In the name of GOD.

Stochastic Process

Fall 2024

Hamid R. Rabiee

Sharif University of Technology

Homework 4 | Estimation Theory | Deadline : 1403/09/17

1. [5] The time interval between the arrival of trains at Sharif Metro Station in the morning follows an exponential distribution with parameter λ. If $x_1, \dots, x_N$ are independent samples of the time between the arrival of two trains at this station in the morning, Find a *minimal sufficient statistic* for λ.

   Solution:
   The PDF of *exponential* distribution is:

$$f(x;\lambda) = \lambda e^{-\lambda x} \ if \ x \geq 0 \ else \ 0$$

   It can be shown that $T(X) = \sum_{i=1}^{N} x_i$ is an *minimal sufficient statistic* for λ. For this purpose, The following condition must hold for any two sets of samples $X = (x_1, \dots, x_N), Y = (y_1, \dots, y_N)$.

$$\frac{f(X;\lambda)}{f(Y;\lambda)} \ is \ independent \ of \ \lambda \quad \Leftrightarrow \quad T(X) = T(Y)$$

   In our case:

$$\frac{f(X;\lambda)}{f(Y;\lambda)} = \frac{\lambda^N \exp\left(-\lambda \sum_i x_i\right)}{\lambda^N \exp\left(-\lambda \sum_i x_i\right)} = \exp\left(-\lambda\left(\sum_i x_i - \sum_i y_i\right)\right).$$

$$\frac{f(X;\lambda)}{f(Y;\lambda)} \ is \ independent \ of \ \lambda \Rightarrow \sum_i x_i - \sum_i y_i = 0 \Rightarrow \sum_i x_i = \sum_i y_i \Rightarrow T(X) = T(Y)$$

$$T(X) = T(Y) \Rightarrow \sum_i x_i = \sum_i y_i \Rightarrow \frac{f(X;\lambda)}{f(Y;\lambda)} = 1$$

   So $T(X) = \sum_{i=1}^{N} x_i$ is a *minimal sufficient statistic.*

2. [10] Let $X_1, \dots, X_N$ be iid with PDF:

$$f(x;\theta) = \begin{cases} \frac{x}{\theta} exp\left(-\frac{x^2}{2\theta}\right), & x > 0 \\ 0, & x \leq 0 \end{cases}$$

   (a) Find a scalar (one-dimensional) *sufficient statistic* for θ using *factorization theorem.*
   (b) Find the *method of moments* estimator of θ.

   Solution:

(a) The joint PDF of $X_1, X_2, \ldots, X_N$ is given as:

$$f(X;\theta) = \begin{cases} \prod_{i=1}^{N} \frac{X_i}{\theta} \exp\left(-\frac{X_i^2}{2\theta}\right), & X_i > 0 \quad \forall i = 1, \ldots, n \\ 0, & \text{otherwise} \end{cases}$$

This simplifies to:

$$f(X;\theta) = \begin{cases} \frac{\prod_i X_i}{\theta^N} \exp\left(-\frac{\sum_i X_i^2}{2\theta}\right), & \min(X_1, X_2, \ldots, X_N) > 0 \\ 0, & \text{otherwise} \end{cases}$$

And:

$$f(X;\theta) = (\prod_i X_i) \left(\frac{1}{\theta^N}\right) \exp\left(-\frac{\sum_i X_i^2}{2\theta}\right) I\left(\min(X_1, X_2, \ldots, X_N) > 0\right)$$

So If $T = \sum_i X_i^2$, $g(T|\theta) = \left(\frac{1}{\theta^N}\right) \exp\left(-\frac{T}{2\theta}\right)$, and $h(X) = (\prod_i X_i) I\left(\min(X_1, X_2, \ldots, X_N) > 0\right)$, we can write PDF as:

$$f(X;\theta) = g(T|\theta)h(X)$$

By the *factorization theorem*, $T = \sum_i X_i^2$ is a **sufficient statistic** for θ.

(b) Expected value is equal to:

$$E_f[x] = \int_{-\infty}^{\infty} x f(x;\theta) dx = \int_0^{\infty} \frac{x^2}{\theta} exp\left(-\frac{x^2}{2\theta}\right) dx$$

$$= -x \cdot exp\left(-\frac{x^2}{2\theta}\right)]_0^{\infty} + \int_0^{\infty} exp\left(-\frac{x^2}{2\theta}\right)$$

Considering the symmetry of $N(0,\theta)$ and its CDF:

$$\int_0^{\infty} \frac{1}{\sqrt{2\pi\theta}} exp\left(-\frac{x^2}{2\theta}\right) = \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\theta}} exp\left(-\frac{x^2}{2\theta}\right) = \frac{1}{2}$$

Hence:

$$E_f[x] = 0 + \frac{1}{2}\sqrt{2\pi\theta} = \sqrt{\frac{\pi\theta}{2}}$$

$$\frac{\sum_{i=1}^{N} X_i}{N} = \sqrt{\frac{\pi\theta}{2}} \Rightarrow \hat{\theta} = \sqrt{\frac{2}{\pi}} \frac{\sum_{i=1}^{N} X_i}{N}$$

3. [20] Suppose that the random variables $Y_1, \ldots, Y_n$ satisfy

$$Y_i = \beta x_i + e_i, \quad i = 1, \ldots, n,$$

where $x_1, \ldots, x_n$ are fixed known constants and $e_1, \ldots, e_n$ are iid samples from $N(0, \sigma^2)$.

(a) Find the MLE of β, and show that it is an unbiased estimator of β.

(b) Calculate the mean and variance of $S = \frac{\sum Y_i}{\sum x_i}$ as an estimator for β, and then compare it to the MLE of β.

Solution:

(a) We know Y_i are iid samples from $N(\beta x_i, \sigma^2)$. So likelihood function is given by:

$$L(\beta, \sigma^2|Y) = \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) \exp\left(-\frac{1}{2\sigma^2}(Y_i - \beta x_i)^2\right)$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i^2 - 2\beta x_i Y_i + \beta^2 x_i^2)\right)$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} Y_i^2 - 2\beta\sum_{i=1}^{n} x_i Y_i + \beta^2 \sum_{i=1}^{n} x_i^2\right)\right)$$

The log-likelihood function is:

$$\log L(\beta, \sigma^2|Y) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \beta x_i)^2$$

Taking the derivative with respect to β and setting it to zero:

$$\frac{\partial}{\partial\beta}\log L = \sum_{i=1}^{n} x_i(Y_i - \beta x_i) = 0$$

$$\Rightarrow \hat{\beta} = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}$$

The second derivative is:

$$\frac{\partial^2}{\partial\beta^2}\log L = -\sum_{i=1}^{n} x_i^2 < 0$$

Thus, $\hat{\beta}$ is the MLE and it is unbiased because:

$$E[\hat{\beta}] = \frac{\sum_{i=1}^{n} x_i E[Y_i]}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} x_i \beta x_i}{\sum_{i=1}^{n} x_i^2} = \beta$$

(b) $Var(\hat{\beta}) = Var\left(\sum_i a_i Y_i\right)$ where $a_i = \frac{x_i}{\sum_j x_j^2}$ are constants. So variance of $\hat{\beta}$ is:

$$Var(\hat{\beta}) = \sum_i a_i^2 Var(Y_i) = \sum_i \left(\frac{x_i}{\sum_j x_j^2}\right)^2 \sigma^2 = \frac{\sum_i x_i^2}{(\sum_j x_j^2)^2}\sigma^2 = \frac{\sigma^2}{\sum_i x_i^2}$$

For S:

$$E[S] = E\left[\frac{\sum Y_i}{\sum x_i}\right] = \frac{1}{\sum x_i}\sum_i E[Y_i] = \frac{1}{\sum x_i}\sum_i \beta x_i = \beta$$

Therefor S like $\hat{\beta}$ is unbiased.

$$Var(S) = Var\left(\frac{\sum Y_i}{\sum x_i}\right) = \frac{1}{\left(\sum x_i\right)^2}\sum_i Var(Y_i) = \frac{\sum_i \sigma^2}{\left(\sum x_i\right)^2} = \frac{n\sigma^2}{n^2\overline{x}^2} = \frac{\sigma^2}{n\overline{x}^2}$$

Note that $\overline{x} = \frac{\sum_i x_i}{n}$. To compare with $\hat{\beta}$ variance, we use following inequality:

$$\sum_i (x_i - \overline{x})^2 = \sum_i x_i^2 - 2\overline{x}\sum_i x_i + \sum_i \overline{x}^2 = \sum_i x_i^2 - 2\overline{x}\,(n\overline{x}) + n\overline{x}^2 = \sum_i x_i^2 - n\overline{x}^2$$

$$\sum_i (x_i - \overline{x})^2 \geq 0 \Rightarrow \sum_i x_i^2 \geq n\overline{x}^2$$

Hence,

$$Var\left(\hat{\beta}\right) = \frac{\sigma^2}{\sum_i x_i^2} \leq \frac{\sigma^2}{n\overline{x}^2} = Var\left(S\right)$$

4. [10] Let $x_1, \ldots, x_N$ are iid samples from a distribution with PDF as follows:

$$f(x;\theta) = 2e^{-|x-\theta|}$$

Find the MLE of θ.

Solution:
The likelihood function is:

$$L(\theta|X) = \prod_{i=1}^{N} 2e^{-|x_i-\theta|} = 2^N e^{-\sum_{i=1}^{N}|x_i-\theta|}.$$

To maximize this function, we should minimize $\sum_{i=1}^{N}|x_i - \theta|$.

$$\hat{\theta} = arg\max_{\theta} L(\theta|X) = arg\min_{\theta} \sum_{i=1}^{N} |x_i - \theta|.$$

We claim that the minimum of the $\sum_{i=1}^{N}|x_i - \theta|$ occurs in the median of $x_1, \ldots, x_N$. Without loss of generality, assume that all the $x_i s$ are sorted in ascending order. Also define $g(t) = \sum_{i=1}^{N}|x_i - t|$.

- Lemma 1: $\hat{\theta}$ lies on one of the $x_i s$. (If n is even, there is more than one optimal solution.)

- Proof: Prove it by the converse. Suppose that $\hat{\theta}$ is between two points x_j and x_{j+1}. If $j \geq (N-j)$, then by moving $\hat{\theta}$ to x_j the value of $g(\hat{\theta})$ decreases by the value $(j-(N-j))(\theta - x_j)$. Otherwise, by moving $\hat{\theta}$ to x_{j+1}, the value of $g(\hat{\theta})$ decreases by the value $((N-j)-j)(x_{j+1}-\theta)$. So we only consider $x_i s$ to find optimal solution.

So we only have to consider $g(x_i) for\ i = x_1, \ldots, x_N$. We have:

$$g(x_k) = \sum_{i=1}^{k-1} x_k - x_i + \sum_{i=k+1}^{N} x_i - x_k = ((k-1)-(N-k))x_k + \sum_{i=1}^{k-1} -x_i + \sum_{i=k+1}^{N} x_i$$
$$= (2k-N-1)x_k + \sum_{i=1}^{k-1} -x_i + \sum_{i=k+1}^{N} x_i$$

Consider the difference $d_k = g(x_k) - g(x_{k+1})$. We have:

$$d_k = (2k-N-1)x_k + \sum_{i=1}^{k-1} -x_i + \sum_{i=k+1}^{N} x_i - \left((2(k+1)-N-1)x_{k+1} + \sum_{i=1}^{k} -x_i + \sum_{i=k+2}^{N} x_i\right)$$

$$= x_k + x_{k+1} + (2k-N-1)x_k - (2k-N+1)x_{k+1} = (2k-N)x_k - (2k-N)x_{k+1} = (2k-N)(x_k - x_{k+1})$$

- Lemma 2: $g(t)$ is convex.

- Proof: Due to the continuity of function $g(t)$, it is only sufficient to check the following condition:

$$g\left(\frac{t_1+t_2}{2}\right) \leq \frac{g(t_1)+g(t_2)}{2}$$

By triangle inequality:

$$|x - \frac{t_1+t_2}{2}| = |\frac{x-t_1}{2} + \frac{x-t_2}{2}| \leq |\frac{x-t_1}{2}| + |\frac{x-t_2}{2}|$$

Hence:

$$g\left(\frac{t_1+t_2}{2}\right) = \sum_i |x_i - \frac{t_1+t_2}{2}| \leq \sum_i |\frac{x_i-t_1}{2}| + |\frac{x_i-t_2}{2}| = \frac{g(t_1)+g(t_2)}{2}$$

Since $g(t)$ is convex and the $x_i s$ are ordered, the minimum will be attained at x_k, where k is the smallest integer, such that $d_k \leq 0$. This is equivalent to $(2k-N)(x_k - x_{k+1}) \leq 0$. As the x_k are ordered ($x_k - x_{k+1} < 0$), this is only possible if $2k - N \geq 0$. Therefore the minimum is attained at x_k for k being the smallest integer k, such that $2k \geq N$, this is $\lceil k = \frac{N}{2} \rceil$, the median of $x_i s$. Note that if N is even, all the numbers in interval $[x_{\frac{N}{2}}, x_{\frac{N}{2}+1}]$ are optimal.

5. [15] Let $X_1, \ldots, X_N$ be iid with PDF:

$$f(x;\theta) = \begin{cases} e^{(\theta - x)}, & x \geq \theta \\ 0, & x < \theta \end{cases}$$

   (a) Find a *complete sufficient statistic* for θ.
   (b) Use this sufficient statistic and calculate $UMVUE$ for θ.

Solution:

(a) The joint PDF of $X_1, X_2, \ldots, X_N$ is given as:

$$f(X;\theta) = \begin{cases} \prod_{i=1}^{N} \exp\left(\theta - X_i\right), & X_i \geq \theta \quad \forall i = 1, \ldots, n \\ 0, & \text{otherwise} \end{cases}$$

This simplifies to:

$$f(X;\theta) = \begin{cases} \exp\left(N\theta\right) \cdot \exp\left(-\sum_{i=1}^{N} X_i\right), & \min(X_1, X_2, \ldots, X_N) \geq \theta \\ 0, & \text{otherwise} \end{cases}$$

And:

$$f(X;\theta) = \exp\left(N\theta\right) \exp\left(-\sum_{i=1}^{N} X_i\right) I\left(\min(X_1, X_2, \ldots, X_N) \geq \theta\right)$$

So If $T = \min(X_1, X_2, \ldots, X_N),\ g(T|\theta) = \exp\left(N\theta\right) I\left(T \geq \theta\right), and\ h(X) = \exp\left(-\sum_{i=1}^{N} X_i\right)$, we can write PDF as:

$$f(X;\theta) = g(T|\theta)h(X)$$

By the *factorization theorem*, $T = \min(x_1, x_2, \ldots, x_N)$ is a **sufficient statistic** for θ.

To prove completeness, we need PDF of T. The CDF of T is:

$$F_T(t) = P(T \leq t) = 1 - P(T > t) = 1 - \prod_{i=1}^{N} P(X_i > t) = 1 - \prod_{i=1}^{N} e^{\theta - t} = 1 - e^{N(\theta - t)}, \quad t \geq \theta.$$

For the above calculations, we used the CDF of $f(x;\theta)$, which is given below.

$$P(x > t) = 1 - \int_{-\infty}^{t} f(u;\theta)du = 1 - \int_{\theta}^{t} \exp(\theta - u)du = 1 - \left(-e^{\theta - t} + e^{\theta - \theta}\right) = e^{\theta - t}.$$

Differentiating the CDF gives the PDF of T:

$$f_T(t;\theta) = Ne^{N(\theta - t)}, \quad t \geq \theta$$

Now assume there exists a function $g(t)$ such that:

$$E_\theta[g(\min(X_1, X_2, \ldots, X_N))] = 0 \quad \textit{for all } \theta.$$

The expectation can be written as:

$$E_\theta[g(T)] = \int_{\theta}^{\infty} g(t)Ne^{N(\theta - t)}\, dt = 0 \quad \forall\theta.$$

Since the result of the integral with respect to θ is constant, its derivative with respect to θ must be zero.

$$\Rightarrow \; -g(\theta)Ne^{N(\theta - \theta)} = 0 \quad \forall\theta \quad \Rightarrow \; g(\theta) = 0 \quad \forall\theta$$

This proves that $T = \min(x_1, x_2, \ldots, x_N)$ is *complete*.

(b) The expected value of $T = \min(X_1, X_2, \ldots, X_N)$ is:

$$E[T] = \int_{\theta}^{\infty} t \cdot Ne^{N(\theta - t)}\, dt.$$

Substitute $u = t - \theta$, so $t = u + \theta$ and $dt = du$. The limits of integration become $u \in [0, \infty)$, and:

$$E[T] = \int_{0}^{\infty} (\theta + u) \cdot Ne^{-Nu}\, du.$$

Split the integral:

$$E[T] = \theta \int_{0}^{\infty} Ne^{-Nu}\, du + \int_{0}^{\infty} u \cdot Ne^{-Nu}\, du.$$

The first term evaluates to:

$$\int_{0}^{\infty} Ne^{-Nu}\, du = -e^{-N(\infty)} + e^{-N(0)} = 1.$$

The second term evaluates to:

$$\int_{0}^{\infty} u \cdot Ne^{-Nu}\, du = -u \cdot e^{-Nu}]_0^{\infty} + \int_{0}^{\infty} e^{-Nu}\, du = 0 + \frac{-e^{-N(\infty)} + e^{-N(0)}}{N} = \frac{1}{N}.$$

Thus:

$$E[T] = \theta + \frac{1}{N}.$$

To construct an unbiased estimator of θ, subtract $\frac{1}{N}$ from T:

$$\hat{\theta} = T - \frac{1}{N}.$$

So $T = \min(X_1, X_2, \ldots, X_N)$ is a sufficient and complete statistic for θ and $\phi(T) = T - \frac{1}{N}$ is an unbiased estimator of θ. By the *Lehmann-Scheffé theorem*, $\min(X_1, X_2, \ldots, X_N) - \frac{1}{N}$ is the **UMVUE**.

6. [20] Let $x_1, \ldots, x_N$ are iid samples from the following distribution:

$$f(x; \alpha, \beta) = \frac{1}{\beta} e^{-\frac{x-\alpha}{\beta}}, \quad x \geq \alpha \ , \quad \beta > 0$$

Find the $UMVUE$ for α *and* β.

(Hint: Show MLE of these parameters are $complete\ sufficient\ statistics$.)

Solution:

The likelihood function is:

$$f(X; \alpha, \beta) = \prod_{i=1}^{N} f(x_i; \alpha, \beta) = \prod_{i=1}^{N} \frac{1}{\beta} e^{-\frac{x_i-\alpha}{\beta}} I(x_i \geq \alpha) = \frac{1}{\beta^N} \exp\left(-\frac{\sum_i (x_i - \alpha)}{\beta}\right) I(\forall_i : x_i \geq \alpha)$$

Define $c = min(x_1, \ldots, x_N)$.

$$f(X; \alpha, \beta) = \frac{1}{\beta^N} \exp\left(-\frac{\sum_i x_i}{\beta} + \frac{N\alpha}{\beta}\right) I(c \geq \alpha)$$

To find the MLE, first focus on α.

$$\hat{\alpha} = arg \max_{\alpha} f(X; \alpha, \beta) = arg \max_{\alpha} \exp\left(\frac{N\alpha}{\beta}\right) I(c \geq \alpha)$$

Since $\exp\left(\frac{N\alpha}{\beta}\right)$, is an increasing function of α, $\hat{\alpha}$ should be the largest number which $I(c \geq \alpha) = 1$. So $\hat{\alpha} = c = min(x_1, \ldots, x_N)$.

For β, Use the derivative of the log-likelihood.

$$\hat{\beta} = arg \max_{\beta} f(X; \alpha, \beta) = arg \max_{\beta} log(f(X; \alpha, \beta))$$

$$L = \log(f(X; \alpha, \beta)) = -N \log(\beta) - \frac{\sum_i x_i}{\beta} + \frac{Nc}{\beta}$$

$$\frac{\partial L}{\partial \beta} = -\frac{N}{\beta} + \frac{\sum_i x_i}{\beta^2} - \frac{Nc}{\beta^2} = 0 \Rightarrow \hat{\beta} = \frac{\sum_i x_i}{N} - c$$

$$\frac{\partial^2 L}{\partial \beta^2}|_{\hat{\beta}} = \frac{N}{\hat{\beta}^2} - 2\frac{\sum_i x_i}{\hat{\beta}^3} + 2\frac{Nc}{\hat{\beta}^3} = \frac{N}{\hat{\beta}^2} - 2\frac{\sum_i x_i - Nc}{\hat{\beta}^3} = \frac{N}{\hat{\beta}^2} - 2\frac{N\hat{\beta}}{\hat{\beta}^3} = -\frac{N}{\hat{\beta}^2} < 0$$

So the MLE is $(\hat{\alpha}, \hat{\beta}) = (c, \frac{\sum_i x_i}{N} - c)$.

The second step is to prove the *completeness* of these statistics.

First for α (Suppose β is fixed):

$$f(X;\alpha,\beta) = \frac{1}{\beta^N}\exp\left(-\frac{\sum_i x_i}{\beta} + \frac{N\alpha}{\beta}\right) I(c \geq \alpha) = \frac{1}{\beta^N}\exp\left(-\frac{\sum_i x_i}{\beta}\right)\exp\left(\frac{N\alpha}{\beta}\right) I(c \geq \alpha)$$

$$h(X) = \frac{1}{\beta^N}\exp\left(-\frac{\sum_i x_i}{\beta}\right), \quad g(T(X) = c|\alpha) = \exp\left(\frac{N\alpha}{\beta}\right) I(c \geq \alpha)$$

Thus, by the *factorization theorem*, $T(X) = c = min(x_1, \ldots, x_N)$ is a *sufficient statistic* for α.

Next, we need PDF of c.

$$F_X(x) = \int_{-\infty}^{x} f(t;\alpha,\beta)dt = \int_{\alpha}^{x} \frac{1}{\beta} e^{-\frac{t-\alpha}{\beta}} dt = -e^{-\frac{t-\alpha}{\beta}}]_{\alpha}^{x} = 1 - e^{-\frac{x-\alpha}{\beta}} \quad x \geq \alpha$$

$$\Rightarrow P_X(x \leq t) = 1 - e^{-\frac{t-\alpha}{\beta}} \ \Rightarrow P_X(x > t) = e^{-\frac{t-\alpha}{\beta}}$$

$$\Rightarrow P_C(c > x) = P_X(\forall_{i=1}^N : x_i > x) = \prod_{i=1}^{N} P_X(x_i > x) = \exp(-N\frac{x-\alpha}{\beta})$$

$$CDF : F_C(x) = 1 - P_C(c > x) = 1 - \exp(-N\frac{x-\alpha}{\beta})$$

$$\Rightarrow PDF : f_C(x;\alpha,\beta) = \frac{N}{\beta}\exp(-N\frac{x-\alpha}{\beta}) \quad if\ x \geq \alpha \quad o.w\ 0$$

Suppose $u(t)$ is a function such that $\forall_\alpha : E_\alpha[u(c)] = 0$.

$$\forall_\alpha : \int_{-\infty}^{\infty} u(t) f_C(t;\alpha,\beta)dt = 0 \Rightarrow \int_{\alpha}^{\infty} u(t)\frac{N}{\beta}\exp(-N\frac{t-\alpha}{\beta})dt = 0$$

$$\Rightarrow \ \forall_\alpha : \int_{\alpha}^{\infty} u(t)\exp(-N\frac{t}{\beta})dt = 0$$

Since the result of the integral with respect to α is constant, its derivative with respect to α must be *zero*.

$$\Rightarrow \ \forall_\alpha : -u(\alpha)\exp(\frac{-N\alpha}{\beta}) = 0 \ \Rightarrow \ u(\alpha) = 0$$

So, $c = min(x_1, \ldots, x_N)$ is a *complete sufficient statistic* for α.

For β, use *exponential family*. (Suppose α is fixed)

$$f(x_i;\alpha,\beta) = \frac{1}{\beta} e^{-\frac{x_i-\alpha}{\beta}} I(x_i \geq \alpha) = I(x_i \geq \alpha)\left(\frac{1}{\beta} e^{\frac{\alpha}{\beta}}\right)\left(e^{\frac{x_i}{-\beta}}\right)$$

$$h(x_i) = I(x_i \geq \alpha), \quad c(\beta) = \frac{1}{\beta} e^{\frac{\alpha}{\beta}}, \quad w_1(\beta) = \frac{-1}{\beta}, \quad t_1(x_i) = x_i$$

By the *theorem*, $\sum_{i=1}^N x_i$ is a *complete sufficient statistic* for β.

Now, these statistics just need to be unbiased.

$$E[c] = \int_{-\infty}^{\infty} t \cdot f_C(t;\alpha,\beta)dt = \int_{\alpha}^{\infty} t\frac{N}{\beta}\exp(-N\frac{t-\alpha}{\beta})dt$$

$$= -t \cdot \exp(-N\frac{t-\alpha}{\beta})]_\alpha^\infty + \int_\alpha^\infty \exp(-N\frac{t-\alpha}{\beta})dt = \alpha + \frac{-\beta}{N}\exp(-N\frac{t-\alpha}{\beta})]_\alpha^\infty = \alpha + \frac{\beta}{N}$$

So, the $UMVUE$ for α is $\hat{\alpha} = c - \frac{\beta}{N}$.

$$E[x] = \int_{-\infty}^\infty t \cdot f_X(t;\alpha,\beta)dt = \int_\alpha^\infty t\frac{1}{\beta}\exp(-\frac{t-\alpha}{\beta})dt$$

$$= -t \cdot \exp(-\frac{t-\alpha}{\beta})]_\alpha^\infty + \int_\alpha^\infty \exp(-\frac{t-\alpha}{\beta})dt = \alpha + -\beta\exp(-\frac{t-\alpha}{\beta})]_\alpha^\infty = \alpha + \beta$$

$$\Rightarrow \; E[\sum_i x_i] = N \cdot E[x] = N\alpha + N\beta$$

So, the $UMVUE$ for β is $\hat{\beta} = \frac{\sum_i x_i}{N} - \alpha$.

Solving this system of linear equations leads to this conclusion:

$$\hat{\alpha} = \frac{N}{N-1}\min(x_1,\ldots,x_N) - \frac{\sum_i x_i}{N(N-1)}, \qquad \hat{\beta} = \frac{\sum_i x_i}{N-1} - \frac{N}{N-1}\min(x_1,\ldots,x_N)$$

7. [15] Let $x_1,\ldots,x_N$ are iid samples from $U(0,\theta)$, which prior distribution of θ is:

$$\pi(\theta|\alpha,\beta) = \frac{\alpha\beta^\alpha}{\theta^{\alpha+1}} \quad \theta \geq \alpha, \beta > 0$$

   (a) Find the MAP estimator for θ.
   (b) Find $Bayes\ Minimum\ Loss$ estimator for θ. Use $Squared\ Error\ Loss$.

Solution:

   (a) PDF of $Uniform$ distribution is:

$$f(x|\theta) = \frac{1}{\theta} \quad if\ \theta \geq x \geq 0 \quad o.w\ 0$$

So Posterior distribution is:

$$P(\theta|X) \propto f(X|\theta)\pi(\theta|\alpha,\beta) = \frac{\alpha\beta^\alpha}{\theta^{\alpha+1}}\prod_{i=1}^N f(x_i|\theta) \quad if\ \theta \geq \alpha,\beta \quad o.w.\ 0$$

$$= \frac{\alpha\beta^\alpha}{\theta^{\alpha+1}}\frac{1}{\theta^N} \quad if\ \theta \geq \alpha,\beta\ \ and\ \ \forall_i : x_i \leq \theta \quad o.w.\ 0$$

To maximize posterior, $\theta \geq \alpha, \beta$ and $max(x_1,\ldots,x_N) \leq \theta$ must be met. In this situation:

$$\hat{\theta} = arg\max_\theta P(\theta|X) = arg\max_\theta \frac{\alpha\beta^\alpha}{\theta^{\alpha+1}}\frac{1}{\theta^N} = arg\max_\theta \frac{\alpha\beta^\alpha}{\theta^{N+\alpha+1}}$$

$$= arg\min_\theta \theta^{N+\alpha+1}$$

Thus, $\hat{\theta}$ is the smallest number that satisfies the conditions.

$$\hat{\theta} = max(\alpha,\beta,x_1,\ldots,x_N)$$

(b)

$$\hat{\theta} = E[\theta|X] = \int_{-\infty}^{\infty} \theta P(\theta|X)d\theta = \int_{-\infty}^{\infty} \theta \frac{P(X|\theta)\pi(\theta)}{P(X)}d\theta = \frac{1}{P(X)}\int_{-\infty}^{\infty} \theta P(X|\theta)\pi(\theta)d\theta$$

First we calculate integral. Define $c = max(\alpha, \beta, x_1, \ldots, x_N)$. From the previous section, we know $P(X|\theta)\pi(\theta)$ is equal to *zero* for $\theta < c$.

$$\int_{-\infty}^{\infty} \theta P(X|\theta)\pi(\theta)d\theta = \int_{c}^{\infty} \theta \frac{\alpha\beta^\alpha}{\theta^{N+\alpha+1}}d\theta = \alpha\beta^\alpha \int_{c}^{\infty} \frac{1}{\theta^{N+\alpha}}d\theta$$

$$= \alpha\beta^\alpha \left(\frac{-\theta^{-(N+\alpha-1)}}{N+\alpha-1}]_c^\infty \right) = \alpha\beta^\alpha \frac{c^{-(N+\alpha-1)}}{N+\alpha-1}$$

To calculate $P(X)$:

$$P(X) = \int_{-\infty}^{\infty} P(X|\theta)\pi(\theta)d\theta = \int_{c}^{\infty} \frac{\alpha\beta^\alpha}{\theta^{N+\alpha+1}}d\theta = \alpha\beta^\alpha \int_{c}^{\infty} \frac{1}{\theta^{N+\alpha+1}}d\theta$$

$$= \alpha\beta^\alpha \left(\frac{-\theta^{-(N+\alpha)}}{N+\alpha}]_c^\infty \right) = \alpha\beta^\alpha \frac{c^{-(N+\alpha)}}{N+\alpha}$$

And finally:

$$\hat{\theta} = \frac{N+\alpha}{N+\alpha-1} max(\alpha, \beta, x_1, \ldots, x_N)$$

8. [5] Suppose $P(\theta;\alpha)$ is a *conjugate prior* for $f(x|\theta)$. Show that the following distribution is a *conjugate prior* for $f(x|\theta)$ too.

$$\sum_{i=1}^{m} \beta_i P(\theta;\alpha_i), \quad s.t. \sum_{i=1}^{m} \beta_i = 1$$

Solution:
By definition, $P(\theta|x)$ and $P(\theta;\alpha)$ are from one family.

$$P(\theta|x) = \frac{f(x|\theta)P(\theta;\alpha)}{\gamma(\alpha)}, \quad \gamma(\alpha) = \int f(x|\theta')P(\theta';\alpha)d\theta'$$

For new posterior:

$$P'(\theta|x) = \frac{f(x|\theta)P(\theta;\alpha,\beta)}{P(x)} = \frac{f(x|\theta)\sum_i \beta_i P(\theta;\alpha_i)}{\int f(x|\theta')P(\theta';\alpha,\beta)d\theta'} = \frac{\sum_i \beta_i f(x|\theta)P(\theta;\alpha_i)}{\int \sum_i \beta_i f(x|\theta')P(\theta';\alpha_i)d\theta'}$$

$$= \frac{\sum_i \beta_i P(\theta|x)\gamma(\alpha_i)}{\sum_i \beta_i \int f(x|\theta')P(\theta';\alpha_i)d\theta'} = \frac{\sum_i \beta_i P(\theta|x)\gamma(\alpha_i)}{\sum_i \beta_i \gamma(\alpha_i)} = \sum_i \frac{\beta_i \gamma(\alpha_i)}{\sum_{j=1}^{m} \beta_j \gamma(\alpha_j)} P(\theta|x) = \sum_{i=1}^{m} \beta_i' P(\theta|x)$$

$$\Rightarrow \beta_i' = \frac{\beta_i \gamma(\alpha_i)}{\sum_{j=1}^{m} \beta_j \gamma(\alpha_j)} \quad \Rightarrow \sum_{i=1}^{m} \beta_i' = 1$$

Since $P(\theta|x)$ and $P(\theta;\alpha_i)$ are from one family and $\sum_{i=1}^{m} \beta_i' = 1$, posterior distribution $P'(\theta|x)$ is in the same probability distribution family as the prior distribution $P(\theta;\alpha,\beta) = \sum_i \beta_i P(\theta;\alpha_i)$.