Name: Kamaur-deen Lawal.

Date: 3/11/21

1. ~~From~~ A frequentist approach would have us assume a prior distribution on θ to be the null hypothesis for the distribution of X, and we have a collection of Samples X. We would also assume θ has a single value that can be estimated. From there we could use the resultant mean and standard deviation to construct a confidence interval for θ.

~~θ~~ However, from a Bayesian perspective, we do not assume θ has a single true value. Instead we assume, θ is a RV sampled from a distribution that we can estimate from observed data. Our prior probability is our initial guess for the distribution of θ. The probability $p(X|\theta)$ then becomes the probability of observing samples X given that our prior distribution of θ is true. Normalizing the prob. of observing the data allows us to compute the posterior probability, $p(\theta|x)$. From here we can construct a α% confidence interval to determine the chance that a sample from the distribution of θ will fall within our confidence interval.

In conclusion, Bayesian methods regard probabilities as forms of uncertainties in expectation of an event. The frequentist approach OTOH regards probabilities as the relative frequencies that an event will take place over repeated trials.

Chapter 3 Problem 2

We start by assuming independent uniform prior distributions on the multi-nomial parameters. Then the posterior distributions are independent multi-nomial

$$(\pi_1, \pi_2, \pi_3) \mid y \sim \text{Dirichlet}(295, 308, 39)$$

$$(\pi_1^*, \pi_2^*, \pi_3^*) \mid y \sim \text{Dirichlet}(289, 333, 20)$$

and $\alpha_1 = \dfrac{\pi_1}{\pi_1 + \pi_2}$, $\alpha_2 = \dfrac{\pi_1^*}{\pi_1^* + \pi_2^*}$. Therefore $\alpha_1 \mid y \sim \text{Beta}(295, 308)$

$\alpha_2 \mid y \sim \text{Beta}(289, 333)$.

Based on the resultant histogram, the posterior probability ~~that~~ that there is a shift towards Bush is $\boxed{19\%}$.

Code:

```
alpha.1 <- rbeta(2000, 295, 308)
alpha.2 <- rbeta(2000, 289, 333)
difference <- alpha.2 - alpha.1
hist(difference)
print(mean(dif>0))
```

Chapter 3 Problem 3.

(a) Data distribution is $p(y|\mu_c, \mu_t, \sigma_c, \sigma_t) = \prod_{i=1}^{32} N(y_{ci}|\mu_c, \sigma_c^2) \prod_{i=1}^{36} N(y_{ti}|\mu_t, \sigma_t^2)$.

Posterior distribution is $p(\mu_c, \mu_t, \log\sigma_c, \log\sigma_t|y) = p(\mu_c, \mu_t, \log\sigma_c, \log\sigma_t) \cdot$

$$p(y|\mu_c, \mu_t, \log\sigma_c, \log\sigma_t)$$

$$= \prod_{i=1}^{32} N(y_{ci}|\mu_c, \sigma_c^2) \prod_{i=1}^{36} N(y_{ti}|\mu_t, \sigma_t^2).$$

The posterior density ends up factoring, so (μ_c, σ_c) and (μ_t, σ_t) are independent in the posterior distribution. The marginal posterior distributions for μ_c and μ_t are $\mu_c|y \sim t_{31}(1.013, .24^2/32)$ and $\mu_t|y \sim t_{35}(1.173, .20^2/36)$.

(b) The resultant histogram of 1000 draws from the posterior density tells us that a 95% posterior interval for the average treatment effect is $[.05, .27]$.

Code:
```
muc <- 1.013 + (.24/sqrt(32))* rt(1000,31).
mut <- 1.173 + (.20/sqrt(36)) * rt(1000, 35)
diff <- mut-muc
hist (diff)
print (sort(diff)[c(25, 976)])
```

Chapter 3 Problem 5

(a) If the measurements are exact, then we get the posterior distributions of the mean and variance as $p(\mu, \sigma^2, y) \sim N(\bar{y}, \sigma^2/n)$ and $p(\sigma^2|y) \sim \text{Inv-}\chi^2(n-1, s^2)$, respectively. The given measurements result in $\bar{y} = 10.4$, $n=5$, and $s^2 = \frac{1}{4}\sum_{i=1}^{5}(y_i - 10.4)^2 = 1.3$.

Therefore $p(\sigma^2|y) \sim \text{Inv-}\chi^2(4, 1.3)$ and $p(\mu, \sigma^2, y) \sim N(10.4, \sigma^2/5)$.

(b). The prior is given by $p(\mu, \sigma^2) \sim \sigma^{-2}$. For some specific value, y_i, the probability of pulling a value $\pm .5$ from a normal distribution with mean μ and standard deviation σ is given by $\Phi\left(\frac{y_i + .5 - \mu}{\sigma}\right) - \Phi\left(\frac{y_i - .5 - \mu}{\sigma}\right)$

Therefore the posterior distribution is $p(\mu, \sigma^2|y) \propto p(\mu, \sigma^2) \cdot p(y|\mu, \sigma^2)$.

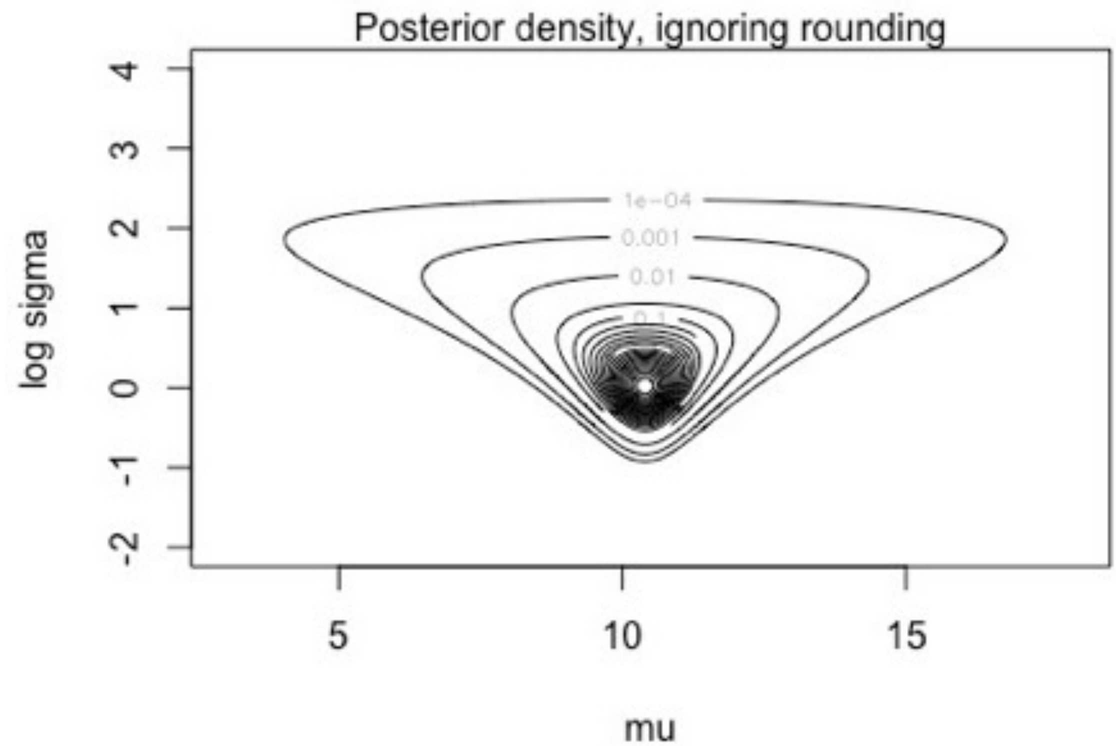
$$= \boxed{\sigma^{-2} \prod_{i=1}^{n} \left(\Phi\left(\frac{y_i + .5 - \mu}{\sigma}\right) - \Phi\left(\frac{y_i - .5 - \mu}{\sigma}\right)\right)}$$

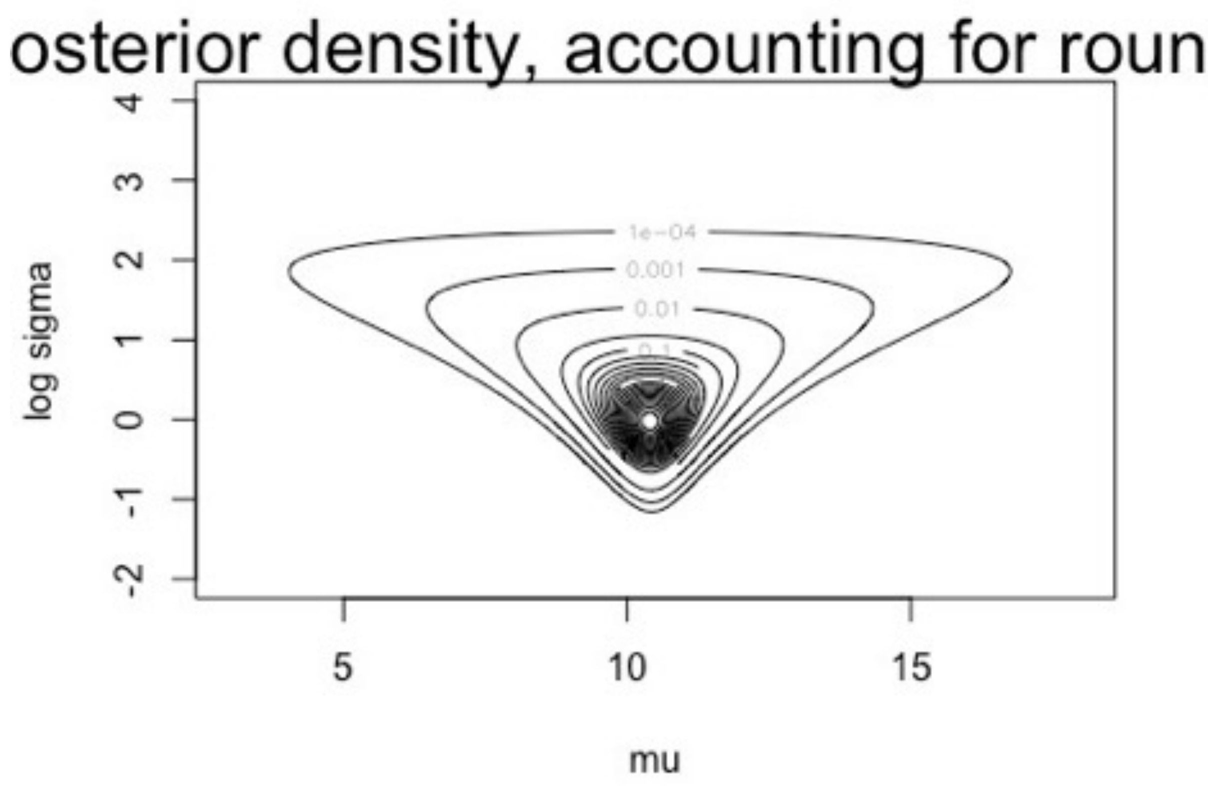
(c) See Resultant Contour plots on the next page

(d) Used R to generate the posterior mean value: $(z_1 - z_2)^2 = .16$.

The code used to solve ~~problems~~ parts c and d is on the next page.

```
# Code for 3.5c and 3.5d
posta <- function(mu,sd,y){
  ldens <- 0
  for (i in 1:length(y)) ldens <- ldens +
    log(dnorm(y[i],mu,sd))
  ldens}

postb <- function(mu,sd,y){
  ldens <- 0
  for (i in 1:length(y)) ldens <- ldens +
    log(pnorm(y[i]+0.5,mu,sd) - pnorm(y[i]-0.5,mu,sd))
  ldens}

summ <- function(x){c(mean(x),sqrt(var(x)), quantile(x, c(.025,.25,.5,.75,.975)))}

nsim <- 2000
y <- c(10,10,12,11,9)
n <- length(y)
ymean <- mean(y)
s2 <- sum((y-mean(y))^2)/(n-1)

mumatrix <- seq(3,18,length=200)
logsdmatrix <- seq(-2,4,length=200)

contours <- c(.0001,.001,.01,seq(.05,.95,.05))

logdens <- outer (mumatrix, exp(logsdmatrix), posta, y)
dens <- exp(logdens - max(logdens))
contour (mugrid, logsdmatrix, dens, levels=contours, xlab="mu", ylab="log sigma", labex=0, cex=2)

mtext ("Posterior density, ignoring rounding", 3)

sd <- sqrt((n-1)*s2/rchisq(nsim,4))
mu <- rnorm(nsim,ymean,sd/sqrt(n))
print (rbind (summ(mu),summ(sd)))

logdens <- outer (mumatrix, exp(logsdmatrix), postb, y)
dens <- exp(logdens - max(logdens))

contour (mumatrix, logsdmatrix, dens, levels=contours, xlab="mu", ylab="log sigma", labex=0, cex=2)
mtext ("Posterior density, accounting for rounding", cex=2, 3)

dens.mu <- apply(dens,1,sum)
muindex <- sample (1:length(mumatrix), nsim, replace=T, prob=dens.mu)

mu <- mumatrix[muindex]
sd <- rep (NA,nsim)

for (i in (1:nsim)) sd[i] <- exp (sample
   (logsdmatrix, 1, prob=dens[muindex[i],]))

print (rbind (summ(mu),summ(sd)))
```


Posterior density, ignoring rounding
log sigma
4
3
2
1
0
-1
-2
1e-04
0.001
0.01
0.1
5
10
15
mu


osterior density, accounting for roun
log sigma
4
3
2
1
0
-1
-2
1e-04
0.001
0.01
0.1
5
10
15
mu

Chapter 3 Problem 7 (pt. 1)

Let Poisson(θ_b) and Poisson(θ_b) denote two models. In addition we have Binomial(p, $b+v$). We will show that the models end up producing the same likelihood given the transformation $p = \dfrac{\theta_b}{\theta_b+\theta_v}$. We start by showing that $b+v$ is a RV w/ a Poisson distribution. Suppose $b+v=n$. Then

$$P(b+v=n) = \sum_{i=0}^{n} P(b+v=n,\ u=i) \quad \text{(law of total probability)}$$

$$= \sum_{i=0}^{n} P(v=n-b,\ u=i)$$

$$= \sum_{i=0}^{n} P(v=n-b)\,P(u=i) \quad \text{(Independence)}$$

$$= \sum_{i=0}^{n} e^{-\theta_v}\frac{\theta_v^{\,n-i}}{(n-i)!}\, e^{-\theta_b}\frac{\theta_b^{\,i}}{i!}$$

$$= e^{-\theta_v-\theta_b}\cdot\frac{1}{n!}\sum_{i=0}^{n}\frac{n!}{(n-i)!}\,\theta_v^{\,n-i}\,\theta_b^{\,i}$$

$$= \frac{e^{-\theta_v-\theta_b}}{n!}\sum_{i=0}^{n}\binom{n}{i}\theta_v^{\,n-i}\,\theta_b^{\,i}$$

$$= \frac{e^{-\theta_v-\theta_b}}{n!}\,(\theta_v+\theta_b)^n$$

$$= \frac{(\theta_v+\theta_b)^n}{n!}\,e^{-(\theta_v+\theta_b)} \sim \text{Poisson}(\theta_b+\theta_v)$$

Chapter 3 Problem 7 (pt. 2)

Therefore $n = b + v \sim \text{Poisson}(\theta_v + \theta_b)$. Now we examine the likelihood of b under the binomial model. As previously stated. $p = \dfrac{\theta_b}{\theta_b + \theta_v}$

$$p(b) = \binom{b+v}{b} p^b (1-p)^v$$

$$= \binom{b+v}{b} \left(\frac{\theta_b}{\theta_b+\theta_v}\right)^b \left(\frac{\theta_b}{\theta_b+\theta_v}\right)^v$$

$$= \frac{(b+v)!}{b!\,v!} \left(\frac{\theta_b^b \theta_v^v}{(\theta_b+\theta_v)^{b+v}}\right)$$

$$= \frac{(b!)^{-1}(v!)^{-1}}{((b+v)!)^{-1}} \left(\frac{\theta_b^b \theta_v^v}{(\theta_b+\theta_v)^{b+v}}\right) \frac{e^{-\theta_v-\theta_b}}{e^{-\theta_v-\theta_b}}$$

$$= \frac{\dfrac{\theta_b^b}{b!} e^{-\theta_b} \, \dfrac{\theta_v^v}{v!} e^{-\theta_v}}{\dfrac{(\theta_v+\theta_b)^{v+b} e^{-\theta_v-\theta_b}}{(v+b)!}} = \frac{\text{Poisson}(\theta_b)\,\text{Poisson}(\theta_v)}{\text{Poisson}(\theta_b+\theta_v)}$$

$$= \frac{p(b)p(v)}{p(b+v)} = \frac{p(b, b+v)}{p(b+v)}$$

$$= p(b \mid b+v)$$

$p(b|b+v)$ is the probability of seeing b successes across a fixed # of trials, n, which we know is a Poisson ~~Process~~ likelihood equivalent to $\text{Poisson}(\theta_b)$. Therefore the binomial model reduces to the likelihood of the Poisson model under the substitution given

Chapter 3 Problem 10.

$$p(\sigma_j^2|y) \propto (\sigma_j^2)^{-n/2-1/2} \cdot \exp\left(-(n-1)s^2/2\sigma_j^2\right)$$

$$\Rightarrow p(1/\sigma_j^2|y) \propto (\sigma_j^2)^2 (1/\sigma_j^2)^{n/2+1/2} \exp\left(-(n-1)s^2/2\sigma_j^2\right) = (1/\sigma_j^2)^{n/2-3/2} \cdot \cancel{\exp}$$

$$\exp\left(-(n-1)s^2/2\sigma_j^2\right)$$

$\Rightarrow (n-1)s^2/\sigma_j^2$ has a χ^2_{n-1} distribution.

Independence assumptions imply that $(n_1-1)s_1^2/\sigma_1^2$ and $(n_2-1)s_2^2/\sigma_2^2$ are independent χ^2 w/ n_1-1 and n_2-1 degrees of freedom. We know that the quotient of two independent χ^2 RVs, each divided by their d.o.f's, has the F distribution. In particular, $\dfrac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}$ has the $F_{n_1-1,\, n_2-1}$ distribution.

Chapter 3 Problem 12

(a) We have enough data where we can simply use a prior distribution that is independent and locally uniform in both input parameters, which means $p(\alpha, \beta) \propto 1$.

(b) We can use standard ordinary least squares regression to generate maximum likelihood estimates for α and β. This leads to an α value of ≈ 29 with a standard error of 2.5, and a β value of -1.1 with a standard error of .6. This gives us the following distributions $\alpha \sim N(29, 2.5)$ and $\beta \sim N(-1.1, .60)$.

(c) $p(y|\theta) \propto \text{Poisson}(\theta)$, $\theta = \alpha + \beta t$. We are using the non-informative prior. $p(\alpha, \beta) \propto 1$. From here we can easily compute the posterior distribution.

$$p(\theta|y) = p(\alpha, \beta | y, t) \propto p(y | \alpha, \beta, t) \cdot p(\alpha, \beta | t)$$

$$\propto (\alpha + \beta t)^y e^{(-\alpha - \beta t)}$$

After adding samples y_i for $i \in [n]$, $p(\alpha, \beta | y_i) \propto (\alpha + \beta t)^{\sum_{i=1}^{n} y_i} \left[e^{-n(\alpha + \beta t)} \right]$

$$\propto \text{Gamma}\left(1 + \sum_{i=1}^{n} y_i, n\right)$$