Mathematics 640: Bayesian Statistics Homework 2

Arif Ali

02/17/16

Exercise 1

$\hat{\mu_c} = 1.013,\ \hat{\sigma_c} = 0.24,\ n_c = 32$

$\hat{\mu_t} = 1.173,\ \hat{\sigma_t} = 0.20,\ n_t = 36$

Part A

$y_{c_i} \sim i.i.d\ N(\mu_c, \sigma_c^2)$

$y_{t_i} \sim i.i.d\ N(\mu_t, \sigma_t^2)$

$p(y|\mu_c, \mu_t, \sigma_t, \sigma_c) = \prod_{i=1}^{32} N(y_i|\mu_c, \sigma_c^2) * \prod_{i=1}^{36} N(y_i|\mu_t, \sigma_t^2)$

Since there is a uniform prior distribution on $(\mu_c, \mu_t, log\sigma_c, log\sigma_t)$, the the posterior distribution of $(\mu_c, \mu_t, log\sigma_c, log\sigma_t)$ follows the distibution:

$\mu|y \sim t(n-1, \bar{y}, s^2/n)$

Because "under the noninformative uniform prior distribution on $(\mu,\ log(\sigma))$, the posterior distribution of μ has the form:

$\frac{\mu-\bar{y}}{s/\sqrt{n}}|y \sim t_{n-1}$" (Gelman pg. 66)

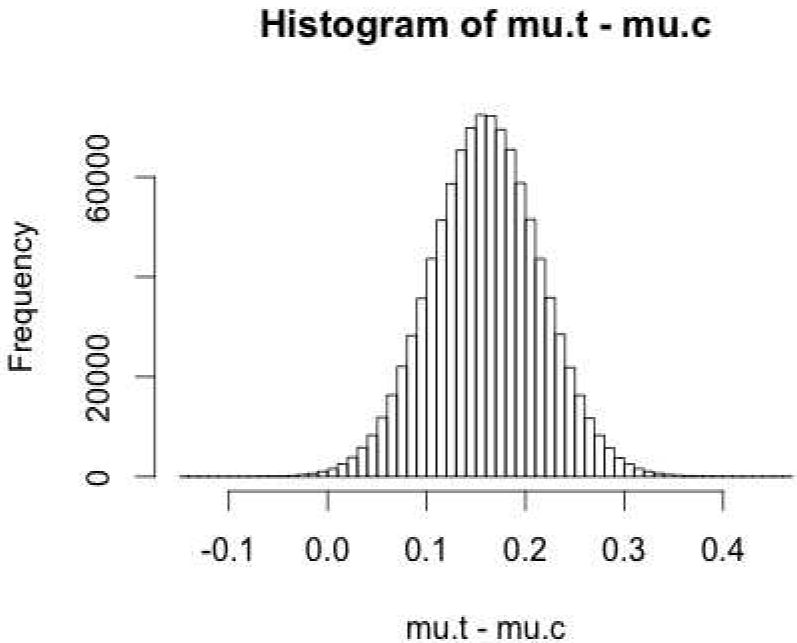
Under the assumption that the samples from the control and treatment are independent, the respective pairs of $y_{c_i} \sim i.i.d\ N(\mu_c, \sigma_c^2)$ and $y_{t_i} \sim i.i.d\ N(\mu_t, \sigma_t^2)$ can also be treated independently.

$\therefore \mu_c|y \sim t(32-1, 1.013, (0.24)^2/32) \equiv \mu_c|y \sim t(31, 1.013, (0.24)^2/32)$

and $\mu_t|y \sim t(36-1, 1.173, (0.20)^2/36) \equiv \mu_c|y \sim t(35, 1.173, (0.20)^2/36)$

Part B

```
mu.c = rt(1e6, df=31)*sqrt(0.24^2/32)+1.013
mu.t = rt(1e6, df=35)*sqrt(0.2^2/36)+1.173
quantile(mu.t-mu.c, c(0.025,0.975))
hist(mu.t-mu.c, breaks = 50)
```



The 95% Posterior interval is (0.05042445 , 0.26938918)

Exercise 2

Part A

$y_1 \ldots y_n \sim iid\ exp\,(\theta)$

$p(y|\theta) = \prod_{i=1}^{n} \theta e^{-\theta * y_i} = (\theta)^n \, e^{-\theta \sum_{i=1}^{n} y_i}$

$log(p(y|\theta)) = n * log(\theta) - \theta \sum_{i=1}^{n} y_i$

$\delta/d\theta(log(p(y|\theta)) = n/\theta$

$\delta^2/d\theta^2(log(p(y|\theta)) = -n/\theta^2$

$I(\theta) = -E\left(-n/\theta^2\right) = \frac{n}{\theta^2}$

$(I(\theta))^{1/2} = \frac{\sqrt{n}}{\theta} = \sqrt{n} * \theta^{-1} = p(\theta)$

$\int_0^\infty \sqrt{n} * \theta^{-1} \delta\theta = \sqrt{n} * \int_0^\infty \theta^{-1} \delta\theta = \infty$

Therefore, this is an improper prior.

Part B

$p(y|\theta, n = 10, \sum_{i=1}^{n} y_i = 1012) = \prod_{i=1}^{n} \theta e^{-\theta * y_i} = (\theta)^{10} \, e^{-1012 * \theta}$

$p(\theta|y) = \frac{(\theta)^{10} e^{-1012*\theta} * \sqrt{n} * \theta^{-1}}{\int_0^\infty (\theta)^{10} e^{-1012*\theta} * \sqrt{n} * \theta^{-1} \delta\theta} = \frac{(\theta)^9 e^{-1012*\theta}}{\int_0^\infty (\theta)^9 e^{-1012*\theta} \delta\theta} =$

$\frac{(\theta)^9 e^{-1012*\theta}}{\int_0^\infty (\theta)^9 e^{-1012*\theta} \delta\theta} = \frac{\frac{1012^{10}}{\Gamma(10)} * (\theta)^9 e^{-1012*\theta}}{\frac{1012^{10}}{\Gamma(10)} * \int_0^\infty (\theta)^{10-1} e^{-1012*\theta} \delta\theta} =$

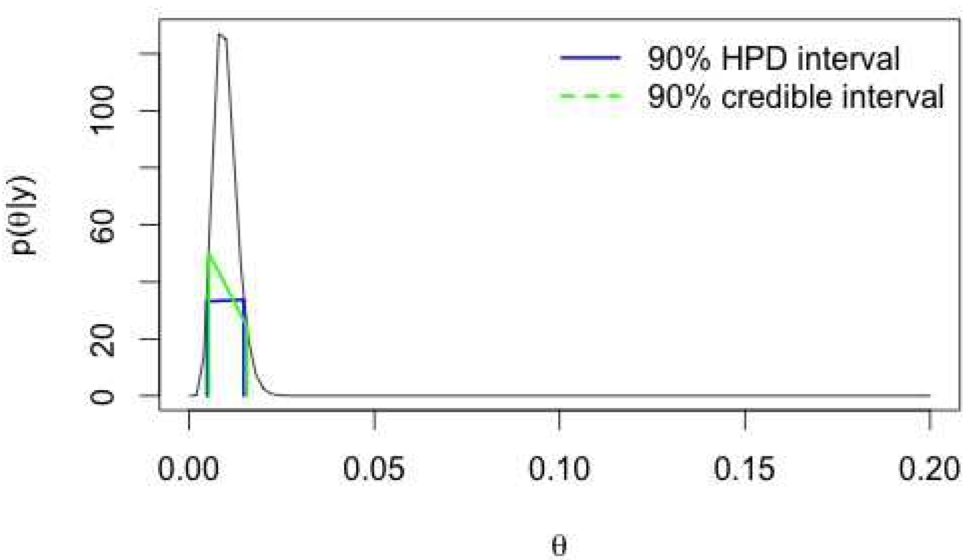
$\frac{\frac{1012^{10}}{\Gamma(10)} * (\theta)^9 e^{-1012*\theta}}{\int_0^\infty \frac{1012^{10}}{\Gamma(10)} * (\theta)^{10-1} e^{-1012*\theta} \delta\theta} = \frac{1012^{10}}{\Gamma(10)} * (\theta)^{10-1} e^{-1012*\theta}$

$\therefore y|\theta \sim gamma(10, 1012)$

Please note that $\int_0^\infty \frac{1012^{10}}{\Gamma(10)} * (\theta)^{10-1} e^{-1012*\theta} \delta\theta = 1 \; since \; \frac{1012^{10}}{\Gamma(10)} * (\theta)^{10-1} e^{-1012*\theta} \sim gamma(10, 1012)$

Part C

```
library(coda)
th = as.mcmc(rgamma(1e6,10,1012))
th.hpd = HPDinterval(th, prob = 0.90)
th.hpd th.ci = qgamma(c(0.05,0.95),10,1012)
curve(dgamma(x, 10,1012), from=0, to=0.2, xlab=expression(theta),
          ylab=expression(paste("p(",theta,"|y)")))
          lines(x=c(th.hpd[1], th.hpd[1], th.hpd[2], th.hpd[2]),
          y = dgamma(c(0, th.hpd[1], th.hpd[2], 0), 10, 1012), col="blue", lwd=2, lty=1)
lines(x=c(th.ci[1], th.ci[1], th.ci[2], th.ci[2]),
          y = dgamma(c(0, th.ci[1], th.ci[2], 0), 10, 1012), col="green", lwd=2, lty=1)
legend("topright", c("90% HPD interval", "90% credible interval"),
                    col=c("blue", "green"), lty=c(1,2), lwd=c(2,2), bty="n")
```



Part D

i.

$p(y|\theta, n = 11, \sum_{i=1}^{n} y_i = 1012 + 95) = (\theta)^{11} \, e^{-(1012+95)*\theta}$

$p(\theta|y_1 \ldots y_{11}) = \frac{(\theta)^{11} e^{-(1012+95)*\theta} * \sqrt{n} * \theta^{-1}}{\int_0^\infty (\theta)^{11} e^{-(1012+95)*\theta} * \sqrt{n} * \theta^{-1} \delta\theta} =$

$\frac{\frac{(1012+95)^{11}}{\Gamma(11)} * (\theta)^{10} e^{-(1012+95)*\theta}}{\frac{(1012+95)^{11}}{\Gamma(1`)} * \int_0^\infty (\theta)^{11-1} e^{-(1012+95)*\theta} \delta\theta} =$

$\frac{(1012+95)^{11}}{\Gamma(11)} * (\theta)^{11-1} \, e^{-(1012+95)*\theta} \therefore y|\theta \sim gamma(11, 1012 + 95)$

ii.

$\pi(\theta) = p(\theta|y_1 \ldots y_{10}) = \frac{1012^{10}}{\Gamma(10)} * (\theta)^{10-1} \, e^{-1012*\theta}$

$L(\theta|y_{11}) = \theta e^{-95\theta}$

$\pi(\theta) * L(\theta|y_{11}) = \frac{1012^{10}}{\Gamma(10)} * (\theta)^{10-1} \, e^{-1012*\theta} * \theta e^{-95\theta} = \frac{1012^{10}}{\Gamma(10)} * (\theta)^{11-1} e^{-(1012+95)\theta}$

$p(\theta|y_1 \ldots y_{11}) = \frac{\pi(\theta) * L(\theta|y_{11})}{\int \pi(\theta) * L(\theta|y_{11}) d\theta} = \frac{\frac{1012^{10}}{\Gamma(10)} * (\theta)^{11-1} e^{-(1012+95)\theta}}{\int \frac{1012^{10}}{\Gamma(10)} * (\theta)^{11-1} e^{-(1012+95)\theta} d\theta} =$

$\frac{\frac{(1012+95)^{11}}{\Gamma(11)} * (\theta)^{11-1} e^{-(1012+95)\theta}}{\int \frac{(1012+95)^{11}}{\Gamma(11)} * (\theta)^{11-1} e^{-(1012+95)\theta} d\theta} =$

$\frac{(1012+95)^{11}}{\Gamma(11)} * (\theta)^{11-1} \, e^{-(1012+95)*\theta} \therefore y|\theta \sim gamma(11, 1012 + 95)$

Part E

$P(\tilde{y} = z|y_1.....y_{11}) = \int p(\tilde{y} = z|\theta) * p(\theta|y_1.....y_{11}) =$

$\int \theta e^{-\theta * z} * \frac{(1012+95)^{11}}{\Gamma(11)} * (\theta)^{11-1} e^{-(1012+95)\theta} \delta\theta =$

$\int \frac{(1012+95)^{11}}{\Gamma(11)} * (\theta)^{12-1} e^{-(1012+95+z)\theta} \delta\theta =$

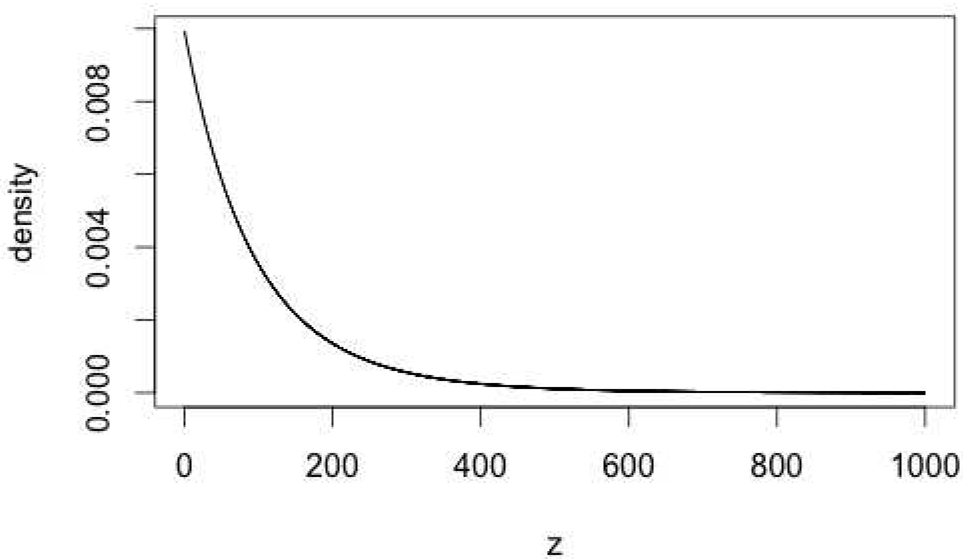
$\frac{11*(1012+95+z)^{12}}{11*(1012+95+z)^{12}} * \int \frac{(1012+95)^{11}}{\Gamma(11)} * (\theta)^{12-1} e^{-(1012+95+z)\theta} \delta\theta =$

$\frac{11*(1012+95)^{11}}{(1012+95+z)^{12}} * \int \frac{(1012+95+z)^{12}}{\Gamma(12)} * (\theta)^{12-1} e^{-(1012+95+z)\theta} \delta\theta =$

$\frac{11*(1012+95)^{11}}{(1012+95+z)^{12}}$

because $\int \frac{(1012+95+z)^{12}}{\Gamma(12)} * (\theta)^{12-1} e^{-(1012+95+z)\theta} \delta\theta \sim Gamma(12, 1012 + 95 + z)$

```
z = seq(0,1000, 0.1)
density=11*(1012+95)^11/(1012+95+z)^12
plot(z,density, type = "l")
```



Exercise 3

Part A

From the Multiparameter lecture, slide 18

$\mu|\sigma^2 \sim N(\mu_0, \sigma^2/\kappa_0)$

$\sigma^2 \sim inv - gamma\left(\frac{\nu_0}{2}, \frac{\nu_0\sigma_0^2}{2}\right)$

corresponds to the joint-prior density

$p(\mu, \sigma^2) = (\sigma^2)^{-\left(\frac{\nu_0+1}{2}-1\right)} exp\left\{-\frac{\kappa_0}{2\sigma^2}\left(\frac{\nu_0\sigma_0^2}{\kappa_0} + (\mu-\mu_0\right)^2\right\}$

$p(\mu, \sigma^2|y) = (\sigma^2)^{-\left(\frac{\nu_n+1}{2}-1\right)} exp\left\{-\frac{\kappa_n}{2\sigma^2}\left(\frac{\sigma_n^2}{\kappa_n} + (\mu-\mu_n\right)^2\right\}$

$\sim Normal - Inverse\ Gamma(\mu_0, \sigma_0^2/\kappa_0; \nu_0/2, \nu_0\sigma_0^2/2)$

In slide 20, the hyperparameters were defined as:

$\mu_n = \frac{\kappa_0}{\kappa_0+n}\mu_0 + \frac{n}{\kappa_0+n}\bar{y}$

$\kappa_n = \kappa_0 + n$

$\nu_n = \nu_0 + n$

$\nu_n\sigma_n^2 = \nu_0\sigma_0^2 + (n-1)s^2 + \frac{\kappa_0 n}{\kappa_0+n}(\bar{y}-\mu_0)^2$

```
school = read.table("school.dat")[,1]
ybar = mean(school); s2 = var(school); n = length(school)
m0 = 0; k0 = 0.1; v0 = 10; s0 = 4
kn = k0 + n
mn = (k0*m0 + n*ybar)/kn
vn = v0 + n
sn = (v0*s0 + (n-1)*s2 + (k0*n/kn)*(ybar - m0)^2)/vn
phi  = rgamma(100000, vn/2, rate=vn/2*sn)
sigma2 = 1/phi
mu   = rnorm(100000, mn, sqrt(sigma2/kn))
```

Part B

```
quantile(mu, c(0.025,0.975))

quantile(sigma2, c(0.025,0.975))
```

The 95% credible interval for is μ **is** $(8.030065, 10.813879)$

The 95% credible interval for is σ^2 **is** $(7.711269, 19.962460)$

Part C

From slide 21 of the Multi-parameter lecture, the marginal posterior distribution is defined as:

$\mu|y \sim t\left(\nu_n, \mu_n, \frac{\sigma_n^2}{\kappa_n}\right)$

$E(\mu|y) = \mu_n$ since $\nu_n = v_0 + n = 10 + 25 = 35 > 1$

$\therefore$the posterior mean is $\frac{(\kappa_0*\mu_0+n*\bar{y})}{\kappa_n} = \frac{0.1*0+25*9.464}{25.1} = 9.426295$

```
mn + qt(c(0.025,0.975), df = vn) * sqrt(sn/kn)
```

The 95% credible interval is $(8.037325, 10.815264)$.

Part D

$p(\tilde{y}|y_1 \dots y_n) = \int_{\sigma^2} \int_{\mu} f(\tilde{y}|\mu, \sigma^2) * f(\mu, \sigma^2|y_1 \dots y_n) d\mu \, d\sigma^2$

where $f(\tilde{y}|\mu, \sigma^2) \sim N(\mu, \sigma^2)$

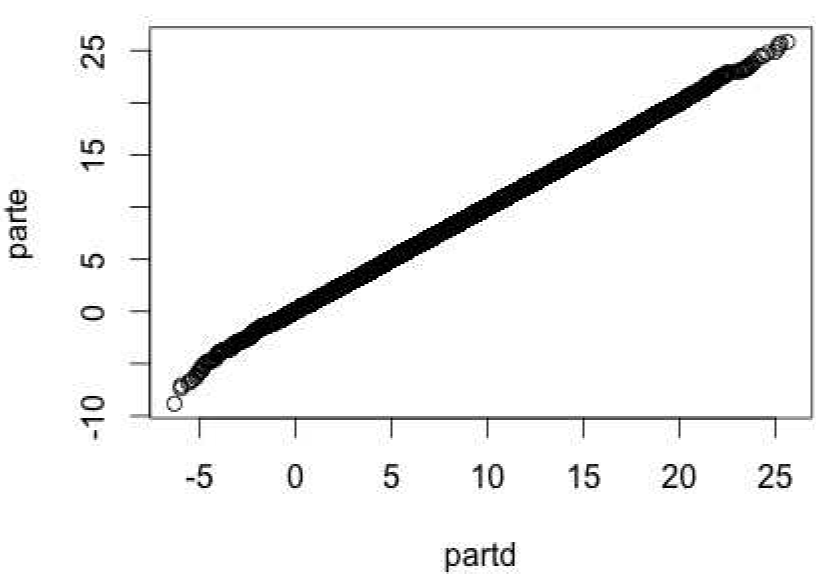
The first step involves obtaining samples of μ and σ^2 from the joint posterior distribution, which was done in Part A. The second step involves inputing those sampled μ and σ^2 into $f(\tilde{y}|\mu, \sigma^2) \sim N(\mu, \sigma^2)$.

```
partd = rnorm(100000, mu, sqrt(sigma2))
```

Part E

From Slide 21, the posterior predictive distribution for a future observation is

```
parte = mn + rt(100000, df = vn)*sqrt(sn*(1+1/kn))
qqplot(partd,parte)
```



Based in the qqplot, the distributions sampled from Part D and Part E are equivalent. This is because the points indicate that that is a direct relationship (no weights) between Part D and Part E.