Stat 574B: HW 3 Problem 1

Nate Hattersley

Octorber 18, 2017

part a.

- **True Posterior**

We know from the previous homework that the conjugate posterior of this binomial sample with a Beta(2,2) prior is Beta(21,8).

- **Normal Approximation**

The large-sample normal approximation to the binomial is simple. If $\hat{p}$ is the sample proportion of successes out of n draws, then we say

$$\hat{\theta}_{MLE} \sim N(\hat{p}, \frac{\hat{p}(1-\hat{p})}{n})$$

Or in our case,

$$\hat{\theta}_{MLE} \sim N(.76, .0854^2)$$

- **Laplace Approximation**

For this derivation we need the likelihood, prior, and mode of the posterior distribution:

$$L(\theta|x) = \theta^{\sum_i x_i}(1-\theta)^{n-\sum_i x_i}$$

$$p(\theta) = \frac{\theta(1-\theta)}{B(2,2)}$$

$$\theta_* = \frac{\alpha^* - 1}{\alpha^* + \beta^* - 2} = \frac{20}{27}$$

The mean of the Laplace approximation is the mode of the posterior, θ_* We find $l(\theta)$ and take the negative reciprocal of its second derivative evaluated at θ_* in order to find the variance of the approximation:

$$l(\theta) = (\sum_i x_i + 1)\log\theta + (n - \sum_i x_i + 1)\log(1-\theta) - \log(\mathrm{B}(2{,}2))$$

$$\frac{d^2 l(\theta)}{d\theta^2} = -\frac{\sum_i x_i + 1}{\theta^2} - \frac{n - \sum_i x_i + 1}{(1-\theta)^2}$$

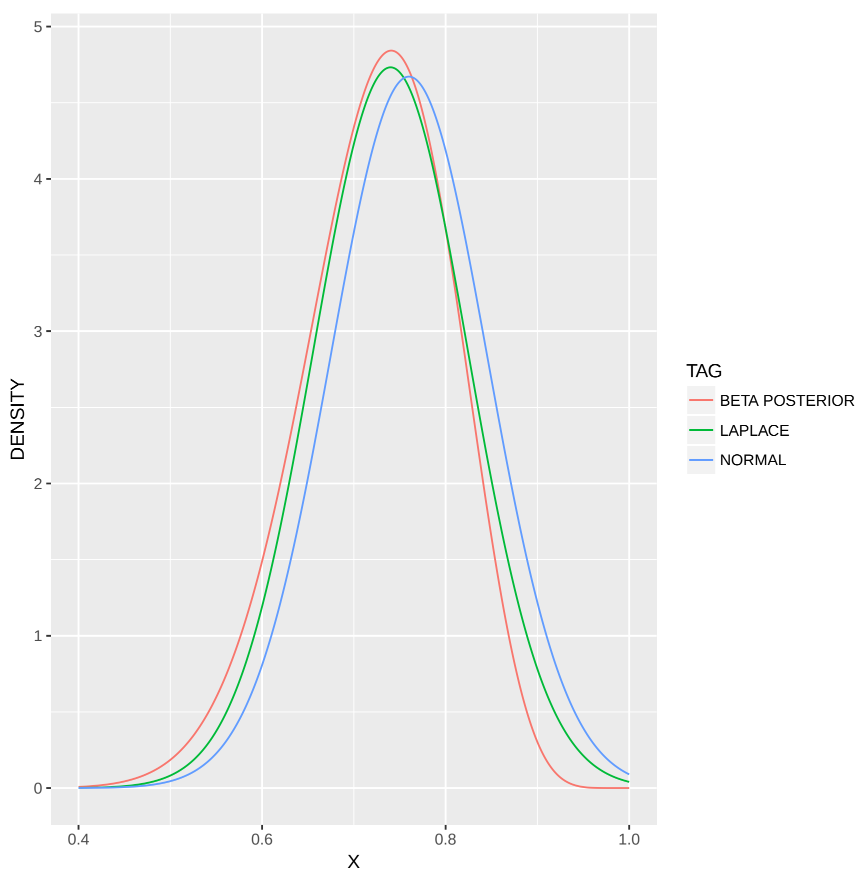
$$\frac{d^2 l(\theta)}{d\theta^2}\Big|_{\theta=\theta_*} = -140.5929$$

Hence we find that the Laplace approximation is:

$$\hat{\theta}_L \sim N(.74, .0843^2)$$

- **Comparison**

Here is a plot of all three densities, on the region of $\theta \in [0.4, 1]$ since that is where most of the density is.



Comparing the three, the Laplace approximation and the true posterior have the same mode, since Laplace is symmetric. The MLE approximation is shifted to the right of the other two since it doesn't include the 2/4 successes/tries of the prior. The true posterior density seems to increase for lower values of θ on the left side of the peak, and decrease for lower values of θ on the right side. So it seems that both approximations have a positive bias with respect to the true posterior.

part b.

Means, standard deviations, and quantiles were all straightforward to compute. Here is a summary of means and standard deviations for each distribution:

Table 1: Summaries

Distribution	Mean	SD
MLE	0.7600	0.0854
Laplace	0.7407	0.0843
Beta Posterior	0.7241	0.0816

Here are the equal-tails quantiles for each distribution given differing levels of α. The size column is just the difference between the upper- and lower-α tails of the $1-\alpha\%$ confidence interval.

Table 2: Quantiles

Distribution	alpha	lower	upper	size
MLE	0.1000	0.6506	0.8694	0.2189
Laplace	0.1000	0.6320	0.8480	0.2161
True Posterior	0.1000	0.6155	0.8259	0.2104
MLE	0.0500	0.6195	0.9005	0.2809
Laplace	0.0500	0.6013	0.8787	0.2773
True Posterior	0.0500	0.5813	0.8491	0.2679
MLE	0.0250	0.5926	0.9274	0.3348
Laplace	0.0250	0.5748	0.9052	0.3304
True Posterior	0.0250	0.5513	0.8678	0.3165
MLE	0.0100	0.5613	0.9587	0.3973
Laplace	0.0100	0.5439	0.9361	0.3922
True Posterior	0.0100	0.5163	0.8875	0.3713

I was right in my hypothesis from part a) that the MLE approximation has a higher mean than the other two. Also, the Beta distribution is skewed left, so its mean is lower than its mode. The variation is also, in order from smallest to largest, Beta, Laplace, MLE. This means that the true posterior distribution gives us the tightest confidence intervals on θ.

Code

```
x <- seq(0.4, 1, 0.001)
n <- length(x)

## Plot from part a
density.df <- data.frame(TAG = rep(c("NORMAL", "LAPLACE", "BETA POSTERIOR"),
    each = n), X = rep(x, times = 3), DENSITY = c(dnorm(x, 0.76,
    0.0854), dnorm(x, 0.74, 0.0843), dbeta(x, 21, 8)))

ggplot(density.df, aes(x = X, y = DENSITY, col = TAG)) + geom_line()


## Alpha-quantiles for each distribution
al <- c(0.1, 0.05, 0.025, 0.01)

mle.quantiles <- data.frame(alpha = al, lower = (lower = qnorm(al,
    0.76, 0.0854)), upper = (upper = qnorm(1 - al, 0.76, 0.0854)),
    size = upper - lower)

laplace.quantiles <- data.frame(alpha = al, lower = (lower = qnorm(al,
    0.74, 0.0843)), upper = (upper = qnorm(1 - al, 0.74, 0.0843)),
    size = upper - lower)

beta.quantiles <- data.frame(alpha = al, lower = (lower = qbeta(al,
    21, 8)), upper = (upper = qbeta(1 - al, 21, 8)), size = upper -
    lower)

all.quantiles <- cbind(Distribution = rep(c("MLE", "Laplace",
    "True Posterior"), each = length(al)), rbind(mle.quantiles,
    laplace.quantiles, beta.quantiles))

all.quantiles <- xtable(all.quantiles, caption = "Quantiles",
    digits = 4)

## Summaries of each distribibuion
means <- xtable(data.frame(Distribution = c("MLE", "Laplace",
    "Beta Posterior"), Mean = c(0.76, 20/27, 21/29), SD = c(0.0854,
    0.0843, 0.0816)), caption = "Summaries", digits = 4)

## Use with Rnw and chunk option results='asis' Outputs
## Summary Table in LATEX
print(means, caption.placement = "top", include.rownames = F)
## Use with Rnw and chunk option results='asis' Outputs
## Quantiles Table in LATEX
print(all.quantiles[order(all.quantiles$alpha, decreasing = T),
    ], caption.placement = "top", include.rownames = F, hline.after = c(-1,
    0, 3, 6, 9, nrow(x)))
```