Chapter 16: Introduction to Bayesian Methods of Inference

16.1 Refer to Table 16.1.

a. $\beta(10,30)$

b. $n = 25$

c. $\beta(10,30)$, $n = 25$

d. Yes

e. Posterior for the $\beta(1,3)$ prior.

16.2 **a.-d.** Refer to Section 16.2

16.3 **a.-e.** Applet exercise, so answers vary.

16.4 **a.-d.** Applex exercise, so answers vary.

16.5 It should take more trials with a beta(10, 30) prior.

16.6 Here, $L(y \mid p) = p(y \mid p) = \binom{n}{y} p^y (1-p)^{n-y}$, where $y = 0, 1, \ldots, n$ and $0 < p < 1$. So,

$$f(y, p) = \binom{n}{y} p^y (1-p)^{n-y} \times \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

so that

$$m(y) = \int_0^1 \binom{n}{y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{y+\alpha-1}(1-p)^{n-y+\beta-1} dp = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(y+\alpha)\Gamma(n-y+\beta)}{\Gamma(n+\alpha+\beta)}.$$

The posterior density of p is then

$$g^*(p \mid y) = \frac{\Gamma(n+\alpha+\beta)}{\Gamma(y+\alpha)\Gamma(n-y+\beta)} p^{y+\alpha-1}(1-p)^{n-y+\beta-1}, \; 0 < p < 1.$$

This is the identical beta density as in Example 16.1 (recall that the sum of n i.i.d. Bernoulli random variables is binomial with n trials and success probability p).

16.7 **a.** The Bayes estimator is the mean of the posterior distribution, so with a beta posterior with $\alpha = y + 1$ and $\beta = n - y + 3$ in the prior, the posterior mean is

$$\hat{p}_B = \frac{Y+1}{n+4} = \frac{Y}{n+4} + \frac{1}{n+4}.$$

b. $E(\hat{p}_B) = \dfrac{E(Y)+1}{n+4} = \dfrac{np+1}{n+4} \neq p$, $V(\hat{p}) = \dfrac{V(Y)}{(n+4)^2} = \dfrac{np(1-p)}{(n+4)^2}$

16.8 **a.** From Ex. 16.6, the Bayes estimator for p is $\hat{p}_B = E(p \mid Y) = \dfrac{Y+1}{n+2}$.

b. This is the uniform distribution in the interval (0, 1).

c. We know that $\hat{p} = Y/n$ is an unbiased estimator for p. However, for the Bayes estimator,

$$E(\hat{p}_B) = \frac{E(Y)+1}{n+2} = \frac{np+1}{n+2} \text{ and } V(\hat{p}_B) = \frac{V(Y)}{(n+2)^2} = \frac{np(1-p)}{(n+2)^2}.$$

Thus, $MSE(\hat{p}_B) = V(\hat{p}_B) + [B(\hat{p}_B)]^2 = \frac{np(1-p)}{(n+2)^2} + \left(\frac{np+1}{n+2} - p\right)^2 = \frac{np(1-p)+(1-2p)^2}{(n+2)^2}$.

d. For the unbiased estimator $\hat{p}$, MSE($\hat{p}$) = V($\hat{p}$) = $p(1 - p)/n$. So, holding n fixed, we must determine the values of p such that

$$\frac{np(1-p)+(1-2p)^2}{(n+2)^2} < \frac{p(1-p)}{n}.$$

The range of values of p where this is satisfied is solved in Ex. 8.17(c).

16.9 **a.** Here, $L(y \mid p) = p(y \mid p) = (1-p)^{y-1}p$, where $y = 1, 2, \ldots$ and $0 < p < 1$. So,

$$f(y, p) = (1-p)^{y-1}p \times \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

so that

$$m(y) = \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha}(1-p)^{\beta+y-2}dp = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+1)\Gamma(y+\beta-1)}{\Gamma(y+\alpha+\beta)}.$$

The posterior density of p is then

$$g^*(p \mid y) = \frac{\Gamma(\alpha+\beta+y)}{\Gamma(\alpha+1)\Gamma(\beta+y-1)} p^{\alpha}(1-p)^{\beta+y-2},\ 0 < p < 1.$$

This is a beta density with shape parameters $\alpha^* = \alpha + 1$ and $\beta^* = \beta + y - 1$.

b. The Bayes estimators are

(1) $\hat{p}_B = E(p \mid Y) = \dfrac{\alpha+1}{\alpha+\beta+Y}$,

(2) $[p(1-p)]_B = E(p \mid Y) - E(p^2 \mid Y) = \dfrac{\alpha+1}{\alpha+\beta+Y} - \dfrac{(\alpha+2)(\alpha+1)}{(\alpha+\beta+Y+1)(\alpha+\beta+Y)}$

$$= \frac{(\alpha+1)(\beta+Y-1)}{(\alpha+\beta+Y+1)(\alpha+\beta+Y)},$$

where the second expectation was solved using the result from Ex. 4.200. (Alternately, the answer could be found by solving $E[p(1-p) \mid Y] = \int_0^1 p(1-p)g^*(p \mid Y)dp$.

16.10 **a.** The joint density of the random sample and θ is given by the product of the marginal densities multiplied by the gamma prior:

$$f(y_1,\dots,y_n,\theta) = \left[\prod_{i=1}^{n}\theta\exp(-\theta y_i)\right]\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\theta^{\alpha-1}\exp(-\theta/\beta)$$

$$= \frac{\theta^{n+\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}}\exp\left(-\theta\sum_{i=1}^{n}y_i-\theta/\beta\right) = \frac{\theta^{n+\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}}\exp\left(-\theta\Big/\frac{\beta}{\beta\sum_{i=1}^{n}y_i+1}\right)$$

b. $m(y_1,\dots,y_n) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}}\int_0^{\infty}\theta^{n+\alpha-1}\exp\left(-\theta\Big/\frac{\beta}{\beta\sum_{i=1}^{n}y_i+1}\right)d\theta$, but this integral resembles that of a gamma density with shape parameter $n + \alpha$ and scale parameter $\frac{\beta}{\beta\sum_{i=1}^{n}y_i+1}$.

Thus, the solution is $m(y_1,\dots,y_n) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}}\Gamma(n+\beta)\left(\frac{\beta}{\beta\sum_{i=1}^{n}y_i+1}\right)^{n+\alpha}$.

c. The solution follows from parts (a) and (b) above.

d. Using the result in Ex. 4.111,

$$\hat{\mu}_B = E(\mu\,|\,\boldsymbol{Y}) = E(1/\theta\,|\,\boldsymbol{Y}) = \frac{1}{\beta^*(\alpha^*-1)} = \left[\frac{\beta}{\beta\sum_{i=1}^{n}Y_i+1}(n+\alpha-1)\right]^{-1}$$

$$= \frac{\beta\sum_{i=1}^{n}Y_i+1}{\beta(n+\alpha-1)} = \frac{\sum_{i=1}^{n}Y_i}{n+\alpha-1}+\frac{1}{\beta(n+\alpha-1)}$$

e. The prior mean for $1/\theta$ is $E(1/\theta) = \frac{1}{\beta(\alpha-1)}$ (again by Ex. 4.111). Thus, $\hat{\mu}_B$ can be written as

$$\hat{\mu}_B = \overline{Y}\left(\frac{n}{n+\alpha-1}\right)+\frac{1}{\beta(\alpha-1)}\left(\frac{\alpha-1}{n+\alpha-1}\right),$$

which is a weighted average of the MLE and the prior mean.

f. We know that $\overline{Y}$ is unbiased; thus $E(\overline{Y}) = \mu = 1/\theta$. Therefore,

$$E(\hat{\mu}_B) = E(\overline{Y})\left(\frac{n}{n+\alpha-1}\right)+\frac{1}{\beta(\alpha-1)}\left(\frac{\alpha-1}{n+\alpha-1}\right) = \frac{1}{\theta}\left(\frac{n}{n+\alpha-1}\right)+\frac{1}{\beta(\alpha-1)}\left(\frac{\alpha-1}{n+\alpha-1}\right).$$

Therefore, $\hat{\mu}_B$ is biased. However, it is asymptotically unbiased since

$$E(\hat{\mu}_B)-1/\theta\to 0.$$

Also,

$$V(\hat{\mu}_B) = V(\overline{Y})\left(\frac{n}{n+\alpha-1}\right)^2 = \frac{1}{\theta^2 n}\left(\frac{n}{n+\alpha-1}\right)^2 = \frac{1}{\theta^2}\frac{n}{(n+\alpha-1)^2} \to 0\,.$$

So, $\hat{\mu}_B \xrightarrow{p} 1/\theta$ and thus it is consistent.

16.11 **a.** The joint density of U and λ is

$$\begin{aligned} f(u,\lambda) = p(u\,|\,\lambda)g(\lambda) &= \frac{(n\lambda)^u \exp(-n\lambda)}{u!} \times \frac{1}{\Gamma(\alpha)\beta^\alpha}\lambda^{\alpha-1}\exp(-\lambda/\beta) \\ &= \frac{n^u}{u!\Gamma(\alpha)\beta^\alpha}\lambda^{u+\alpha-1}\exp(-n\lambda-\lambda/\beta) \\ &= \frac{n^u}{u!\Gamma(\alpha)\beta^\alpha}\lambda^{u+\alpha-1}\exp\left[-\lambda\Big/\left(\frac{\beta}{n\beta+1}\right)\right] \end{aligned}$$

b. $m(u) = \dfrac{n^u}{u!\Gamma(\alpha)\beta^\alpha}\displaystyle\int_0^\infty \lambda^{u+\alpha-1}\exp\left[-\lambda\Big/\left(\frac{\beta}{n\beta+1}\right)\right]d\lambda$, but this integral resembles that of a gamma density with shape parameter $u + \alpha$ and scale parameter $\dfrac{\beta}{n\beta+1}$. Thus, the solution is $m(u) = \dfrac{n^u}{u!\Gamma(\alpha)\beta^\alpha}\Gamma(u+\alpha)\left(\dfrac{\beta}{n\beta+1}\right)^{u+\alpha}$.

c. The result follows from parts (a) and (b) above.

d. $\hat{\lambda}_B = E(\lambda\,|\,U) = \alpha^*\beta^* = (U+\alpha)\left(\dfrac{\beta}{n\beta+1}\right)$.

e. The prior mean for λ is $E(\lambda) = \alpha\beta$. From the above,

$$\hat{\lambda}_B = \left(\sum_{i=1}^n Y_i + \alpha\right)\left(\frac{\beta}{n\beta+1}\right) = \overline{Y}\left(\frac{n\beta}{n\beta+1}\right) + \alpha\beta\left(\frac{1}{n\beta+1}\right),$$

which is a weighted average of the MLE and the prior mean.

f. We know that $\overline{Y}$ is unbiased; thus $E(\overline{Y}) = \lambda$ Therefore,

$$E(\hat{\lambda}_B) = E(\overline{Y})\left(\frac{n\beta}{n\beta+1}\right) + \alpha\beta\left(\frac{1}{n\beta+1}\right) = \lambda\left(\frac{n\beta}{n\beta+1}\right) + \alpha\beta\left(\frac{1}{n\beta+1}\right).$$

So, $\hat{\lambda}_B$ is biased but it is asymptotically unbiased since

$$E(\hat{\lambda}_B) - \lambda \to 0.$$

Also,

$$V(\hat{\lambda}_B) = V(\bar{Y})\left(\frac{n\beta}{n\beta+1}\right)^2 = \frac{\lambda}{n}\left(\frac{n\beta}{n\beta+1}\right)^2 = \lambda\frac{n\beta}{(n\beta+1)^2} \to 0.$$

So, $\hat{\lambda}_B \xrightarrow{p} \lambda$ and thus it is consistent.

16.12 First, it is given that $W = vU = v\sum_{i=1}^{n}(Y_i - \mu_0)^2$ is chi–square with n degrees of freedom. Then, the density function for U (conditioned on v) is given by

$$f_U(u \mid v) = |v| f_W(uv) = v\frac{1}{\Gamma(n/2)2^{n/2}}(uv)^{n/2-1}e^{-uv/2} = \frac{1}{\Gamma(n/2)2^{n/2}}u^{n/2-1}v^{n/2}e^{-uv/2}.$$

a. The joint density of U and v is then

$$\begin{aligned} f(u,v) = f_U(u \mid v)g(v) &= \frac{1}{\Gamma(n/2)2^{n/2}}u^{n/2-1}v^{n/2}\exp(-uv/2)\times\frac{1}{\Gamma(\alpha)\beta^\alpha}v^{\alpha-1}\exp(-v/\beta) \\ &= \frac{1}{\Gamma(n/2)\Gamma(\alpha)2^{n/2}\beta^\alpha}u^{n/2-1}v^{n/2+\alpha-1}\exp(-uv/2 - v/\beta) \\ &= \frac{1}{\Gamma(n/2)\Gamma(\alpha)2^{n/2}\beta^\alpha}u^{n/2-1}v^{n/2+\alpha-1}\exp\left[-v\Big/\left(\frac{2\beta}{u\beta+2}\right)\right]. \end{aligned}$$

b. $m(u) = \dfrac{1}{\Gamma(n/2)\Gamma(\alpha)2^{n/2}\beta^\alpha}u^{n/2-1}\displaystyle\int_0^\infty v^{n/2+\alpha-1}\exp\left[-v\Big/\left(\frac{2\beta}{u\beta+2}\right)\right]dv$, but this integral resembles that of a gamma density with shape parameter $n/2 + \alpha$ and scale parameter $\dfrac{2\beta}{u\beta+2}$. Thus, the solution is $m(u) = \dfrac{u^{n/2-1}}{\Gamma(n/2)\Gamma(\alpha)2^{n/2}\beta^\alpha}\Gamma(n/2+\alpha)\left(\dfrac{2\beta}{u\beta+2}\right)^{n/2+\alpha}$.

c. The result follows from parts (a) and (b) above.

d. Using the result in Ex. 4.111(e),

$$\hat{\sigma}_B^2 = E(\sigma^2 \mid U) = E(1/v \mid U) = \frac{1}{\beta^*(\alpha^*-1)} = \frac{1}{n/2+\alpha-1}\left(\frac{U\beta+2}{2\beta}\right) = \frac{U\beta+2}{\beta(n+2\alpha-2)}.$$

e. The prior mean for $\sigma^2 = 1/v = \dfrac{1}{\beta(\alpha-1)}$. From the above,

$$\hat{\sigma}_B^2 = \frac{U\beta+2}{\beta(n+2\alpha-2)} = \frac{U}{n}\left(\frac{n}{n+2\alpha-2}\right) + \frac{1}{\beta(\alpha-1)}\left(\frac{2(\alpha-1)}{n+2\alpha-2}\right).$$

16.13 **a.** (.099, .710)
b. Both probabilities are .025.

c. $P(.099 < p < .710) = .95$.
d.-g. Answers vary.
h. The credible intervals should decrease in width with larger sample sizes.

16.14 **a.-b.** Answers vary.

16.15 With $y = 4$, $n = 25$, and a beta(1, 3) prior, the posterior distribution for p is beta(5, 24). Using R, the lower and upper endpoints of the 95% credible interval are given by:

```
> qbeta(.025,5,24)
[1] 0.06064291
> qbeta(.975,5,24)
[1] 0.3266527
```

16.16 With $y = 4$, $n = 25$, and a beta(1, 1) prior, the posterior distribution for p is beta(5, 22). Using R, the lower and upper endpoints of the 95% credible interval are given by:

```
> qbeta(.025,5,22)
[1] 0.06554811
> qbeta(.975,5,22)
[1] 0.3486788
```

This is a wider interval than what was obtained in Ex. 16.15.

16.17 With $y = 6$ and a beta(10, 5) prior, the posterior distribution for p is beta(11, 10). Using R, the lower and upper endpoints of the 80% credible interval for p are given by:

```
> qbeta(.10,11,10)
[1] 0.3847514
> qbeta(.90,11,10)
[1] 0.6618291
```

16.18 With $n = 15$, $\sum_{i=1}^{n} y_i = 30.27$, and a gamma(2.3, 0.4) prior, the posterior distribution for θ is gamma(17.3, .030516). Using R, the lower and upper endpoints of the 80% credible interval for θ are given by

```
> qgamma(.10,shape=17.3,scale=.0305167)
[1] 0.3731982
> qgamma(.90,shape=17.3,scale=.0305167)
[1] 0.6957321
```

The 80% credible interval for θ is (.3732, .6957). To create a 80% credible interval for $1/\theta$, the end points of the previous interval can be inverted:

$$.3732 < \theta < .6957$$
$$1/(.3732) > 1/\theta > 1/(.6957)$$

Since $1/(.6957) = 1.4374$ and $1/(.3732) = 2.6795$, the 80% credible interval for $1/\theta$ is (1.4374, 2.6795).

16.19 With $n = 25$, $\sum_{i=1}^{n} y_i = 174$, and a gamma(2, 3) prior, the posterior distribution for λ is gamma(176, .0394739). Using R, the lower and upper endpoints of the 95% credible interval for λ are given by

```
> qgamma(.025,shape=176,scale=.0394739)
[1] 5.958895
> qgamma(.975,shape=176,scale=.0394739)
[1] 8.010663
```

16.20 With $n = 8$, $u = .8579$, and a gamma(5, 2) prior, the posterior distribution for v is gamma(9, 1.0764842). Using R, the lower and upper endpoints of the 90% credible interval for v are given by

```
> qgamma(.05,shape=9,scale=1.0764842)
[1] 5.054338
> qgamma(.95,shape=9,scale=1.0764842)
[1] 15.53867
```

The 90% credible interval for v is (5.054, 15.539). Similar to Ex. 16.18, the 90% credible interval for $\sigma^2 = 1/v$ is found by inverting the endpoints of the credible interval for v, given by (.0644, .1979).

16.21 From Ex. 6.15, the posterior distribution of p is beta(5, 24). Now, we can find $P^*(p \in \Omega_0) = P^*(p < .3)$ by (in R):

```
> pbeta(.3,5,24)
[1] 0.9525731
```

Therefore, $P^*(p \in \Omega_a) = P^*(p \geq .3) = 1 - .9525731 = .0474269$. Since the probability associated with H_0 is much larger, our decision is to not reject H_0.

16.22 From Ex. 6.16, the posterior distribution of p is beta(5, 22). We can find $P^*(p \in \Omega_0) = P^*(p < .3)$ by (in R):

```
> pbeta(.3,5,22)
[1] 0.9266975
```

Therefore, $P^*(p \in \Omega_a) = P^*(p \geq .3) = 1 - .9266975 = .0733025$. Since the probability associated with H_0 is much larger, our decision is to not reject H_0.

16.23 From Ex. 6.17, the posterior distribution of p is beta(11, 10). Thus, $P^*(p \in \Omega_0) = P^*(p < .4)$ is given by (in R):

```
> pbeta(.4,11,10)
[1] 0.1275212
```

Therefore, $P^*(p \in \Omega_a) = P^*(p \geq .4) = 1 - .1275212 = .8724788$. Since the probability associated with H_a is much larger, our decision is to reject H_0.

16.24 From Ex. 16.18, the posterior distribution for θ is gamma(17.3, .0305). To test

$$H_0: \theta > .5 \text{ vs. } H_a: \theta \leq .5,$$

we calculate $P^*(\theta \in \Omega_0) = P^*(\theta > .5)$ as:

```
> 1 - pgamma(.5,shape=17.3,scale=.0305)
[1] 0.5561767
```

Therefore, $P^*(\theta \in \Omega_a) = P^*(\theta \geq .5)$ = 1 – .5561767 = .4438233. The probability associated with H_0 is larger (but only marginally so), so our decision is to not reject H_0.

16.25 From Ex. 16.19, the posterior distribution for λ is gamma(176, .0395). Thus, $P^*(\lambda \in \Omega_0) = P^*(\lambda > 6)$ is found by

```
> 1 - pgamma(6,shape=176,scale=.0395)
[1] 0.9700498
```

Therefore, $P^*(\lambda \in \Omega_a) = P^*(\lambda \leq 6)$ = 1 – .9700498 = .0299502. Since the probability associated with H_0 is much larger, our decision is to not reject H_0.

16.26 From Ex. 16.20, the posterior distribution for v is gamma(9, 1.0765). To test:

$$H_0: v < 10 \text{ vs. } H_a: v \geq 10,$$

we calculate $P^*(v \in \Omega_0) = P^*(v < 10)$ as

```
> pgamma(10,9, 1.0765)
[1] 0.7464786
```

Therefore, $P^*(\lambda \in \Omega_a) = P^*(v \geq 10)$ = 1 – .7464786 = .2535214. Since the probability associated with H_0 is larger, our decision is to not reject H_0.