Estimation Case Study: The Exponential Distribution

Gene Quinn

Overview

In this case study we will focus on the problem of estimating the parameter θ of an exponential distribution.

Overview

In this case study we will focus on the problem of estimating the parameter θ of an exponential distribution.

In the process, we will:

- Derive the maximum likelihood estimator $\hat{\theta}_{mle}$ for θ.
- Derive the method of moments estimator for $\hat{\theta}_{mom}$ for θ.
- Show that they are the same: $\hat{\theta}_{mle} = \hat{\theta}_{mom} = \hat{\theta}$
- Derive the mean and variance of $\hat{\theta}$
- Show that $\hat{\theta}$ is efficient for θ (it achieves the Cramer-Rao lower bound)

The Expnential Distribution

We will find it expedient to write the pdf of the exponential distribution in the following form:

$$f_X(x;\theta) \;=\; \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right) \qquad x \in (0,\infty)$$

The Expnential Distribution

We will find it expedient to write the pdf of the exponential distribution in the following form:

$$f_X(x;\theta) \;=\; \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right) \qquad x \in (0,\infty)$$

As it suits our purpose we will treat f as a function of x alone or as a function of x and θ.

The Expnential Distribution

We will find it expedient to write the pdf of the exponential distribution in the following form:

$$f_X(x;\theta) \;=\; \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right) \qquad x \in (0,\infty)$$

As it suits our purpose we will treat f as a function of x alone or as a function of x and θ.

We assume at the outset that we have a random sample of size n,

$$x = \{x_1, \ldots, x_n\}$$

from the exponential population under study and wish to estimate the parameter θ.

Method of Moments Estimate

We obtain the method of moments estimate for θ by equating the expected value of x,

$$\mathsf{E}(x) \;=\; \int_0^\infty x \cdot f(x;\theta)\,dx \quad = \quad \theta$$

to the first *sample* moment,

$$\overline{x} \;=\; \frac{1}{n}\sum_{i=1}^{n} x_i$$

Method of Moments Estimate

The resulting equation is:

$$\theta = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Method of Moments Estimate

The resulting equation is:

$$\theta = \frac{1}{n}\sum_{i=1}^{n} x_i$$

We obtain the method of moments estimate for θ by solving this equation for θ and defining $\hat{\theta}_{mom}$ to be the result.

Method of Moments Estimate

The resulting equation is:

$$\theta = \frac{1}{n}\sum_{i=1}^{n} x_i$$

We obtain the method of moments estimate for θ by solving this equation for θ and defining $\hat{\theta}_{mom}$ to be the result.

Of course in this case there is nothing to solve because E$(x) = \theta$, so all we have to do is declare that

$$\hat{\theta}_{mom} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Maximum Likelihood Estimate

The maximum likelihood estimate for θ is the value of θ that maximizes the likelihood function of the sample.

Recall that the likelihood function of the sample is obtained as the product of the density function values for each x_i, considered as a function of θ:

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} \frac{\exp\left(-\frac{x_i}{\theta}\right)}{\theta}$$

Maximum Likelihood Estimate

The maximum likelihood estimate for θ is the value of θ that maximizes the likelihood function of the sample.

Recall that the likelihood function of the sample is obtained as the product of the density function values for each x_i, considered as a function of θ:

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} \frac{\exp\left(-\frac{x_i}{\theta}\right)}{\theta}$$

$$= \left(\frac{1}{\theta^n}\right) \exp\left(-\frac{\sum x_i}{\theta}\right)$$

Maximum Likelihood Estimate

To find the value of θ that maximizes $L(\theta)$, we differentiate

$$L(\theta) = \left(\frac{1}{\theta^n}\right) \exp\left(-\frac{\sum x_i}{\theta}\right)$$

with respect to θ and set the result equal to zero.

Maximum Likelihood Estimate

To find the value of θ that maximizes $L(\theta)$, we differentiate

$$L(\theta) = \left(\frac{1}{\theta^n}\right) \exp\left(-\frac{\sum x_i}{\theta}\right)$$

with respect to θ and set the result equal to zero.

After some simplification, the derivative is

$$\frac{dL(\theta)}{d\theta} = -\exp\left(-\frac{\sum x_i}{\theta}\right) \theta^{-(n+2)} \left(-\sum_{i=1}^{n} x_i + n\theta\right)$$

Maximum Likelihood Estimate

For positive values of θ,

$$\frac{dL(\theta)}{d\theta} = -\exp\left(-\frac{\sum x_i}{\theta}\right)\theta^{-(n+2)}\left(-\sum_{i=1}^{n} x_i + n\theta\right)$$

can only be zero if the third factor is zero,

$$-\sum_{i=1}^{n} x_i + n\theta = 0$$

Maximum Likelihood Estimate

For positive values of θ,

$$\frac{dL(\theta)}{d\theta} = -\exp\left(-\frac{\sum x_i}{\theta}\right)\theta^{-(n+2)}\left(-\sum_{i=1}^{n} x_i + n\theta\right)$$

can only be zero if the third factor is zero,

$$-\sum_{i=1}^{n} x_i + n\theta = 0$$

Solving this equation for θ gives the maximum likelihood estimate,

$$\hat{\theta}_{mle} = \frac{\sum x_i}{n}$$

Properties of the Estimator

Since the method of moments and maximum likelihood estimates are the same, we will dispense with the distinction and simply refer to

$$\hat{\theta} = \frac{\sum x_i}{n}$$

Properties of the Estimator

Since the method of moments and maximum likelihood estimates are the same, we will dispense with the distinction and simply refer to

$$\hat{\theta} = \frac{\sum x_i}{n}$$

We now address the properties of $\hat{\theta}$.

In particular, we would like to know:

- is $\hat{\theta}$ unbiased?
- is $\hat{\theta}$ efficient for θ?

Properties of the Estimator

- is $\hat{\theta}$ unbiased?
- is $\hat{\theta}$ efficient for θ?

Recall that $\hat{\theta}$ is unbiased if

$$\mathsf{E}(\hat{\theta}) \;=\; \theta$$

Properties of the Estimator

- is $\hat{\theta}$ unbiased?
- is $\hat{\theta}$ efficient for θ?

Recall that $\hat{\theta}$ is unbiased if

$$\mathsf{E}(\hat{\theta}) = \theta$$

If $\hat{\theta}$ is unbiased, it is efficient for θ if Var$(\hat{\theta}$) is equal to the Cramer-Rao Lower Bound.

Both questions require us to evaluate the moments of $\hat{\theta}$.

Properties of the Estimator

How do we find the expected value and variance of

$$\hat{\theta} = \frac{\sum x_i}{n} ?$$

Properties of the Estimator

How do we find the expected value and variance of

$$\hat{\theta} = \frac{\sum x_i}{n}\ ?$$

One approach is to make use of the fact that the x_is are independently distributed with common density function

$$f(x_i, \theta) = \frac{1}{\theta} \exp\left(-\frac{x_i}{\theta}\right)$$

Properties of the Estimator

Here is an overview of the approach:

First, find the moment-generating function $M_{x_i}(t)$ of each x_i.

Properties of the Estimator

Here is an overview of the approach:

First, find the moment-generating function $M_{x_i}(t)$ of each x_i.

Next, make use of the theorem that says the moment-generating function of x_i/n is $M_{x_i}(t/n)$

Properties of the Estimator

Here is an overview of the approach:

First, find the moment-generating function $M_{x_i}(t)$ of each x_i.

Next, make use of the theorem that says the moment-generating function of x_i/n is $M_{x_i}(t/n)$

Finally, make use of the theorem that says the moment-generating function of the sum of independently distributed random variables is the product of their individual mgf's.

Properties of the Estimator

Here is an overview of the approach:

First, find the moment-generating function $M_{x_i}(t)$ of each x_i.

Next, make use of the theorem that says the moment-generating function of x_i/n is $M_{x_i}(t/n)$

Finally, make use of the theorem that says the moment-generating function of the sum of independently distributed random variables is the product of their individual mgf's.

The final result is that the moment-generating function of $\hat{\theta}$ is

$$M_{\hat{\theta}}(t) \ = \ \prod_{i=1}^{n} M_{x_i}(t/n)$$

Properties of the Estimator

Now we find the moment-generating function for x_i, $M_{x_i}(t)$, by evaluating the integral

$$M_{x_i}(t) = \int_0^\infty e^{tx_i} f(x_i;\theta)\, dx_i = \int_0^\infty e^{tx_i} \frac{1}{\theta} \exp\left(-\frac{x_i}{\theta}\right) dx_i$$

Properties of the Estimator

Now we find the moment-generating function for x_i, $M_{x_i}(t)$, by evaluating the integral

$$M_{x_i}(t) = \int_0^\infty e^{tx_i} f(x_i;\theta)\,dx_i = \int_0^\infty e^{tx_i} \frac{1}{\theta} \exp\left(-\frac{x_i}{\theta}\right) dx_i$$

The result is:

$$M_{x_i}(t) = \frac{1}{1 - t\theta}$$

Properties of the Estimator

Now we use a theorem to find the moment-generating function for x_i/n,

$$M_{x_i/n}(t) \;=\; M_{x_i}\left(\frac{t}{n}\right) \;=\; \frac{1}{1-\frac{t\theta}{n}}$$

Properties of the Estimator

Now we use a theorem to find the moment-generating function for x_i/n,

$$M_{x_i/n}(t) = M_{x_i}\left(\frac{t}{n}\right) = \frac{1}{1 - \frac{t\theta}{n}}$$

Finally, since the x_i are independently distributed and

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \frac{x_i}{n}$$

we can use the theorem that says the moment-generating function of the sum will be the product of the moment-generating functions of the individual terms.

Properties of the Estimator

So the final result is

$$M_{\hat{\theta}}(t) = \prod_{i=1}^{n} M_{x_i/n}(t) = \prod_{i=1}^{n} M_{x_i}\left(\frac{t}{n}\right) = \prod_{i=1}^{n} \frac{1}{1-\frac{t\theta}{n}}$$

$$= \left(\frac{1}{1-\frac{t\theta}{n}}\right)^n$$

Properties of the Estimator

To find E$(\hat{\theta})$, we differentiate $M_{\hat{\theta}}(t)$ with respect to t

$$\frac{dM_{\hat{\theta}}(t)}{dt} = \frac{d}{dt}\left(\frac{1}{1-\frac{t\theta}{n}}\right)^n = \frac{\left(\frac{n}{n-t\theta}\right)^n n\theta}{n-t\theta}$$

Properties of the Estimator

To find E$(\hat{\theta})$, we differentiate $M_{\hat{\theta}}(t)$ with respect to t

$$\frac{dM_{\hat{\theta}}(t)}{dt} = \frac{d}{dt}\left(\frac{1}{1-\frac{t\theta}{n}}\right)^n = \frac{\left(\frac{n}{n-t\theta}\right)^n n\theta}{n-t\theta}$$

Now evaluate the result at $t = 0$ to obtain E$(\hat{\theta})$:

$$E(\hat{\theta}) = \left.\frac{dM_{\hat{\theta}}(t)}{dt}\right|_{t=0} = \frac{\left(\frac{n}{n-0\cdot\theta}\right)^n n\theta}{n-0\cdot\theta} = \frac{(1)^n n\theta}{n} = \theta$$

Properties of the Estimator

To find E$(\hat{\theta})$, we differentiate $M_{\hat{\theta}}(t)$ with respect to t

$$\frac{dM_{\hat{\theta}}(t)}{dt} = \frac{d}{dt}\left(\frac{1}{1-\frac{t\theta}{n}}\right)^n = \frac{\left(\frac{n}{n-t\theta}\right)^n n\theta}{n-t\theta}$$

Now evaluate the result at $t=0$ to obtain E$(\hat{\theta})$:

$$E(\hat{\theta}) = \left.\frac{dM_{\hat{\theta}}(t)}{dt}\right|_{t=0} = \frac{\left(\frac{n}{n-0\cdot\theta}\right)^n n\theta}{n-0\cdot\theta} = \frac{(1)^n n\theta}{n} = \theta$$

This establishes that $\hat{\theta}$ is an unbiased estimator of θ:

$$\mathsf{E}(\hat{\theta}) = \mathsf{E}\left[\frac{\sum x_i}{n}\right] = \theta$$

Properties of the Estimator

To find E$(\hat{\theta}^2)$, we differentiate $M_{\hat{\theta}}(t)$ *twice* with respect to t and evaluate the result at $t = 0$:

$$\frac{d^2 M_{\hat{\theta}}(t)}{dt^2} = \frac{d^2}{dt^2}\left(\frac{1}{1-\frac{t\theta}{n}}\right)^n = \frac{\left(\frac{n}{n-t\theta}\right)^n \theta^2 n(n+1)}{(n-t\theta)^2}$$

Properties of the Estimator

To find E$(\hat{\theta}^2)$, we differentiate $M_{\hat{\theta}}(t)$ *twice* with respect to t and evaluate the result at $t = 0$:

$$\frac{d^2 M_{\hat{\theta}}(t)}{dt^2} = \frac{d^2}{dt^2}\left(\frac{1}{1-\frac{t\theta}{n}}\right)^n = \frac{\left(\frac{n}{n-t\theta}\right)^n \theta^2 n(n+1)}{(n-t\theta)^2}$$

Now evaluate the result at $t = 0$ to obtain E$(\hat{\theta}^2)$:

$$E(\hat{\theta}^2) = \left.\frac{d^2 M_{\hat{\theta}}(t)}{dt^2}\right|_{t=0} = \frac{\left(\frac{n}{n-0\cdot\theta}\right)^n \theta^2 n(n+1)}{(n-0\cdot\theta)^2} = \frac{n+1}{n}\theta^2$$

Properties of the Estimator

Finally we calculate $\text{Var}(\hat{\theta})$ as

$$\text{Var}(\hat{\theta}) = \text{E}(\hat{\theta}^2) - \left[\text{E}(\hat{\theta})\right]^2$$

$$= \left(\frac{n+1}{n}\right)\theta^2 - [\theta]^2 = \frac{\theta^2}{n}$$

Properties of the Estimator

Finally we calculate Var$(\hat{\theta})$ as

$$\text{Var}(\hat{\theta}) \;=\; \text{E}(\hat{\theta}^2) - \Big[\text{E}(\hat{\theta})\Big]^2$$

$$= \left(\frac{n+1}{n}\right)\theta^2 - [\theta]^2 \;=\; \frac{\theta^2}{n}$$

All that remains is to determine whether Var$(\hat{\theta})$ equals the Cramer-Rao lower bound for the variance of an unbiased estimator for θ.

Properties of the Estimator

Finally we calculate Var$(\hat{\theta})$ as

$$\text{Var}(\hat{\theta}) \;=\; \text{E}(\hat{\theta}^2) - \left[\text{E}(\hat{\theta})\right]^2$$

$$= \left(\frac{n+1}{n}\right)\theta^2 - [\theta]^2 \;=\; \frac{\theta^2}{n}$$

All that remains is to determine whether Var$(\hat{\theta})$ equals the Cramer-Rao lower bound for the variance of an unbiased estimator for θ.

If it does, we will say that $\hat{\theta}$ is an efficient estimator for the parameter θ.

Cramer-Rao Inequality

To determine the Cramer-Rao lower bound for the variance of an unbiased estimator for θ, we can either evaluate

$$\left\{ n\mathsf{E}\left[\left(\frac{\partial \ln f_X(x;\theta)}{\partial \theta} \right)^2 \right] \right\}^{-1}$$

or

$$\left\{ -n\mathsf{E}\left[\frac{\partial^2 \ln f_X(x;\theta)}{\partial \theta^2} \right] \right\}^{-1}$$

Cramer-Rao Inequality

To determine the Cramer-Rao lower bound for the variance of an unbiased estimator for θ, we can either evaluate

$$\left\{ n\mathsf{E}\left[\left(\frac{\partial \ln f_X(x;\theta)}{\partial \theta} \right)^2 \right] \right\}^{-1}$$

or

$$\left\{ -n\mathsf{E}\left[\frac{\partial^2 \ln f_X(x;\theta)}{\partial \theta^2} \right] \right\}^{-1}$$

Either way, we should get the same result.

Cramer-Rao Inequality

Since

$$f(x;\theta) \;=\; \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right)$$

on substitution the second form of the lower bound becomes

$$\left\{-n\mathsf{E}\left[\frac{\partial^2}{\partial\theta^2}\ln\left(\frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right)\right)\right]\right\}^{-1}$$

Cramer-Rao Inequality

Since

$$f(x;\theta) \;=\; \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right)$$

on substitution the second form of the lower bound becomes

$$\left\{-n\mathsf{E}\left[\frac{\partial^2}{\partial\theta^2}\ln\left(\frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right)\right)\right]\right\}^{-1}$$

The second order partial derivative in square brackets is

$$\frac{\partial^2}{\partial\theta^2}\ln\left(\frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right)\right) \;=\; \frac{\theta-2x}{\theta^3}$$

Cramer-Rao Inequality

The expected value of this quantity is

$$\mathsf{E}\left(\frac{\theta - 2x}{\theta^3}\right) = \frac{\theta - 2\mathsf{E}(x)}{\theta^3} = \frac{\theta - 2\theta}{\theta^3} = -\frac{1}{\theta^2}$$

Cramer-Rao Inequality

The expected value of this quantity is

$$\mathsf{E}\left(\frac{\theta - 2x}{\theta^3}\right) = \frac{\theta - 2\mathsf{E}(x)}{\theta^3} = \frac{\theta - 2\theta}{\theta^3} = -\frac{1}{\theta^2}$$

on substitution the expression for the Cramer-Rao lower bound becomes

$$\left\{-n\mathsf{E}\left[\frac{\theta - 2x}{\theta^3}\right]\right\}^{-1} = \left\{(-n)\left(-\frac{1}{\theta^2}\right)\right\}^{-1} = \frac{\theta^2}{n}$$

Cramer-Rao Inequality

Now having established that

$$\text{Var}(\hat{\theta}) = \text{Var}\left(\frac{\sum x_i}{n}\right) = \frac{\theta^2}{n}$$

together with the fact that, by the Cramer-Rao inequality *any* unbiased estimator for θ has

$$\text{Var}(\hat{\theta}) \geq \frac{\theta^2}{n}$$

we conclude that

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

is an efficient estimator for θ.

Cramer-Rao Inequality

In other words, if we consider all possible unbiased estimators for θ,

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

has the smallest possible variance.