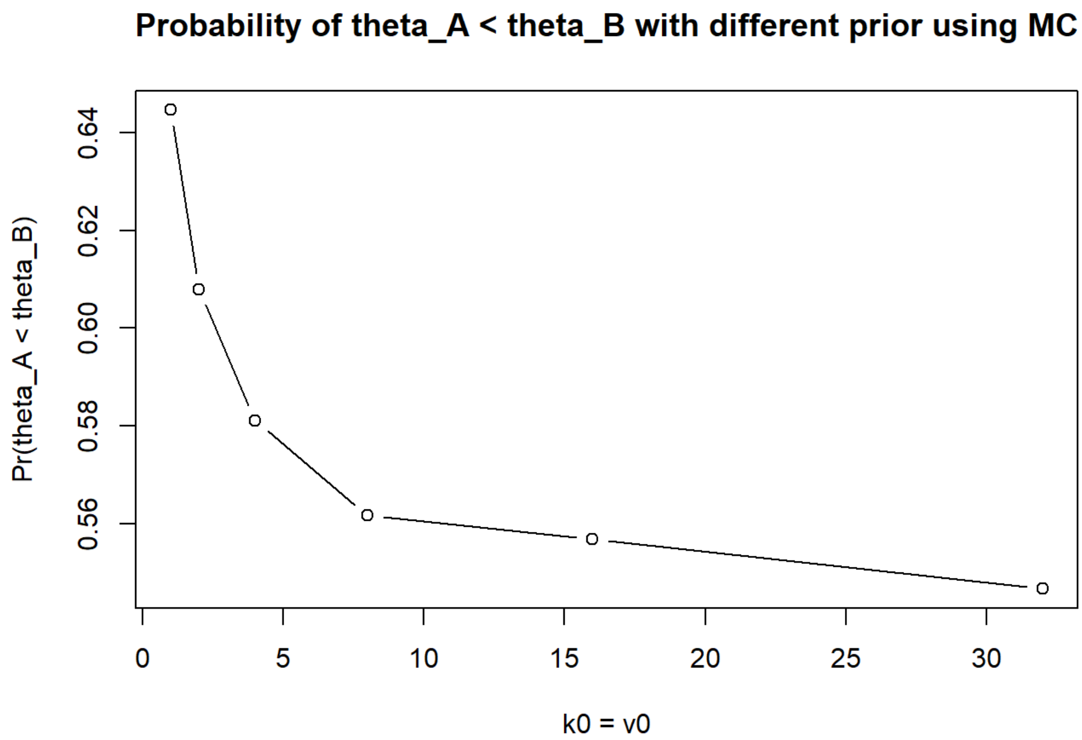
Homework3

Bingying Liu

February 4, 2020

Question 1: Hoff 5.2

```r
yA <- 75.2; yB <- 77.5
sA <- 7.3; sB <- 8.1
nA <- 16; nB <- 16
mu0 <- 10; sigma20 <- 100

N = 5
Pr = c()
for (i in 0:N){
  k0 = v0 = 2^i

  k_nA <- k0 + nA; k_nB <- k0 + nB
  v_nA <- v0 + nA; v_nB <- k0 + nB

  mu_nA <- (k0*mu0 + nA*yA)/k_nA
  mu_nB <- (k0*mu0 + nB*yB)/k_nB
  sigma2_nA <- (1/v_nA) * (v0*sigma20 + sA^2*(nA-1) + (nA*k0/k_nA)*(yA-mu0)^2)
  sigma2_nB <- (1/v_nB) * (v0*sigma20 + sB^2*(nB-1) + (nB*k0/k_nB)*(yB-mu0)^2)

  # Posterior
  S = 10000
  tauA_postsample <- rgamma(S, k_nA/2, k_nA*sigma2_nA/2)
  thetaA_postsample <- rnorm(S, mu_nA, sqrt(1/(k_nA*tauA_postsample)))
  tauB_postsample <- rgamma(S, k_nB/2, k_nB*sigma2_nB/2)
  thetaB_postsample <- rnorm(S, mu_nB, sqrt(1/(k_nB*tauB_postsample)))

  # MC Sampling
  Pr <- c(Pr, mean(thetaA_postsample < thetaB_postsample))
}

plot(2^(0:(N)), Pr, "b", xlab = "k0 = v0", ylab = "Pr(theta_A < theta_B)",
     main="Probability of theta_A < theta_B with different prior using MC")
```



- Describe how you might use this plot to convey the evidence that $theta_A < theta_B$ to people of a variety of prior opinions.

The sample mean of B is larger than sample mean of A and both of them are larger than the prior mean. As prior belief becomes stronger, in this case the greater the prior sample size k_0 and prior degrees of freedom v_0, the more influence it is to be incorporated into posterior mean and variance. As shown in the plot, the probability of $\theta_A - \theta_B$ is greater than 0.53 and converges to 0.53 for larger prior sample size or prior degrees of freedom. This means θ_A is less than θ_B with a large confidence.

Q2:

$Y = (y_1, y_2, \cdots, y_n)$

$y_i \sim N(\theta, \sigma^2)$.

Jeffrey's prior: $\pi(\theta, \sigma^2) \propto (\sigma^2)^{-\frac{3}{2}}$.

Derive posterior under this prior and state whether it's proper. What happens when $n=1$ vs. $n>1$?

$$p(\theta, \sigma^2 | Y) \propto p(Y|\theta,\sigma^2) \cdot p(\theta, \sigma^2)$$

$$\propto (\sigma^2)^{-\frac{3}{2}} \cdot p(y_1 \cdots, y_n | \theta, \sigma^2)$$

$$\propto (\sigma^2)^{-\frac{3}{2}} (\sigma^2)^{-\frac{n}{2}} \cdot \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \theta)^2\right)$$

$$\propto (\sigma^2)^{-\frac{n+3}{2}} \exp\left(\frac{\sum_{i=1}^n y_i^2 - 2\theta \cdot \sum_{i=1}^n y_i + n\theta^2}{-2\sigma^2}\right)$$

$$\propto \left(\frac{1}{\sigma^2}\right)^{\frac{n+3}{2}} \exp\left(\frac{n\left(\theta - 2\theta \frac{\sum_{i=1}^n y_i}{n} + \frac{\sum_{i=1}^n y_i}{n^2}\right) + \left(\sum_{i=1}^n y_i^2 - \frac{\sum_{i=1}^n y_i^2}{n}\right)}{-2\sigma^2}\right)$$

$$\propto \exp\left(\frac{\left(\theta - \frac{\sum_{i=1}^n y_i}{n}\right)^2}{-2\frac{\sigma^2}{n}}\right) \cdot \left(\frac{1}{\sigma^2}\right)^{\frac{n+5}{2}-1} \exp\left(\frac{\sum_{i=1}^n y_i^2 - \frac{\sum_{i=1}^n y_i^2}{n}}{-2\sigma^2}\right).$$

When $n=1$:

$$p(\theta, \sigma^2 | Y) \propto \exp\left(\frac{(\theta - \sum_{i=1}^n y_i)^2}{-2\sigma^2}\right) \cdot \left(\frac{1}{\sigma^2}\right)^2.$$

This posterior doesn't contain the same distribution as prior's (normal & gamma).

Also, since $p(\sigma^2|Y) = \int p(\theta, \sigma^2 | Y)\, d\theta \propto \int \exp\frac{(\theta - \sum_{i=1}^n y_i)^2}{-2\sigma^2} \left(\frac{1}{\sigma^2}\right)^2 d\theta$ is a function of both θ and σ^2, this posterior is not proper.

When $n \neq 1$:

$$p(\theta, \sigma^2 | Y) \propto N\left(\frac{\sum_{i=1}^n y_i}{n}, \frac{\sigma^2}{n}\right) \times \text{Gamma}\left(\frac{n+5}{2}, \frac{1}{2}\left(\sum_{i=1}^n y_i^2 - \frac{\sum_{i=1}^n y_i^2}{2}\right)\right)$$

The posterior's distribution for Jeffrey's prior is a Normal-inverse-gamma distrib with parameters derived above. Thus when $n \neq 1$, it's a proper posterior distribution

Q3: $y|\theta \sim Po(\theta)$.

1) Find the Jeffrey's prior for θ. $\pi(\theta) \propto \sqrt{I(\theta)}$

$\because$ $I(\theta) = -E\left[\frac{\partial^2}{\partial^2\theta} \log p(y|\theta) \middle| \theta\right]$ and $p(y|\theta) \sim Po(\theta)$.

$\therefore$ $\log p(y|\theta) = \log \frac{\theta^y \cdot e^{-\theta}}{y!} = y\log\theta - \theta - \log y!$

$\therefore$ $\frac{\partial^2}{\partial^2\theta} \log p(y|\theta) = \frac{\partial}{\partial\theta}\left(\frac{y}{\theta} - 1\right) = -y \cdot \theta^{-2} = -\frac{y}{\theta^2}$.

$\therefore$ $I(\theta) = -E\left[\frac{\partial^2}{\partial^2\theta} \log p(y|\theta) \middle| \theta\right] = -E\left[-\frac{y}{\theta^2}\right] = \frac{\theta}{\theta^2} = \frac{1}{\theta}$?

Then, we have Jeffrey's prior $\propto \sqrt{I(\theta)} = \frac{1}{\sqrt{\theta}}$.

Since $\int_0^{+\infty} \frac{1}{\sqrt{\theta}} d\theta = \left.\frac{\theta^{\frac{1}{2}}}{\frac{1}{2}}\right|_0^{\infty} = 2\sqrt{\theta}\Big|_0^{\infty} = \infty$, this distribution doesn't integrate to 1, it's improper.

2). What values of a and b for the gamma density $Ga(a,b)$ will result in a close match to the Jeffrey's density you found?

$f(\theta) = \frac{1}{\sqrt{\theta}} = \theta^{-\frac{1}{2}} = \theta^{\frac{1}{2}-1} \cdot e^{-0\cdot\theta} \propto Ga(\frac{1}{2}, 0)$.

$\therefore$ $Ga(\frac{1}{2}, 0)$ will result in a close match.

Q4: y_i, $\begin{cases} x_i=0 & \text{for control subjects.} \\ x_i=1 & \text{for treated subjects.} \end{cases}$, $y_i \sim \text{Poisson}(\lambda \cdot \gamma^{x_i})$. $\begin{cases} \lambda = E[y_i \mid x_i=0] \\ \gamma: \text{multiplicative change in the mean in treated gp.} \end{cases}$

$\lambda \sim Ga(1,1)$.

Is the joint posterior for λ and γ conjugate? $\gamma \sim Ga(1,1)$.

$p(\lambda,\gamma \mid y_1 \cdots y_n) \propto p(y_1, \cdots y_n \mid \lambda, \gamma) \cdot p(\lambda) \cdot p(\gamma)$. since λ and γ are independent in assumpti[on]

$$\propto \prod_{i=1}^{m} e^{-\lambda\gamma} \cdot (\lambda\gamma)^{y_i} \cdot \prod_{i=1}^{n-m} e^{-\lambda} \cdot (\lambda)^{y_i} \cdot \lambda^{1-1} \cdot e^{-\lambda} \cdot \gamma^{1-1} \cdot e^{-\gamma}.$$

m is number of subjects in treated gp, $m \in [0,n]$

$$\propto e^{-m\lambda\gamma} \cdot (\lambda\gamma)^{\sum_{i=1}^{m} y_i} \, e^{-(n-m)\lambda} \cdot \lambda^{\sum_{i=1}^{n-m} y_i} \cdot e^{-(\lambda+\gamma)}.$$

$$= \left(e^{-m\lambda\gamma-\gamma} \cdot \gamma^{\sum_{i=1}^{m} y_i}\right) \cdot \left(e^{-(n-m+1)\lambda} \cdot \lambda^{\sum_{i=1}^{m} y_i + \sum_{i=1}^{n-m} y_i}\right).$$

$$= \left(e^{-(m\lambda+1)\gamma} \cdot \gamma^{\left(\sum_{i=1}^{m} y_i+1\right)-1}\right) \cdot \left(e^{-(n-m+1)\lambda} \, \lambda^{\left(\sum_{i=1}^{n} y_i+1\right)-1}\right).$$

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

Yes, the joint posterior for λ and γ is conjugate to the priors, since it's a Gamma ~ Gamma distribution.