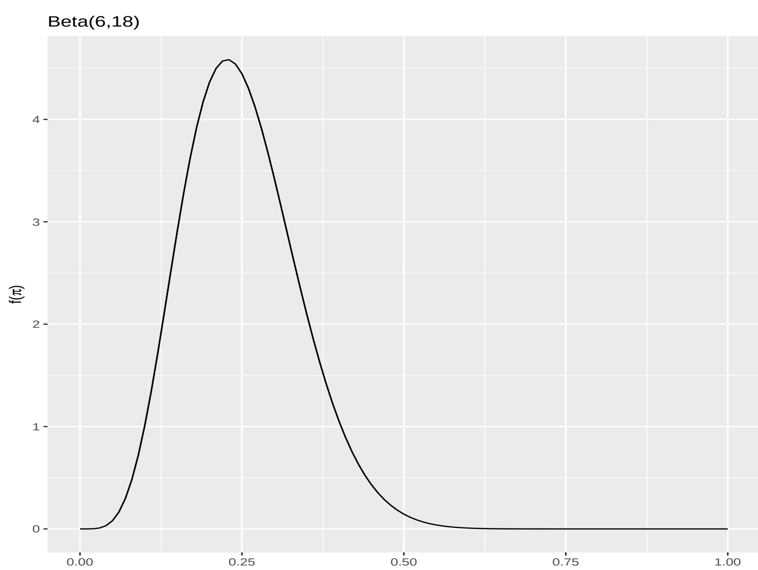


Figure 3: Beta(6,18) prior.

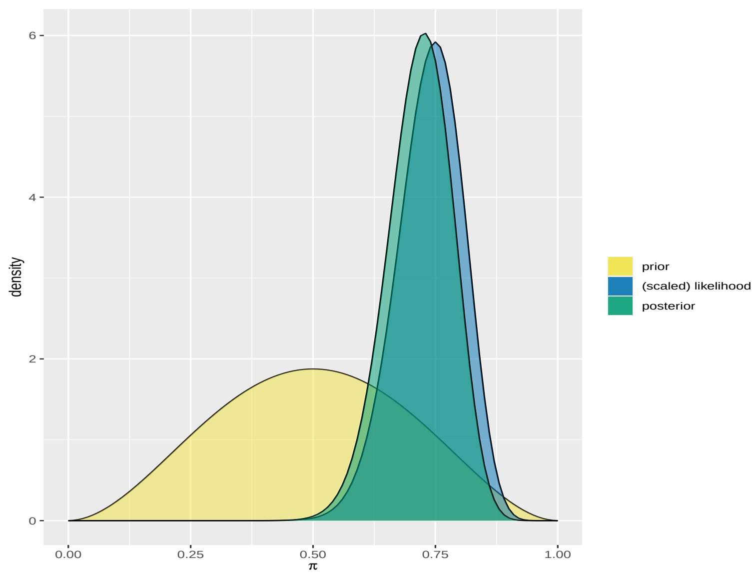


Figure 4: Patrick's prior, likelihood, and posterior.

So the posterior model is Beta$(21, 53)$.

(c) From the above output: Posterior mean $= 0.283$, mode$=0.278$, sd $= 0.052$.

(d) The sample proportion from the data is $15/50 = 0.3$, so the posterior mean is slightly closer to the sample proportion than to the prior mean. Based on this, the posterior reflects the data slightly more than it reflects the prior. Another comment is that the "weight" on the sample proportion in the calculation of the posterior mean is $\frac{n}{\alpha+\beta+n} = \frac{50}{74} = 0.676$, which implies there is more weight on the data.

10. (a) From R:

```
> plot_beta_binomial(alpha=3,beta=3,y=30,n=40)
> summarize_beta_binomial(alpha=3,beta=3,y=30,n=40)
      model alpha beta      mean      mode         var        sd
1     prior     3    3 0.5000000 0.5000000 0.035714286 0.1889822
2 posterior    33   13 0.7173913 0.7272727 0.004313639 0.0656783
```

Patrick's posterior is Beta$(33, 13)$, and his point estimate of $\pi$ (posterior mean) is 0.717.

(b) From R:

```
> plot_beta_binomial(alpha=3,beta=3,y=15,n=20)
> summarize_beta_binomial(alpha=3,beta=3,y=15,n=20)
      model alpha beta      mean      mode         var         sd
1     prior     3    3 0.5000000 0.5000000 0.035714286 0.18898224
2 posterior    18    8 0.6923077 0.7083333 0.007889546 0.08882312
```

Harold's posterior is Beta$(18, 8)$, and his point estimate of $\pi$ (posterior mean) is 0.692.

(c) Since Harold's sample size is smaller than Patrick's, his Bayesian point estimate is closer to the prior mean of 0.5 than Patrick's is.

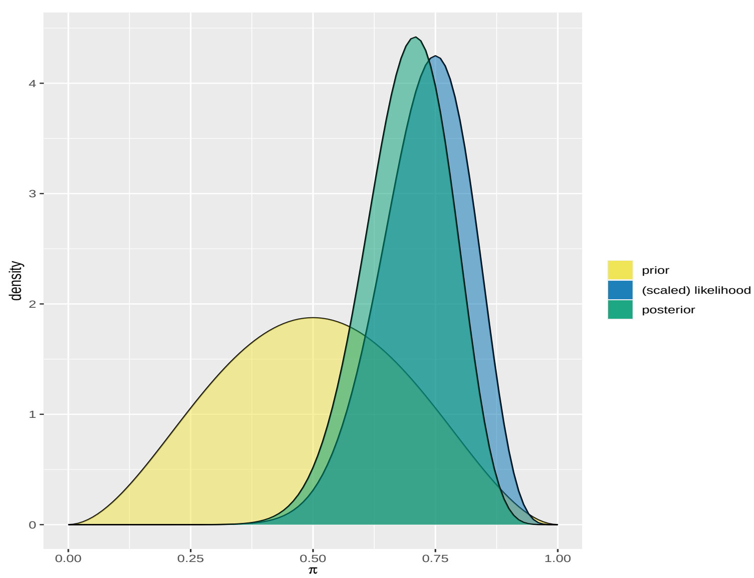


Figure 5: Harold's's prior, likelihood, and posterior.