Estimation

Assume data $y_1, \ldots, y_n$ is generated by random variables $Y_1, Y_2, \ldots, Y_n$ which have a joint distribution specified by parameter $\theta_1, \theta_2, \ldots, \theta_p$.

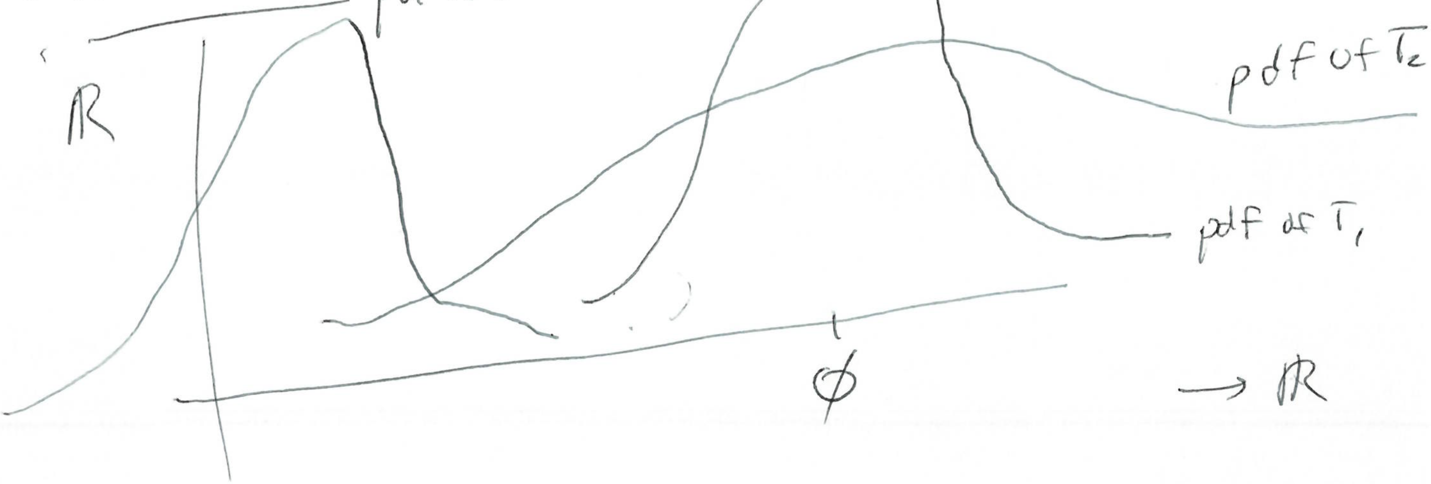
We want ~~to~~ to estimate a known function of the parameters $\phi(\theta_1, \ldots, \theta_p)$. We use a statistic $T: \mathbb{R}^n \to \mathbb{R}$ to estimate $\phi(\theta_1, \ldots, \theta_p)$, for example, $T(y_1, \ldots, y_n) = \frac{1}{n}\sum_{i=1}^{n} y_i$

The ~~random~~ random variable $T(Y_1, \ldots, Y_n)$ is an <u>estimator</u> of ϕ and $T(y_1, \ldots, y_n)$ is an <u>estimate</u> of ϕ.

We want to find good estimators of ϕ.

Therefore, we wand $E(T(Y_1, \ldots, Y_n)) \approx \phi$ and $Var(T(Y_1, \ldots, Y_n))$ to be small

Illustration:



T_1 and T_2 are unbiased for ϕ

but $Var(T_2) >> Var(T_1)$ and T_2 is a poor estimator.

~~T_3 is a~~ $Var(T_3)$ is small, but $E(T_3) << \phi$

so T_3 is a poor estimator.

T_1 is the best estimator.

Let $\underline{Y}$ denote $(Y_1, \dots, Y_n)$

Let $\underline{Y}$ ~~denote $(Y_1, \dots, Y_n)$~~.

We say $T(\underline{Y})$ is <u>unbiased</u> for ϕ

if $E(T(\underline{Y})) = \phi$ for all $\theta_1, \dots, \theta_p$

We defined the <u>bias</u> of T to be

$$\text{bias}(T(\underline{Y})) = E(T(\underline{Y})) - \phi$$

$\text{bias}(T(\underline{Y})) = 0 \ \forall \theta_1, \dots, \theta_p$ if and only if $T(\underline{Y})$ is unbiased for ϕ.

$\text{bias}(T(\underline{Y}))$ is a function of $\theta_1, \dots, \theta_p$.

<u>Example</u>

Suppose Y_i are i.i.d. Poisson(λ), $\phi(\lambda) = \lambda$

Consider two estimators:

$$T_1(\underline{Y}) = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

$$T_2(\underline{Y}) = \frac{1}{2}$$

$$\text{bias}(T_1(\underline{Y})) = E\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) - \lambda$$

$$= \frac{1}{n}\sum_{i=1}^{n} E(Y_i) - \lambda$$

$$= \frac{1}{n}\sum_{i=1}^{n} \lambda - \lambda = \frac{1}{n}(n\lambda) - \lambda$$
$$= \lambda - \lambda = 0.$$

for all λ

$T_1(\underline{Y})$ is unbiased for λ.

~~but~~ $\text{bias}(T_2(\underline{Y})) = E\left(\frac{1}{2}\right) - \lambda$
$$= \frac{1}{2} - \lambda$$

Even though $\text{bias}(T_2(\underline{Y})) = 0$ for $\lambda = 1/2$,
$T_2(\lambda)$ is biased.

We would like a positive measure of bias.

<u>Define</u>

The <u>Mean Square Error (MSE)</u> of an estimator $T(\underline{Y})$ for ϕ is
$$MSE(T(\underline{Y})) = E\left(\left(T(\underline{Y}) - \phi\right)^2\right)$$

MSE is always non-negative.

Lemma

$$MSE(T(\underline{Y})) = Var(T(\underline{Y})) + (bias(T(\underline{Y})))^2$$

*

MSE combines variance and bias.

It is a function of $\theta_1, \dots, \theta_p$.

Example

Y_i iid Poisson(λ), $T_1(\underline{Y}) = \frac{1}{n}\sum_{i=1}^{n} Y_i$, $T_2(\underline{Y}) = \frac{1}{2}$

We showed $bias(T_1(\underline{Y})) = 0$, $bias(T_2(\underline{Y})) = \frac{1}{2} - \lambda$

$$Var(T_1(\underline{Y})) = Var\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) = \frac{1}{n^2}\cancel{E}Var\left(\sum_{i=1}^{n} Y_i\right)$$

$$= \frac{1}{n^2} n\,Var(Y_i) = \frac{\lambda}{n}.$$

$$MSE(T_1(\underline{Y})) = \frac{\lambda}{n} + 0^2 = \frac{\lambda}{n}$$

$$Var(T_2(\underline{Y})) = Var\left(\frac{1}{2}\right) = 0$$

$$MSE(T_2(\underline{Y})) = 0 + \left(\frac{1}{2} - \lambda\right)^2 = \left(\frac{1}{2} - \lambda\right)^2$$

Note that $\lim_{n\to\infty} MSE(T_1(\underline{Y})) = 0$ but

$$\lim_{n\to\infty} MSE(T_2(\underline{Y})) = \left(\frac{1}{2} - \lambda\right)^2 \neq 0.$$

<u>Definition</u>

If $\lim_{n\to\infty} MSE(T(\underline{Y})) = 0$, then we say

$T(\underline{Y})$ is <u>consistent</u>.

If $T(\underline{Y})$ is consistent, then the estimates of ϕ get better and better as $n \to \infty$.

There is a lower bound on how small the variance of an unbiased estimator can be. It is called the <u>Cramér-Rao lower bound</u>.

Suppose there is only one parameter θ, so $p=1$. Given independent data $y_1, \ldots, y_n$, the <u>likelihood function</u> is

$$L(\theta; y_1, \ldots, y_n) = \prod_{i=1}^{n} f_{Y_i}(y_i)$$

where $f_{Y_i}(y)$ is the p.d.f. of Y_i if Y_i are ~~continu~~ continuous random variables and where $L(\theta; y_1, \ldots, y_n) = \prod_{i=1}^{n} P(Y_i = y_i)$

if Y_i are discrete random variables with p.m.f. $P(Y_i = y)$

The Cramér-Rao Lower Bound (CRLB) for the variance of unbiased estimators of $\phi(\theta)$ is

$$CRLB(\phi) = \frac{\left(\frac{d\phi(\theta)}{d\theta}\right)^2}{E\left(-\frac{d^2}{d\theta^2} \ln L(\theta; Y_1, \ldots, Y_n)\right)}$$

Theorem

If $T(\underline{Y})$ is an unbiased estimator for ϕ, then $Var(T(\underline{Y})) \geq CRLB(\phi)$

Corollary

If $Var(T(\underline{Y})) = CRLB(\phi)$, then $T(\underline{Y})$ has the smallest variance of all unbiased estimators of ϕ.

We call an unbiased estimator for ϕ a Minimum Variance Unbiased Estimator (MVUE) if it has the smallest possible variance.

Example

Let Y_i be iid Exponential(θ), $\phi(\theta) = \frac{1}{\theta}$

Y_i is continuous with p.d.f. $f_{Y_i}(y) = \theta e^{-\theta y}, \; y > 0$

$$L(\theta; y_1, \dots, y_n) = \prod_{i=1}^{n} \theta e^{-\theta y_i} = \theta^n e^{-\theta \sum_{i=1}^{n} y_i}$$

$$\ln L(\theta; y_1, \dots, y_n) = n \ln \theta - \theta \sum_{i=1}^{n} y_i$$

$$\frac{d^2}{d\theta^2} \ln(L(\theta; y_1, \dots, y_n)) = -\frac{n}{\theta^2}$$

$$E\left(-\frac{d^2}{d\theta^2} \ln L(\theta; Y_1, \dots, Y_n)\right) = \frac{n}{\theta^2}$$

$$\frac{d\phi(\theta)}{d\theta} = \frac{d\,\theta^{-1}}{d\theta} = -\theta^{-2}$$

$$CRLB = \frac{(-\theta^{-2})^2}{n\theta^{-2}} = \frac{1}{n\theta^2}$$

Let us show that $\frac{1}{n}\sum_{i=1} Y_i$ is MVUE.

First $E\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(Y_i)$

$= \frac{1}{n} n \frac{1}{\theta} = \frac{1}{\theta} = \phi(\theta)$

$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ is unbiased for ϕ

Second, $Var\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)$

$= \frac{1}{n^2}\sum_{i=1}^{n} Var(Y_i)$

$= \frac{1}{n^2} n \frac{1}{\theta^2} = \frac{1}{n\theta^2}$

$= CRLB(\phi)$

So $\bar{Y}$ is MVUE

Example

Let Y_i be iid Poisson (λ)

Y_i are discrete $\quad P(Y_i)\ P(Y_i = y) = \frac{e^{-\lambda}\lambda^{y}}{y!}, \quad y = 0, 1, 2, \ldots$

$$L(\lambda; y_1, \ldots y_n) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{y_i}}{y_i!} = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} y_i}}{\prod_{i=1}^{n} y_i!}$$

$$\ln L(\lambda; y_1, \ldots, y_n) = -n\lambda + \left(\sum_{i=1}^{n} y_i\right)\ln\lambda - \ln\left(\prod_{i=1}^{n} y_i!\right)$$

$$\frac{d^2}{d\lambda^2}\ln\left(L(\lambda; y_1, \ldots, y_n)\right) = -\frac{\sum_{i=1}^{n} y_i}{\lambda^2}$$

$$E\left(-\frac{d^2}{d\lambda^2}\ln L(\lambda; Y_1, Y_2, \ldots, Y_n)\right)$$

$$= E\left(\frac{\sum_{i=1}^{n} Y_i}{\lambda^2}\right) = \frac{n\lambda}{\lambda^2} = \frac{n}{\lambda}$$

Suppose $\phi(\lambda) = \lambda$.

Then $CRLB(\phi) = \dfrac{\left(\frac{d\lambda}{d\lambda}\right)^2}{n/\lambda}$

$$= \frac{(1)^2}{n/\lambda} = \frac{\lambda}{n}.$$

We will show $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ is ~~unbiased~~ MVUE for ϕ.

We know $E(\bar{Y}) = \lambda$

$$Var\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) = \frac{1}{n^2} n\lambda = \frac{\lambda}{n}$$
$$= CRLB(\lambda).$$

So $\bar{Y}$ is MVUE.

Methods of ~~Estm~~ Estimation

There are ways of coming up with estimators.
We will cover

1. Method of Moments
2. Maximum Likelihood.

Method of Moments

Def: The kth population moment of Y is $\mu_k = E(Y^k)$ $k = 1, 2, 3, \ldots$

The μ_k depend on $\theta_1, \ldots, \theta_p$.

Def The kth sample moment is

$$m_k = \frac{1}{n}\sum_{i=1}^{n} y_i^k$$

The m_k just depend on $y_1, \dots, y_k$.

Def The method of moments estimators of $\theta_1, \dots, \theta_p$ are obtained by solving the equations

$$\mu_k = m_k \quad k = 1, 2, \dots, p$$

for $\theta_1, \dots, \theta_p$.

We denote the resulting estimates ~~by~~ of $\theta_1, \dots, \theta_p$ by

$$\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_p$$

The method of moments ~~estimator~~ estimate for $\phi(\theta_1, \dots, \theta_p)$ is $\phi(\tilde{\theta}_1, \dots, \tilde{\theta}_p)$

!

Example

Let Y_i be iid. Poisson(λ) $\quad p=1$

$$\mu_1 = E(Y) = \lambda$$

$$m_1 = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}$$

Set $\mu_1 = m_1 \Rightarrow \tilde{\lambda} = \bar{y}$

estimator is $\tilde{\lambda} = \bar{Y}$

Example

Let Y_i be iid ~~Poisson~~ Exponential(λ), $p=$

$$\mu_1 = E(Y) = \frac{1}{\lambda}$$

$$m_1 = \bar{y}$$

Set $\frac{1}{\lambda} = \bar{y} \Rightarrow \tilde{\lambda} = \frac{1}{\bar{y}}$

estimator $\tilde{\lambda} = \frac{1}{\bar{Y}}$

For some distributions like Uniform$[-\theta, \theta]$ or $N(0, \sigma^2)$, μ_1 does not involve the parameter. In these cases, we must use ~~higher~~ second moments.

Example

$Y_i \sim \text{Uniform}[-\theta, \theta]$

$$\mu_1 = E(Y_i) = \frac{-\theta + \theta}{2} = 0$$ doesn't depend on θ

$$\mu_2 = E(Y_i^2) = \frac{1}{2\theta}\int_{-\theta}^{\theta} y^2\,dy = \frac{\theta^2}{3}$$

$$m_2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2$$

set $$\frac{\theta^2}{3} = \frac{1}{n}\sum_{i=1}^{n} y_i^2$$

$$\tilde{\theta} = \sqrt{\frac{3}{n}\sum_{i=1}^{n} y_i^2}$$

or, we could ~~also~~ set

$$\cancel{Var(Y) = \frac{1}{n-1}\left(m_2 - m_1^2\right)}$$

$$Var(Y) = \frac{1}{n-1}\left(m_2 - m_1^2\right)$$

$$\frac{\theta^2}{3} = m_2 - m_1^2$$

$$\tilde{\theta} = \sqrt{3(m_2 - m_1^2)}$$

Example

Y_i are iid $N(\mu, \sigma^2)$ $\theta_1 = \mu, \theta_2 = \sigma^2$

$$\mu_1 = E[Y_i] = \mu$$

$$\mu_2 = E[Y_i^2] = E[Y_i^2] - (E[Y_i])^2 + (E[Y_i])^2$$

$$= Var(Y_i) + (E[Y_i])^2$$

$$= \sigma^2 + \mu^2$$

We set $\mu_1 = m_1$, $\mu_2 = m_2$ $\Rightarrow$

$$\mu = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$\sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2$$

$$\Rightarrow \tilde{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}$$

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \bar{y}^2$$

$$= \frac{1}{n}\sum_{i=1}^{n} (y_i - \bar{y})^2$$

Usual sample variance is

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} (y_i - \bar{y})^2$$

We know $\frac{1}{n-1}\sum_{i=1}^{n} (Y_i - \bar{Y})$ is ~~unbia~~ unbiased for σ^2

so $\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (Y_i - \bar{Y})^2$ is biased for σ^2.

Method of moments estimators may be biased.