HW#5

2.19 (Carlin and Louis book)

Consider two independent binomial distributions, $X_W \sim Bin(n_W, \theta_W)$, the number of whites selected, and , $X_B \sim Bin(n_B, \theta_B)$, the number of blacks selected.

Suppose we put independent uninformative priors on the θ parameters, the joint prior is:

$\pi(\theta_W)\pi(\theta_B) = 1$.

Then the joint posterior is:

$$p(\theta_W, \theta_B \mid X_W, X_B) = \theta_{w}^{X_W}(1-\theta_W)^{n_W - X_W}\,\theta_{B}^{X_B}(1-\theta_B)^{n_B - X_B} = \theta_{w}^{41}(1-\theta_W)^{39}\theta_{B}^{14}(1-\theta_B)^{30},$$

= beta(42,40).beta(15,31)

Let $\Delta = \theta_W - \theta_B$. We can simulate the posterior probability that $\Delta > 0$ by Monte Carlo simulation as follows. Take a random sample of size G from beta(42,40). Take separately, another random sample of size G from beta(15,31). Calculate $\Delta_j = \theta_{Wj} - \theta_{Bj}$, $j = 1, \ldots, G$. Then

$$\widehat{p}(\Delta > 0 \mid data) = \frac{1}{G}\sum_{j=1}^{G} I(\Delta_j > 0)$$

In R we use the following commands:

```
> G=10000

>delta=rbeta(G,42,40)-rbeta(G,15,31)

>p_hat=sum(delta>0)/G

>p_hat

[1] 0.9812
```

This gives us .98 as an estimate for $P(\Delta > 0 \mid data)$.

Next we calculate a Bayes Factor for $H_0 : \theta_W = \theta_B$ versus $H_1 : \theta_W \neq \theta_B$. Again assuming uniform priors, we get

$$BF = \frac{M_0(X_W, X_B)}{M_1(X_W, X_B)}$$

$$= \frac{\int \binom{n_W}{X_W}\binom{n_B}{X_B}\theta^{X_W+X_B}(1-\theta)^{(n_W-X_W)+(n_B-X_W)}d\theta}{\int\int \binom{n_W}{X_W}\theta_W^{X_W}(1-\theta_W)^{(n_W-X_W)}\binom{n_B}{X_B}\theta_B^{X_B}(1-\theta_B)^{(n_B-X_W)}d\theta_W d\theta_B}$$

$$= \frac{\int_0^1 \theta^{55}(1-\theta)^{69}d\theta}{\int_0^1\int_0^1 \theta_w^{41}(1-\theta_W)^{39}\theta_B^{14}(1-\theta_B)^{30}d\theta_W d\theta_B}$$

$$= \frac{\frac{\Gamma(56)\Gamma(70)}{\Gamma(126)}}{\frac{\Gamma(15)\Gamma(31)}{\Gamma(46)}\frac{\Gamma(42)\Gamma(40)}{\Gamma(82)}}$$

```
>
BF=(gamma(56)*gamma(70)/gamma(126))/((gamma(15)*gamma(31)/gamma(46))*(gamma(42)*gamma
(40)/gamma(82)))

> BF

[1] 0.5071705
```

Thus using R we get a Bayes Factor of .507. This means that the evidence in favor of the model that allows for discrimination is only 2:1, this is not very impressive.

Alternatively, we can consider a single binomial distribution, that is consider

$X_W + X_B \sim Bin(n_W + n_B, \theta)$. Then the Bayes Factor is:

$$BF = \frac{M_0(X_W, X_B)}{M_1(X_W, X_B)}$$

$$= \frac{\int \binom{n_W+n_B}{X_W+X_B}\theta^{X_W+X_B}(1-\theta)^{(n_W-X_W)+(n_B-X_W)}d\theta}{\int\int \binom{n_W}{X_W}\theta_W^{X_W}(1-\theta_W)^{(n_W-X_W)}\binom{n_B}{X_B}\theta_B^{X_B}(1-\theta_B)^{(n_B-X_W)}d\theta_W d\theta_B}$$

$$= \frac{\binom{124}{55}}{\binom{44}{14}\binom{80}{41}}$$ [old Bayes Factor]

```
>BF_one=(choose(124,55)/(choose(44,14)*choose(80,41)))*BF

>BF_one

[1] 29.16
```

Thus using R we get a Bayes Factor of 29.16, which is more impressive than previous answer but in opposite direction. Thus there is strong evidence in support of the two samples being drawn from the same binomial distribution versus two independent binomials.

Albert book

4.4.

a)

```
y=c(10,11,12,11,9)

S=sum((y-mean(y))^2)

n=5

sigma2=S/rchisq(1000,n-1)

mu=rnorm(1000,mean=mean(y),sd=sqrt(sigma2)/sqrt(n))

>mean(mu)

[1] 10.59050

>var(mu)

[1] 0.5241484
```

b) An unrounded observation is $y_i + u_i$, where u_i is uniform (-.5,.5).

The standard noninformative prior is $g(\mu,\sigma^2) \propto \frac{1}{\sigma^2}$. Then the posterior density of the mean and variance is given by $g(\mu,\sigma^2 \mid y) \propto \frac{1}{(\sigma^2)^{n/2+1}} \exp\left(-\frac{1}{2\sigma^2}(S + n(\mu - \bar{y})^2)\right)$

where $S = \sum (y - \bar{y})^2$.

```
u=runif(5,-.5,.5)

y=y+u

S=sum((y-mean(y))^2)

n=5

sigma2=S/rchisq(1000,n-1)

mu=rnorm(1000,mean=mean(y),sd=sqrt(sigma2)/sqrt(n))

>  mean(mu)

[1] 10.58802

>var(mu)
```

[1] 0.3690483 *************The posterior of mu from unrounded values has smaller variance than that from rounded values.*********************

4.6.

(a) The joint posterior of (λ_1, λ_2) is given by:

$$g(\lambda_1, \lambda_2 \mid y_1, y_2) \propto (e^{-4\lambda_1}\lambda_1^{y_1} e^{-4\lambda_2}\lambda_2^{y_2})(\lambda_1^{143} e^{-2.4\lambda_1}\lambda_2^{99} e^{-2.5\lambda_2}) = (e^{-6.4\lambda_1}\lambda_1^{y_1+143})(e^{-6.5\lambda_2}\lambda_2^{y_2+99})$$

Thus (λ_1, λ_2) have independent gamma posterior distributions.

(b)

```
y1=260;y2=165

lambda1=rgamma(1000,shape=y1+144,scale=1/(4+1/2.4))

 lambda2=rgamma(1000,shape=y2+100,scale=1/(4+(1/2.5))

#lambda1=rgamma(1000,shape=y1+144,rate=6.4)

 #lambda2=rgamma(1000,shape=y2+100,rate=6.5)
```

(c) measure=lambda1-1.5*lambda2

```
length(measure[measure>0])/1000

[1] 0.566 (for rate)   # [1] .651 (for scale)
```