HW3 - Econometrics

Himanshu, MDS202327

2025-04-08

Question 1

(a) Write down the *p.m.f.* for Y_1 and Y_2.

Let θ be an unknown parameter satisiying $0 \leq \theta \leq 1$. Consider the following two experiments involving θ.

$\varepsilon_1 = \{Y_1, \theta, f(y_1 \mid \theta)\}$, a binomial experiment in which a coin is flipped T_1 times, where T_1 is predetermined and Y_1 is number of 'heads' observed in T_1 flips.

$\varepsilon_2 = \{Y_2, \theta, f(y_2 \mid \theta)\}$, a negative binomial experiment in which a coin is flipped until m 'tails' are observed, where $m > 0$ is predetermined and Y_2 is number of 'heads' observed in the process flips.

Now, the *p.m.f.* for Y_1 is given by

$$f_1(y_1 \mid \theta) = \binom{T_1}{Y_1} \theta^{Y_1} (1-\theta)^{T_1 - Y_1}$$

and, the *p.m.f.* for Y_2 is given by

$$f_1(y_2 \mid \theta) = \binom{m + Y_2 - 1}{Y_2} \theta^{m} (1-\theta)^{Y_2}$$

(b) Suppose $T_1 = 12$ and $m = 3$, and that the two experiments yield the following results. $y_1 = y_2 = 9$. Based on this information, write down the likelihood for the two experiments. What can you say based on the Likelihood Principle?

Considering the results of the experiment,

$f_1(y_1 \mid \theta_1) = \binom{12}{9} \theta_1^9 (1-\theta_1)^3 \propto \theta_1^9 (1-\theta_1)^3$ and,

$f_2(y_2 \mid \theta_2) = \binom{11}{9} \theta_2^3 (1-\theta_2)^9 \propto \theta_2^3 (1-\theta_2)^9 \propto (1-\theta_1)^3 \theta_1^9 \propto f_1(y_1 \mid \theta_1)$

where θ_1 and θ_2 are probability of 'heads' and 'tails' respectively as for experiments ε_1 and ε_2, 'heads' and 'tails' are notion of 'success' respectively.

Since, both likelihoods $f_1(y_1 \mid \theta_1)$ and $f_2(y_2 \mid \theta_2)$ are proportional, according to the Likelihood Principle, similar inferences can be made about θ (θ_1 or θ_2) through either experiments.

Question 2

(a) Show that the Pareto distribution is a conjugate prior for the uniform distribution,

$$\boldsymbol{\pi(\theta) = \begin{cases} ak^a\theta^{-(a+1)} & \theta >= k, a > 0 \\ 0 & \text{otherwise} \end{cases}}$$

The uniform distribution is given by,

$$f(y_i \mid \theta) = \begin{cases} \frac{1}{\theta} & 0 \le y_i \le \theta \\ 0 & \text{otherwise} \end{cases} = \frac{1}{\theta} I(y_i \le \theta)$$

Hence, the likelihood is given by

$$f(y \mid \theta) = \prod_{i=1}^{n} f(y_i \mid \theta) = \frac{1}{\theta^n} I(\max_i(y_i) \le \theta)$$

The prior distribution can be written as,

$$\pi(\theta) = ak^a\theta^{-(a+1)} \cdot I(k \le \theta) I(a > 0) \propto \frac{1}{\theta^{a+1}} I(k \le \theta)$$

The posterior distribution is given by,

$$\begin{aligned} \pi(\theta \mid y) &\propto f(y \mid \theta) \cdot \pi(\theta) \\ &\propto \frac{1}{\theta^n} I(\max(y_i) \le \theta) \cdot \frac{1}{\theta^{a+1}} I(k \le \theta) \\ &\propto \frac{1}{\theta^{a+n+1}} I(\max(k, \max_i(y_i)) \le \theta) \\ &\propto \text{Pareto}(a+n, \max(k, \max_i(y_i))) \end{aligned}$$

(b) Show that $\hat{\theta} = \max(y_1, \cdots, y_n)$ is the MLE of θ, where y_i are random samples from $f(y_i \mid \theta)$.

The likelihood is given by,

$$l(y \mid \theta) = \prod_{i=1}^{n} f(y_i \mid \theta) = \frac{1}{\theta^n} I(\max_i(y_i) \le \theta)$$

MLE of θ is given by $\hat{\theta}$ which maximize the likelihood $l(y \mid \theta)$ and since $\frac{1}{\theta^n}$ is a decreasing function for $\theta > 0$. $l(y \mid \theta)$ gets maximized at the least value attained by θ i.e. $\max_i(y_i)$. Hence, MLE(θ), $\hat{\theta} = \max(y_1, \cdots, y_n)$.

(c) Find the posterior distribution of $\boldsymbol{\theta}$ and its expected value.

As shown in part a) the posterior distribution $\sim \text{Pareto}(a+n, \max(k, \max_i(y_i)))$

$$\pi(\theta \mid y) = \frac{a+n}{\theta^{a+n+1}} \cdot \max(k, \max_i(y_i))^{a+n} \cdot I(\theta \geq \max(k, \max_i(y_i)))$$

For simplicity, let $k_1 = \max(k, \max_i(y_i))$

The posterior distribution can be written as,

$$\pi(\theta \mid y) = \frac{a+n}{\theta^{a+n+1}} \cdot k_1^{a+n} \cdot I(\theta \geq k_1)$$

Now, to calculate the expected value of the posterior,

$$\begin{aligned}
\mathbb{E}(\theta \mid y) = \mathbb{E}(\pi(\theta \mid y)) &= \int_{k_1}^{\infty} \theta \cdot \frac{a+n}{\theta^{a+n+1}} \cdot k_1^{a+n} \cdot d\theta \\
&= \frac{k_1(a+n)}{(a+n-1)} \int_{k_1}^{\infty} \frac{(a+n-1)}{\theta^{(a+n-1)+1}} \cdot k_1^{(a+n-1)} \cdot d\theta \\
&= \frac{k_1(a+n)}{(a+n-1)} \\
&= \frac{(a+n)}{(a+n-1)} \cdot \max(k, \max_i(y_i))
\end{aligned}$$

Question 3

Consider the following sets of data obtained after tossing a die 100 and 1000 times respectively,

n	**1**	**2**	**3**	**4**	**5**	**6**
100	19	12	17	18	20	14
1000	190	120	170	180	200	140

Suppose you are interested in $\boldsymbol{\theta_1}$, the probability of obtaining a one spot. Assume your prior for all the probabilities is a Dirichlet distribution, where each $\boldsymbol{\alpha_i = 2}$. Compute the posterior distribution for $\boldsymbol{\theta_1}$ for each of the sample sizes in the table. Plot the resulting distribution and compare the results. Comment on the effect of having a larger sample.

The prior distribution is given by,

$$\pi(\theta) = \text{Dirichlet}(\alpha_1 = 2, \ldots, \alpha_6 = 2) = \frac{\Gamma(\sum_{i=1}^{6} 2)}{\prod_{i=1}^{6} \Gamma(2)} \cdot \theta_1 \ldots \theta_6 = \Gamma(12) \cdot \theta_1 \ldots \theta_6$$

The likelihood is given by, $f(y \mid \theta) = \theta_1^{y_1} \ldots \theta_6^{y_6}$ where $\sum_{i=1}^{6} y_i = n$

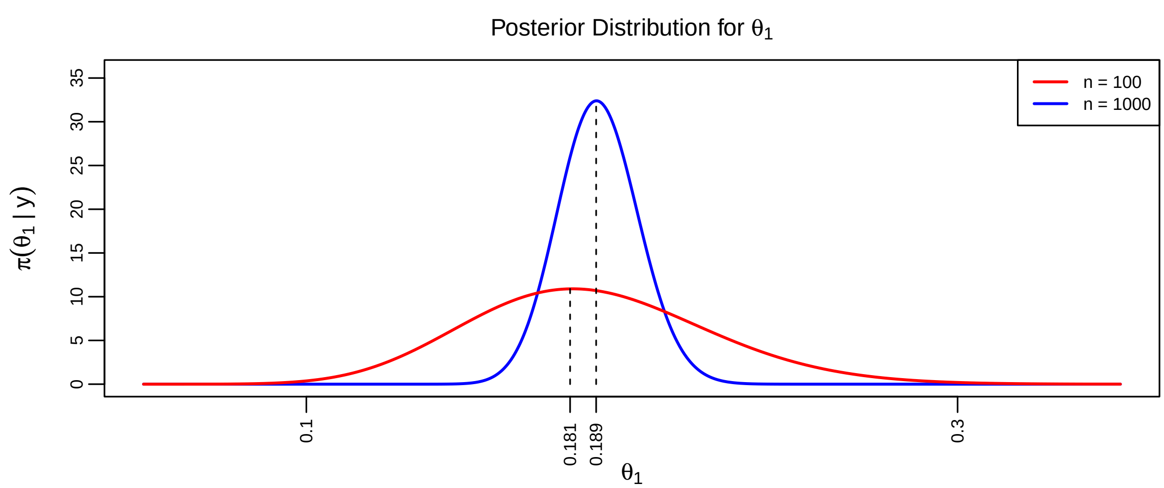
The posterior distribution is given by,

$$\pi(\theta \mid y) = \Gamma(12) \cdot \theta_1 \ldots \theta_6 \cdot \theta_1^{y_1} \ldots \theta_6^{y_6} \propto \theta_1^{y_1+1} \ldots \theta_6^{y_6+1} \propto \text{Dirichlet}(y_1 + 1, \ldots, y_6 + 1)$$

The marginal posterior distribution for θ_1 is given by,

$$\pi(\theta_1 \mid y) \sim \text{Beta}(y_1 + \alpha_1, \sum_{i=2}^{6}(y_i + \alpha_i)) \sim \text{Beta}(y_1 + 2, \sum_{i=2}^{6}(y_i + 2)) \sim \text{Beta}(y_1 + 2, 10 + (y_2 + \ldots + y_6))$$

For $n = 100, \pi(\theta_1 \mid y) \sim \text{Beta}(21, 91)$ and for $n = 1000, \pi(\theta_1 \mid y) \sim \text{Beta}(192, 820)$



As sample size grows, the posterior gets more concenterated around $\hat{\theta} = 0.19$, since the likelihood get more dominant over the prior with increased sample size.

Question 4

(a) Derive Jefferey's prior for θ.

Jeffrey's prior is defined as the square root of the determinant of the information matrix.

Jefferey's Prior, $\mathbb{J}(\theta) = \sqrt{\mathbb{I}(\theta)} \quad where \quad \mathbb{I}(\theta) = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\log f(y_i \mid \theta)\right] \quad$ and $\quad f(y_i \mid \theta) \sim \exp(\theta) = \frac{1}{\theta}\exp\left(-\frac{y_i}{\theta}\right)$

$$\begin{aligned}
\mathbb{I}(\theta) &= -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\log\left(\frac{1}{\theta}\exp\left(-\frac{y_i}{\theta}\right)\right)\right] \\
&= -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\left(-\log(\theta) - \frac{y_i}{\theta}\right)\right] \\
&= \mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\left(\log(\theta) + \frac{y_i}{\theta}\right)\right] \\
&= \mathbb{E}\left[\frac{\partial}{\partial\theta}\left(\frac{1}{\theta} - \frac{y_i}{\theta^2}\right)\right] \\
&= \mathbb{E}\left[\frac{2y_i}{\theta^3} - \frac{1}{\theta^2}\right] \\
&= \frac{2\mathbb{E}[y_i]}{\theta^3} - \frac{1}{\theta^2} \\
&= \frac{2\theta}{\theta^3} - \frac{1}{\theta^2} \\
&= \frac{1}{\theta^2}
\end{aligned}$$

Hence, Jefferey's prior, $\mathbb{J}(\theta) = \sqrt{\mathbb{I}(\theta)} = \frac{1}{\theta}$

(b) Derive Jeffrey's prior for $\alpha = \theta^{-1}$. What do you observe?

Jefferey's prior for α is given by,

$$\mathbb{J}(\alpha) = \mathbb{J}(\theta)\left|\frac{d\theta}{d\alpha}\right| = \frac{1}{\theta}\cdot\left|\frac{d}{d\alpha}(\alpha^{-1})\right| = \alpha\cdot\left|-\alpha^{-2}\right| = \alpha^{-1}$$

The Jefferey's prior is invariant under reparameterization i.e. $\mathbb{J}(\theta) \propto \theta^{-1} \Leftrightarrow \mathbb{J}(\alpha) \propto \alpha^{-1}$.

(c) Find the posterior density of θ corresponding to the prior density in (a). Be specific in noting the family to which it belongs.

The prior distribution for θ is given by $\pi(\theta) = \theta^{-1}$

The likelihood is given by,

$$f(y \mid \theta) = \prod_{i=1}^{n} f(y_i \mid \theta) = \theta^{-n}\exp\left[-\theta^{-1}\sum_{i=1}^{n} y_i\right]$$

The posterior distribution is given by,

$$\pi(\theta \mid y) \propto \pi(\theta)\cdot f(y \mid \theta) \propto \theta^{-1}\cdot\theta^{-n}\exp\left[-\theta^{-1}\sum_{i=1}^{n} y_i\right] \propto \theta^{-(n+1)}\exp\left[-\theta^{-1}\sum_{i=1}^{n} y_i\right] \propto IG\left(n, \sum_{i=1}^{n} y_i\right)$$

Question 5

(a) Derive the conditional posterior distribution of β and show that $\pi(\beta \mid \sigma^2, y) \sim N(\bar{\beta}, B_1)$, where $B_1 = [\sigma^{-2}X'X + B_0^{-1}]^{-1}$ and $\bar{\beta} = B_1[\sigma^{-2}X'y + B_0^{-1}\beta_0]$

The likelihood is given by,

$$f(y \mid \beta, \sigma^2) \propto \exp\left[-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)\right]\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}}$$

The prior distribution of β is given by,

$$\pi(\beta) \propto \exp\left[-\frac{1}{2}(\beta - \beta_0)'B_0^{-1}(\beta - \beta_0)\right]$$

The conditional posterior distribution of β is given by,

$$\begin{aligned}
\pi(\beta \mid \sigma^2, y) &\propto f(y \mid \beta, \sigma^2) \cdot \pi(\beta) \\
&\propto \exp\left[-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)\right]\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} \cdot \exp\left[-\frac{1}{2}(\beta - \beta_0)'B_0^{-1}(\beta - \beta_0)\right] \\
&\propto \exp\left[-\frac{1}{2}\Big\{\sigma^{-2}(y - X\beta)'(y - X\beta) + (\beta - \beta_0)'B_0^{-1}(\beta - \beta_0)\Big\}\right]
\end{aligned}$$

Expanding the terms inside the curly bracket,

$$\begin{aligned}
(y - X\beta)'(y - X\beta) &= y'y - 2\beta'X'y + \beta'X'X\beta \\
(\beta - \beta_0)'B_0^{-1}(\beta - \beta_0) &= \beta'B_0^{-1}\beta - 2\beta'B_0^{-1}\beta_0 + \beta_0'B_0^{-1}\beta_0
\end{aligned}$$

Substituting back and grouping together,

$$\begin{aligned}
\pi(\beta \mid \sigma^2, y) &\propto \exp\left[-\frac{1}{2}\Big\{\beta'(\sigma^{-2}X'X + B_0^{-1})\beta - 2\beta'(\sigma^{-2}X'y + B_0^{-1}\beta_0) + \beta_0'B_0^{-1}\beta_0 + \sigma^{-2}y'y\Big\}\right] \\
&\propto \exp\left[-\frac{1}{2}\Big\{\beta'(\sigma^{-2}X'X + B_0^{-1})\beta - 2\beta'(\sigma^{-2}X'y + B_0^{-1}\beta_0)\Big\}\right]
\end{aligned}$$

Let $A = \sigma^{-2}X'X + B_0^{-1}$, $b = \sigma^{-2}X'y + B_0^{-1}\beta_0$ and $\bar{\beta} = A^{-1}b$

$$\begin{aligned}
\pi(\beta \mid \sigma^2, y) &\propto \exp\left[-\frac{1}{2}\Big\{\beta'A\beta - 2\beta'b\Big\}\right] \\
&\propto \exp\left[-\frac{1}{2}\Big\{\beta'A\beta - 2\beta'b + \bar{\beta}'A\bar{\beta} - \bar{\beta}'A\bar{\beta}\Big\}\right] \\
&\propto \exp\left[-\frac{1}{2}\Big\{\big(\beta'A\beta - 2\beta'(A\bar{\beta}) + \bar{\beta}'A\bar{\beta}\big) - \bar{\beta}'A\bar{\beta}\Big\}\right] \\
&\propto \exp\left[-\frac{1}{2}\Big\{\big(\beta - \bar{\beta}\big)' A \big(\beta - \bar{\beta}\big) - \bar{\beta}'A\bar{\beta}\Big\}\right] \\
&\propto \exp\left[-\frac{1}{2}\Big\{\big(\beta - \bar{\beta}\big)' A \big(\beta - \bar{\beta}\big)\Big\}\right] \\
&\propto N(\bar{\beta}, B_1) \qquad \text{where } B_1 = A^{-1} = [\sigma^{-2}X'X + B_0^{-1}]^{-1} \text{ and } \bar{\beta} = B_1[\sigma^{-2}X'y + B_0^{-1}\beta_0]
\end{aligned}$$

(b) Derive the conditional posterior distribution of $\boldsymbol{\sigma^2}$ and show that $\boldsymbol{\pi(\sigma^2 \mid, \beta, y) \sim IG(\frac{\alpha_1}{2}, \frac{\delta_1}{2})}$, where $\boldsymbol{\alpha_1 = \alpha_0 + n}$ and $\boldsymbol{\delta_1 = \delta_0 + (y - X\beta)'(y - X\beta)}$

The likelihood is given by,

$$f(y \mid \beta, \sigma^2) \propto \exp\left[-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)\right]\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}}$$

The prior distribution of σ^2 is given by,

$$\pi(\sigma^2) \propto \exp\left[-\frac{\delta_0}{2\sigma^2}\right]\left(\frac{1}{\sigma^2}\right)^{\left(\frac{\alpha_0}{2}+1\right)}$$

The conditional posterior distribution of σ^2 is given by,

$$\begin{aligned}
\pi(\beta \mid \sigma^2, y) &\propto f(y \mid \beta, \sigma^2) \cdot \pi(\sigma^2) \\
&\propto \exp\left[-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)\right]\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} \cdot \exp\left[-\frac{\delta_0}{2\sigma^2}\right]\left(\frac{1}{\sigma^2}\right)^{\left(\frac{\alpha_0}{2}+1\right)} \\
&\propto \exp\left[-\frac{1}{2\sigma^2}\Big(\delta_0 + (y - X\beta)'(y - X\beta)\Big)\right] \cdot \left(\frac{1}{\sigma^2}\right)^{\left(\frac{\alpha_0+n}{2}+1\right)} \\
&\propto \exp\left[-\frac{\delta_1}{2\sigma^2}\right] \cdot \left(\frac{1}{\sigma^2}\right)^{\left(\frac{\alpha_1}{2}+1\right)} \\
&\propto IG\left(\frac{\alpha_1}{2}, \frac{\delta_1}{2}\right) \qquad \text{where } \alpha_1 = \alpha_0 + n \text{ and } \delta_1 = \delta_0 + (y - X\beta)'(y - X\beta)
\end{aligned}$$