Alberto Andrés Valdés González.
Degree: Mathematical Engineer.
Work position: Data Scientist.
Mail: anvaldes@uc.cl/alberto.valdes.gonzalez.96@gmail.com
Location: Santiago, Chile.

Fisher Information

Properties of estimators

Unbiased:

$$\boxed{\mathbb{E}[\hat{\theta}] = \theta}$$

Fisher information

$$\boxed{I(\theta) = \mathbb{E}\left[\left(\frac{\partial ln[f(\vec{x}|\theta)]}{\partial \theta}\right)^2\right] = \mathbb{E}\left[-\frac{\partial^2 ln[f(\vec{x}|\theta)]}{\partial \theta^2}\right]}$$

If we have more than one estimador, that is to say $\theta_1, \theta_2, ..., \theta_n$ then:

$$\boxed{I(\theta) = -\begin{bmatrix} \frac{\partial^2 ln[f(\vec{x}|\theta)]}{\partial \theta_1^2} & \frac{\partial^2 ln[f(\vec{x}|\theta)]}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 ln[f(\vec{x}|\theta)]}{\partial \theta_1 \partial \theta_n} \\ \frac{\partial^2 ln[f(\vec{x}|\theta)]}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 ln[f(\vec{x}|\theta)]}{\partial \theta_2^2} & \cdots & \frac{\partial^2 ln[f(\vec{x}|\theta)]}{\partial \theta_2 \partial \theta_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 ln[f(\vec{x}|\theta)]}{\partial \theta_n \partial \theta_1} & \frac{\partial^2 ln[f(\vec{x}|\theta)]}{\partial \theta_n \partial \theta_2} & \cdots & \frac{\partial^2 ln[f(\vec{x}|\theta)]}{\partial \theta_n \partial \theta_n} \end{bmatrix}}$$

Cramer-Rao lower bound

Let $\hat{\theta}$ be an **unbiased estimator**, then:

$$\boxed{\mathbb{V}(\hat{\theta}) \geq \frac{1}{I(\theta)}}$$

Properties of estimators

Efficient:

We say that unbiased estimator $\hat{\theta}$ is **efficient** for θ if its variance reaches the Cramer-Rao lower bond. That is to say:

$$\boxed{\mathbb{V}(\hat{\theta}) = \frac{1}{I(\theta)}}$$

Asymptotic normality

If $\hat{\theta}$ is **MLE** and **unbiased** then:

$$(\hat{\theta} - \theta) \sim N\left(0, I(\theta)^{-1}\right) \text{ as } n \to +\infty$$

Example: $X_i \overset{\text{iid}}{\sim} Exp(\lambda)$. Compute $I(\lambda)$, $\mathbb{E}[\hat{\lambda}]$ and $\mathbb{V}(\hat{\lambda})$.

Remember:

$$\boxed{l(\lambda) = -n \cdot ln(\lambda) - \left(\frac{1}{\lambda}\right) \cdot \sum_{i=1}^{n} x_i}$$

$\Rightarrow$

$$\frac{\partial l(\lambda)}{\partial \lambda} = -n \cdot \frac{1}{\lambda} + \frac{1}{\lambda^2} \cdot \sum_{i=1}^{n} x_i$$

$\Rightarrow$

$$\frac{\partial^2 l(\lambda)}{\partial \lambda^2} = n \cdot \frac{1}{\lambda^2} - \frac{2}{\lambda^3} \cdot \sum_{i=1}^{n} x_i$$

$\Rightarrow$

$$\boxed{-\frac{\partial^2 l(\lambda)}{\partial \lambda^2} = -n \cdot \frac{1}{\lambda^2} + \frac{2}{\lambda^3} \cdot \sum_{i=1}^{n} x_i}$$

$\Rightarrow$

$$\mathbb{E}\left[-\frac{\partial^2 l(\lambda)}{\partial \lambda^2}\right] = \mathbb{E}\left[-n \cdot \frac{1}{\lambda^2} + \frac{2}{\lambda^3} \cdot \sum_{i=1}^{n} x_i\right] = -n \cdot \frac{1}{\lambda^2} + \frac{2}{\lambda^3} \cdot \sum_{i=1}^{n} \mathbb{E}[x_i] = -n \cdot \frac{1}{\lambda^2} + \frac{2}{\lambda^3} \cdot \sum_{i=1}^{n} \lambda$$

$$= -n \cdot \frac{1}{\lambda^2} + 2 \cdot \frac{1}{\lambda^3} \cdot n \cdot \lambda = \frac{n}{\lambda^2}$$

$\Rightarrow$

$$\boxed{I(\lambda) = \frac{n}{\lambda^2}} \quad (1)$$

On the other hand:

$$\boxed{\hat{\lambda} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i}$$

$\Rightarrow$

$$\mathbb{E}[\hat{\lambda}] = \mathbb{E}\left[\frac{1}{n} \cdot \sum_{i=1}^{n} x_i\right] = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbb{E}[x_i] = \frac{1}{n} \cdot \sum_{i=1}^{n} \lambda = \frac{1}{n} \cdot n \cdot \lambda = \lambda$$

$\Rightarrow$

$$\boxed{\mathbb{E}[\hat{\lambda}] = \lambda} \quad (2)$$

Also we have that:

$$\mathbb{V}[\hat{\lambda}] = \mathbb{V}\left[\frac{1}{n} \cdot \sum_{i=1}^{n} x_i\right] = \frac{1}{n^2} \cdot \sum_{i=1}^{n} \mathbb{V}[x_i] = \frac{1}{n^2} \cdot \sum_{i=1}^{n} \lambda^2 = \frac{1}{n^2} \cdot n \cdot \lambda^2 = \frac{\lambda^2}{n}$$

$\Rightarrow$

$$\boxed{\mathbb{V}[\hat{\lambda}] = \frac{\lambda^2}{n}} \quad (3)$$

Conlusion 1: How $\mathbb{E}[\hat{\lambda}] = \lambda$ then the estimator is **unbiased**.

Conlusion 2: How $\mathbb{V}[\hat{\lambda}] = \dfrac{1}{I(\lambda)}$ then the estimator is **efficient**.

Conlusion 3: How $\hat{\lambda}$ is **MLE** then $(\hat{\lambda} - \lambda) \sim N\left(0, \frac{\lambda^2}{n}\right)$ as $n \to +\infty$.

Example: $X_i \overset{\text{iid}}{\sim} Poisson(\lambda)$. Compute $I(\lambda)$, $\mathbb{E}[\hat{\lambda}]$ and $\mathbb{V}(\hat{\lambda})$.

Remember:

$$\boxed{l(\lambda) = -n \cdot \lambda + \left(\sum_{i=1}^{n} x_i\right) \cdot ln(\lambda) - \sum_{i=1}^{n} ln[(x_i)!\ ]}$$

$\Rightarrow$

$$\frac{\partial l(\lambda)}{\partial \lambda} = -n + \left(\sum_{i=1}^{n} x_i\right) \cdot \frac{1}{\lambda}$$

$\Rightarrow$

$$\frac{\partial^2 l(\lambda)}{\partial \lambda^2} = -\frac{1}{\lambda^2} \cdot \left(\sum_{i=1}^{n} x_i\right)$$

$\Rightarrow$

$$\boxed{-\frac{\partial^2 l(\lambda)}{\partial \lambda^2} = \frac{1}{\lambda^2} \cdot \left(\sum_{i=1}^{n} x_i\right)}$$

$\Rightarrow$

$$\mathbb{E}\left[-\frac{\partial^2 l(\lambda)}{\partial \lambda^2}\right] = \mathbb{E}\left[\frac{1}{\lambda^2} \cdot \left(\sum_{i=1}^{n} x_i\right)\right] = \frac{1}{\lambda^2} \cdot \sum_{i=1}^{n} \mathbb{E}[x_i] = \frac{1}{\lambda^2} \cdot \sum_{i=1}^{n} \lambda = \frac{1}{\lambda^2} \cdot n \cdot \lambda = \frac{n}{\lambda}$$

$\Rightarrow$

$$\boxed{I(\lambda) = \frac{n}{\lambda}} \quad (1)$$

On the other hand:

$$\boxed{\hat{\lambda} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i}$$

$\Rightarrow$

$$\mathbb{E}[\hat{\lambda}] = \mathbb{E}\left[\frac{1}{n} \cdot \sum_{i=1}^{n} x_i\right] = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbb{E}[x_i] = \frac{1}{n} \cdot \sum_{i=1}^{n} \lambda = \frac{1}{n} \cdot n \cdot \lambda = \lambda$$

$\Rightarrow$

$$\boxed{\mathbb{E}[\hat{\lambda}] = \lambda} \quad (2)$$

Also we have that:

$$\mathbb{V}[\hat{\lambda}] = \mathbb{V}\left[\frac{1}{n} \cdot \sum_{i=1}^{n} x_i\right] = \frac{1}{n^2} \cdot \sum_{i=1}^{n} \mathbb{V}[x_i] = \frac{1}{n^2} \cdot \sum_{i=1}^{n} \lambda = \frac{1}{n^2} \cdot n \cdot \lambda = \frac{\lambda}{n}$$

$\Rightarrow$

$$\boxed{\mathbb{V}[\hat{\lambda}] = \frac{\lambda}{n}} \quad (3)$$

Conlusion 1: How $\mathbb{E}[\hat{\lambda}] = \lambda$ then the estimator is **unbiased**.

Conlusion 2: How $\mathbb{V}[\hat{\lambda}] = \dfrac{1}{I(\lambda)}$ then the estimator is **efficient**.

Conlusion 3: How $\hat{\lambda}$ is **MLE** then $(\hat{\lambda} - \lambda) \sim N\left(0, \frac{\lambda}{n}\right)$ as $n \to +\infty$.

Example: $X_i \overset{\text{iid}}{\sim} Bernoulli(p)$. Compute $I(p)$, $\mathbb{E}[\hat{p}]$ and $\mathbb{V}(\hat{p})$.

Remember:

$$\boxed{l(p) = \left(\sum_{i=1}^{n} x_i\right) \cdot ln(p) + \left(n - \sum_{i=1}^{n} x_i\right) \cdot ln(1-p)}$$

$\Rightarrow$

$$\frac{\partial l(p)}{\partial p} = \left(\sum_{i=1}^{n} x_i\right) \cdot \frac{1}{p} - \left(n - \sum_{i=1}^{n} x_i\right) \cdot \frac{1}{1-p}$$

$\Rightarrow$

$$\frac{\partial^2 l(p)}{\partial p^2} = -\left(\sum_{i=1}^{n} x_i\right) \cdot \frac{1}{p^2} - \left(n - \sum_{i=1}^{n} x_i\right) \cdot \frac{1}{(1-p)^2}$$

$\Rightarrow$

$$\boxed{-\frac{\partial^2 l(p)}{\partial p^2} = \left(\sum_{i=1}^{n} x_i\right) \cdot \frac{1}{p^2} + \left(n - \sum_{i=1}^{n} x_i\right) \cdot \frac{1}{(1-p)^2}}$$

$\Rightarrow$

$$\begin{aligned}
\mathbb{E}\left[-\frac{\partial^2 l(p)}{\partial p^2}\right] &= \mathbb{E}\left[\left(\sum_{i=1}^{n} x_i\right) \cdot \frac{1}{p^2} + \left(n - \sum_{i=1}^{n} x_i\right) \cdot \frac{1}{(1-p)^2}\right] \\
&= \frac{1}{p^2} \cdot \sum_{i=1}^{n} \mathbb{E}[x_i] + \frac{1}{(1-p)^2} \cdot \left(n - \sum_{i=1}^{n} \mathbb{E}[x_i]\right)
\end{aligned}$$

$$= \frac{1}{p^2} \cdot \sum_{i=1}^{n} p + \frac{1}{(1-p)^2} \cdot \left(n - \sum_{i=1}^{n} p \right) = \frac{1}{p^2} \cdot n \cdot p + \frac{1}{(1-p)^2} \cdot (n - n \cdot p)$$

$$= \frac{n}{p} + \frac{n}{(1-p)} = \frac{n}{p \cdot (1-p)}$$

$\Rightarrow$

$$\boxed{I(p) = \frac{n}{p \cdot (1-p)}} \quad (1)$$

On the other hand:

$$\boxed{\hat{p} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i}$$

$\Rightarrow$

$$\mathbb{E}[\hat{p}] = \mathbb{E}\left[\frac{1}{n} \cdot \sum_{i=1}^{n} x_i \right] = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbb{E}[x_i] = \frac{1}{n} \cdot \sum_{i=1}^{n} p = \frac{1}{n} \cdot n \cdot p = p$$

$\Rightarrow$

$$\boxed{\mathbb{E}[\hat{p}] = p} \quad (2)$$

Also we have that:

$$\mathbb{V}[\hat{p}] = \mathbb{V}\left[\frac{1}{n} \cdot \sum_{i=1}^{n} x_i \right] = \frac{1}{n^2} \cdot \sum_{i=1}^{n} \mathbb{V}[x_i] = \frac{1}{n^2} \cdot \sum_{i=1}^{n} p \cdot (1-p) = \frac{1}{n^2} \cdot n \cdot p \cdot (1-p) = \frac{p \cdot (1-p)}{n}$$

$\Rightarrow$

$$\boxed{\mathbb{V}[\hat{\lambda}] = \frac{p \cdot (1-p)}{n}} \quad (3)$$

Conlusion 1: How $\mathbb{E}[\hat{p}] = p$ then the estimator is **unbiased**.

Conlusion 2: How $\mathbb{V}[\hat{p}] = \dfrac{1}{I(p)}$ then the estimator is **efficient**.

Conlusion 3: How $\hat{p}$ is **MLE** then $(\hat{p} - p) \sim N\left(0, \frac{p \cdot (1-p)}{n}\right)$ as $n \to +\infty$.

Example: $X_i \overset{\text{iid}}{\sim} Normal(\mu, \sigma)$. Compute $I(\mu, \sigma)$.

Remember:

$$\boxed{l(\mu,\sigma) = -\frac{n}{2} \cdot ln(2\pi) - n \cdot ln(\sigma) - \frac{1}{2} \cdot \sum_{i=1}^{n} \left[\frac{x_i - u}{\sigma}\right]^2}$$

$\Rightarrow$

$$\frac{\partial l(\mu,\sigma)}{\partial \mu} = \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n} [x_i - \mu]$$

$$\frac{\partial l(\mu,\sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \cdot \sum_{i=1}^{n} [x_i - \mu]^2$$

$\Rightarrow$

$$\frac{\partial^2 l(\mu,\sigma)}{\partial \mu^2} = -\frac{n}{\sigma^2}$$

$$\frac{\partial^2 l(\mu,\sigma)}{\partial \sigma \partial \mu} = -\frac{2}{\sigma^3} \cdot \sum_{i=1}^{n} [x_i - \mu]$$

$$\frac{\partial^2 l(\mu,\sigma)}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{3}{\sigma^4} \cdot \sum_{i=1}^{n} [x_i - \mu]^2$$

Now:

$$\mathbb{E}\left[-\frac{\partial^2 l(\mu,\sigma)}{\partial \mu^2}\right] = \mathbb{E}\left[\frac{n}{\sigma^2}\right] = \frac{n}{\sigma^2}$$

$\Rightarrow$

$$\boxed{\mathbb{E}\left[-\frac{\partial^2 l(\mu,\sigma)}{\partial \mu^2}\right] = \frac{n}{\sigma^2}} \quad (1)$$

$$\mathbb{E}\left[\frac{\partial^2 l(\mu,\sigma)}{\partial \sigma \partial \mu}\right] = \mathbb{E}\left[\frac{2}{\sigma^3} \cdot \sum_{i=1}^{n} [x_i - \mu]\right] = \frac{2}{\sigma^3} \cdot \sum_{i=1}^{n} \mathbb{E}[X_i - \mu]$$

$$= \frac{2}{\sigma^3} \cdot \sum_{i=1}^{n} (\mathbb{E}[X_i] - \mu) = \frac{2}{\sigma^3} \cdot \sum_{i=1}^{n} \mathbb{E}[X_i - \mu] = \frac{2}{\sigma^3} \cdot \sum_{i=1}^{n} (\mu - \mu) = 0$$

$\Rightarrow$

$$\boxed{\mathbb{E}\left[-\frac{\partial^2 l(\mu,\sigma)}{\partial\sigma\partial\mu}\right] = \mathbb{E}\left[-\frac{\partial^2 l(\mu,\sigma)}{\partial\mu\partial\sigma}\right] = 0} \quad (2)$$

$$\mathbb{E}\left[-\frac{\partial^2 l(\mu,\sigma)}{\partial\sigma^2}\right] = \mathbb{E}\left[-\frac{n}{\sigma^2} + \frac{3}{\sigma^4}\cdot\sum_{i=1}^{n}[x_i-\mu]^2\right] = -\frac{n}{\sigma^2} + \frac{3}{\sigma^4}\cdot\sum_{i=1}^{n}\mathbb{E}[(x_i-\mu)^2]$$
$$= -\frac{n}{\sigma^2} + \frac{3}{\sigma^4}\cdot\sum_{i=1}^{n}\mathbb{V}(x_i) = -\frac{n}{\sigma^2} + \frac{3\cdot n\cdot\sigma^2}{\sigma^4} = \frac{2n}{\sigma^2}$$

$\Rightarrow$

$$\boxed{\mathbb{E}\left[-\frac{\partial^2 l(\mu,\sigma)}{\partial\sigma^2}\right] = \frac{2n}{\sigma^2}} \quad (3)$$

$\Rightarrow$

$$\boxed{I(\mu,\sigma) = \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{2n}{\sigma^2} \end{bmatrix}}$$
