(1) **Cramér-Rao lower bound - Simulation**

In Homework 7 Exercise 2 a density $f(x|\theta) = \theta x^{\theta-1}$ for $0 < x < 1$ and $\theta > 0$ was given. The goal was to find a sutable function g of the parameter θ such that there exists an unbiased estimator of $g(\theta)$ which attains the Cramér-Rao lower bound.

A unbiased statistic which attains the Cramér-Rao lower bound is for $g(\theta) = \frac{1}{\theta}$ given by

$$S_n(X_1, ..., X_n) = -\frac{1}{n}\sum_{i=1}^{n} \ln(X_i).$$

Implement the following steps in `R`:

(a) Write pdf `dhw`, cdf `phw`, quantile `qhw` and random sampling function `rhw` for the above distribution parameterized by θ (see for example `?runif, ?rnorm`).

Hint: Given an strict monotone continuous cdf F, then $F^{-1}(U)$ is distributed with cdf F for $U \sim U(0,1)$.

(b) Fix an arbitrary θ and perform a simulation with growing sample size n = 500, 1000, 1500, ..., 10000 each with 100 replications for the estimation of $g(\theta)$ with the statistic S_n.

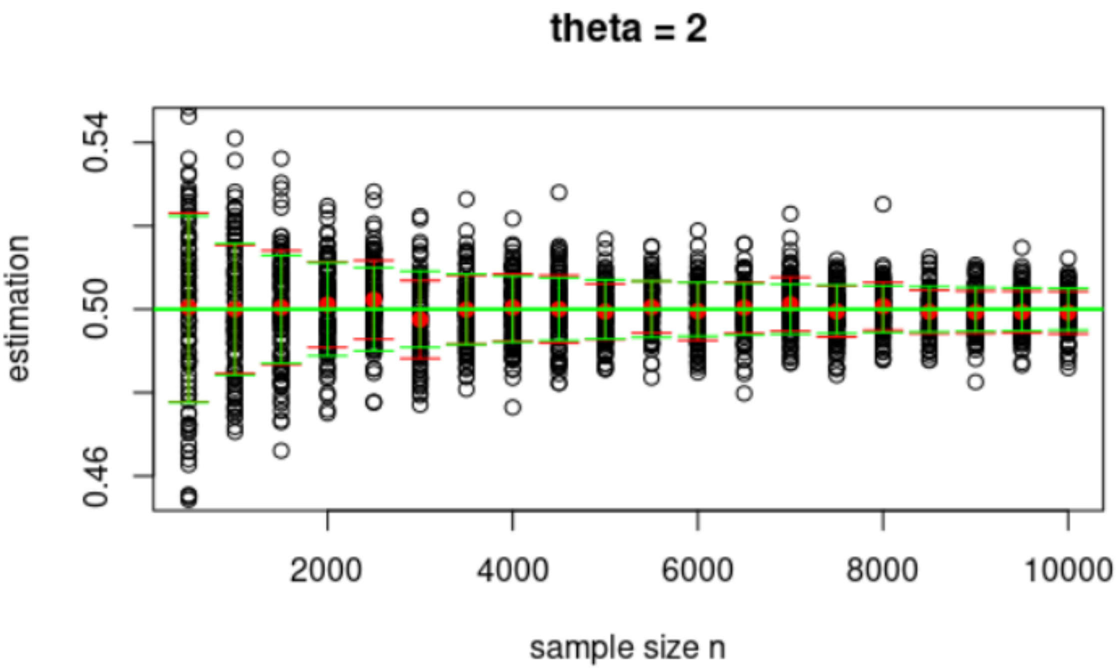
(c) Create a scatter plot of all the estimates over the sample size, add the sample mean and standard deviation aggregated over the sample size to the plot. Finally, add the theoretical mean and standard deviation of the statistic S_n.

a) $\theta > 0$; $f_\theta(x) = \theta x^{\theta-1} \mathbb{1}_{(0,1)}(x)$, hence $F_\theta(x) = \begin{cases} 0, & \text{if } x \le 0 \\ x^\theta, & \text{if } 0 < x < 1 \\ 1, & \text{if } 1 \le x \end{cases}$

Consider any two numbers $x, p \in (0,1)$. We have $F_\theta(x) = p \Leftrightarrow x^\theta = p \Leftrightarrow x = p^{\frac{1}{\theta}}$.

Hence, the Quantile function $Q_\theta : [0,1] \to [0,1]$ is given by $Q_\theta(p) := p^{\frac{1}{\theta}}$

c) We know from last weeks Homework, that $\mathbb{E}(S_n) = \frac{1}{\theta}$ and $\mathbb{V}ar(S_n) = \frac{1}{n\theta^2}$.



(2) **Sufficient statistic and point estimator statistics**
Let $X_1, \dots X_n$ be a random sample from a population with pdf

$$f(x|\theta) = \begin{cases} \frac{\theta}{x^2}, & \theta \le x \\ 0, & \text{otherwise} \end{cases}$$

with unknown $\theta > 0$. Use the Factorization theorem to obtain a sufficient statistic for θ.

The likelihood is given by $L(x|\theta) = \begin{cases} \theta^n \prod_{i=1}^{n} x_i^{-2} & , \text{if } \theta \le \min\{x_i \mid 1 \le i \le n\} \\ 0 & , \text{otherwise} \end{cases}$

$T(x) := \left(\prod_{i=1}^{n} x_i^{-2}, \min\{x_i \mid 1 \le i \le n\} \right)$ $\qquad h(x) := 1$

$g(y|\theta) := \begin{cases} \theta^n y_1 & , \text{if } \theta \le y_2 \\ 0 & , \text{otherwise} \end{cases}$

We have $L(x,\theta) = g(T(x)|\theta)\, h(x)$, hence $T(x)$ is a sufficient statistic for θ by the factorization theorem.

(3) **Minimal sufficient statistic 1**

Let $X_1, \ldots, X_n$ be a random sample from a population with $\mathcal{N}(\mu, \mu)$ distribution, where $\mu > 0$ is unknown.

(a) Show that the statistic $\sum X_i^2$ is minimal sufficient in the $\mathcal{N}(\mu, \mu)$ family.

(b) Show that the statistic $(\sum X_i, \sum X_i^2)$ is sufficient but not minimal sufficient in the $\mathcal{N}(\mu, \mu)$ family.

a) $L(x|\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\mu}} \exp\left(\frac{-(x_i-\mu)^2}{2\mu}\right) = (2\pi\mu)^{-\frac{n}{2}} \exp\left(\sum_{i=1}^{n} \frac{1}{2\mu}\left(2x_i\mu - x_i^2 - \mu^2\right)\right)$

$= (2\pi\mu)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\mu}\sum_{i=1}^{n} x_i^2 - \frac{n\mu}{2}\right) \exp\left(\sum_{i=1}^{n} x_i\right) = g(T(x)|\mu)\, h(x)$, where

$g(z|\mu) = (2\pi\mu)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\mu} z - \frac{n\mu}{2}\right)$ and $h(x) = \exp\left(\sum_{i=1}^{n} x_i\right)$

Hence, $T(x) = \sum_{i=1}^{n} x_i^2$ is sufficient.

$\frac{L(x|\mu)}{L(y|\mu)} = \exp\left(\frac{1}{2\mu}\left(\sum_{i=1}^{n} y_i^2 - \sum_{i=1}^{n} x_i^2\right)\right)$ is constant as a function of μ, if and only if $T(y) = \sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} x_i^2 = T(x)$, hence $T(X)$ is minimal sufficient.

b) $S(x) = \left(\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i^2\right)$ is clearly sufficient, since one component is $T(X)$, simply take $\tilde{g}(z|\mu) = (2\pi\mu)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\mu} z_2 - \frac{n\mu}{2}\right)$, but it is not minimal sufficient, since $\pi_2(S(X)) = T(X)$.

(4) **Minimal sufficient statistic 2**
Let $X_1, \dots, X_n$ be a random sample from a population with pdf

$$f(x|\theta) = \begin{cases} \frac{2x}{\theta}, & 0 < x < \theta \\ 0, & \text{otherwise} \end{cases},$$

with unknown parameter $\theta > 0$. Find a minimal sufficient statistic for θ.

$$L(x|\theta) = \begin{cases} \left(\frac{2}{\theta}\right)^n \prod_{i=1}^{n} x_i & , \text{if } 0 < \min\{x_i \mid i=1,\dots,n\} \leq \max\{x_i \mid i=1,\dots,n\} < \theta \\ 0 & , \text{otherwise} \end{cases}$$

If we constrain x to $(\mathbb{R}^+)^n$ and define $T(x) := \max\{x_i \mid i=1,\dots,n\}$, $h(x) = \prod_{i=1}^{n} x_i$, and

$g(z|\theta) = \begin{cases} \left(\frac{2}{\theta}\right)^n z_1 & , \text{if } z_2 < \theta \\ 0 & , \text{else} \end{cases}$ then $L(x|\theta) = g(T(x)|\theta)\, h(x)$, hence $T(X)$ is sufficient

If $T(x) = T(y)$, then $\frac{L(x|\theta)}{L(y|\theta)} = \frac{h(x)}{h(y)}$ is constant as a function of $\theta \in (T(x), \infty)$.

If $T(y) < T(x)$, then we choose $\theta_1, \theta_2 \in \mathbb{R}^+$ such that $T(x) < \theta_2$ and $T(y) < \theta_1 < T(x)$, hence $\frac{L(x,\theta_1)}{L(y,\theta_1)} = 0$ and $\frac{L(x,\theta_2)}{L(x,\theta_2)} = \frac{h(x)}{h(y)} > 0$, hence $\frac{L(x|\theta)}{L(y|\theta)}$ is not constant as a function of θ. We conclude that $T(X)$ is a minimal sufficient statistic.

(5) **Sufficiency, bias, Rao-Blackwell theorem**
Let $X_1, \ldots, X_n$ be i.i.d. $Poi(\lambda)$, with unknown $\lambda > 0$.

(a) Show that $Y = \sum_{n=1}^{n} X_i$ is a sufficient statistic for λ.
(b) Find an unbiased estimator of $p_r = P(X = r)$, which depends only on X_1.
Find $P(X_1 = r | Y = k)$ both for $k \geq r$ and $k < r$.
Hence use the Rao-Blackwell theorem to improve your estimator of p_r.

a) $$L(x|\lambda) = \begin{cases} \lambda^{\sum_{i=1}^n x_i} \left(\prod_{i=1}^n x_i!\right)^{-1} \exp(-n\lambda) & , \text{if } x = 0, 1, \ldots \\ 0 & , \text{otherwise} \end{cases}$$

From now on we only consider $x \in \{0, 1, \ldots\}$. Let $g(z|\lambda) := \lambda^z \exp(-n\lambda)$, and $h(x) := \left(\prod_{i=1}^n x_i!\right)$. We have $L(x|\lambda) = g(Y, \lambda) h(X)$, hence Y is a sufficient statistic for λ.

b) $p_r = \mathbb{P}(X = r) = \frac{\lambda^r}{r!} e^{-\lambda}$, $\quad \hat{p}_r(x) := \begin{cases} 1 & , \text{if } x_1 = r \\ 0 & , \text{otherwise} \end{cases}$, then $\hat{p}_r(X)$ is an unbiased estimator of p_r, since $\mathbb{E}_\lambda(\hat{p}_r(X)) = \frac{\lambda^r}{r!} e^{-\lambda} = p_r$

Let $k < r$, if there was w with $k = Y(w) = \sum_{i=1}^n x_i(w)$ and $X_1(w) = r$, then $r = X_1(w) \leq \sum_{i=1}^n x_i(w) = k$ ↯, hence $\mathbb{P}(X_1 = r | Y = k) = 0$

Let $r \leq k$, then

$$\mathbb{P}(X_1 = r | Y = k) = \frac{\mathbb{P}(X_1 = r, \sum_{i=2}^n x_i = k - r)}{\mathbb{P}(Y = k)} = \frac{\mathbb{P}(X_1 = r)\, \mathbb{P}(\sum_{i=2}^n x_i = k - r)}{\mathbb{P}(Y = k)}$$

$$= \frac{\lambda^r}{r!} e^{-\lambda} \frac{((n-1)\lambda)^{k-r}}{(k-r)!} e^{-(n-1)\lambda} \left(\frac{(n\lambda)^k}{k!} e^{-n\lambda}\right)^{-1}$$

$$= \frac{k!}{r!(k-r)!} \frac{(n-1)^{k-r}}{n^k} = \binom{k}{r} \frac{(n-1)^{k-r}}{n^k}$$

We used that $\sum_{i=2}^n X_i \sim Poi((n-1)\lambda)$ and $Y \sim Poi(n\lambda)$

Hence $$\phi(k) = \mathbb{E}(X_1 = r | Y = k) = \mathbb{P}(X_1 = r | Y = k) = \begin{cases} \binom{k}{r} \frac{(n-1)^{k-r}}{n^k} & , \text{if } 0 \leq r \leq k \\ 0 & , \text{otherwise} \end{cases}$$

By the Rao-Blackwell Theorem, $\phi(Y)$ is an unbiased estimator of p_r with $\mathrm{Var}_{p_r}(\phi(Y)) \leq \mathrm{Var}(\hat{p}_r(X))$.