Statistical Methods MATH2715 info

Teaching material is all online!

- On Minerva `http://minerva.leeds.ac.uk`
- On GitHub https://github.com/luisacutillo78/Statistical-Methods-Lecture-Notes

Resources

- Mathematical Statistics and Data Analysis - 3rd ed. (by J. A. Rice);
- http://www1.maths.leeds.ac.uk/statistics/R/Rintro.pdf;
- https://www.datacamp.com/courses/free-introduction-to-r.

Where We've Been, Where We're Going

In the previous Lecture

- Confidence intervals
- examples and exercises at the whiteboard

Today

- Estimators
- Methods of Moments
- Maximum Likelihood
- Properties of an estimator
- Whiteboard examples

Estimators

Suppose we have $X = X_1, X_2, \ldots, X_n$ drawn from a distribution with some parameter θ

Definition

Estimators

An *estimator* $\widehat{\theta}_n$ of θ is a function of the observed data which (we hope) forms a useful approximation to the parameter:

$$\widehat{\theta}_n = g(X_1, X_2, \ldots, X_n).$$

Note that $\widehat{\theta}_n$ can depend only on the observed data, and not on any unknown parameters.

Definitions: estimator and Estimate

Estimator

Given a random sample $\underline{X} = \{X_1, \ldots, X_n\}$, the objective is to find an **Estimator** $\hat{\theta} = g(\underline{X})$ for the parameter θ of interest.

Estimate

Once we have a real observed sample $\underline{x} = \{x_1, \ldots, x_n\}$, $\hat{\theta} = g(\underline{x})$ is an **Estimate** of the parameter θ of interest.

In this lecture, we will learn two methods for point estimation of a parameter:

1: Method of Moments,

2: Method of Maximum Likelihood.

Method of Moments

kth moment

The kth moment of a probability law is defined as $\mu_k = E(X^k)$

kth sample moment

Given a iid random sample $\underline{X} = \{X_1, \ldots, X_n\}$,The kth sample moment is defined as

$$\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

Method of Moments consists in

- view $\hat{\mu}_k$ as an estimate of μ_k
- find the expression of the unknown parameters in terms of the lowest possible order moments
- substitute sample moments into the expression

Method of Moments

One Parameter Case, θ

- find the expression of the unknown parameter in terms of mean $E(X)$
- Set $E[X] = \frac{1}{n}\sum_{i=1}^{n} X_i$
- get the estimate of the parameter of interest

Whiteboard: Examples 1, 2 and 3.

Two Parameters Case, θ_1, θ_2

- find the expression of the unknown parameters in terms of $E(X)$ and $E(X^2)$
- Set $E[X] = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $E[X^2] = \frac{1}{n}\sum_{i=1}^{n} X_i^2$
- get the estimate of the parameters of interest.

Whiteboard: Examples 4, 5, 6, 7.

Maximum Likelihood

Method of Moments limitation

- The shape of the underlying distribution is not taken into account
- Alternative method: the method of **Maximum Likelihood** takes into account the probability function of the population!

The basic idea starts with the joint distribution of $X = X_1, X_2, \ldots, X_n$ depending upon a parameter θ,

$$f\left(\mathbf{x}; \theta\right) = f\left(x_1, x_2, \ldots, x_n; \theta\right).$$

For fixed θ, probability statements can be made about X. If we have observations, x, but θ is unknown, we regard information about θ as being contained in the likelihood

$$l(\theta; \mathbf{x}) = f\left(\mathbf{x}; \theta\right),$$

where l is regarded as a function of θ with $\mathbf{x}$ fixed.

Likelihood of a parameter

We define the likelihood of the parameter θ given the observed sample $\underline{x} = \{x_i\}_{i=1}^{n}$ as

$$l(\theta; \underline{x}) \quad \propto \quad \prod_{i=1}^{n} f_X(x_i; \theta).$$

That is, the likelihood of each possible parameter value θ is the probability of this value of θ given the observed sample that we got.

Maximum Likelihood estimator of a parameter

The maximum likelihood (mle) of θ is that value of θ that maximises the likelihood function, i.e. that makes the observed data most probable or likely.

Whiteboard: Example 8

Maximum Likelihood estimator of a parameter

Invariance Property

$$\hat{\theta} \quad \text{MLE for } \theta \quad \Rightarrow \quad g(\hat{\theta}) = \widehat{g(\theta)} \quad \text{MLE of } g(\theta).$$

Log Likelihood

Instead of maximising the likelihood, sometimes it is better to maximise

$$L(\theta; \underline{x}) \quad = \quad Log(l(\theta; \underline{x}))$$

which leads to the same estimate since $Log(\cdot)$ is monotonically increasing.

Recall:

- $Log(1) = 0$, $Log(e) = 1$;
- $Log(ab) = Log(a) + Log(b)$, $Log(a^b) = bLog(a)$;
- $\frac{dLog(f(x))}{dx} = \frac{f'(x)}{f(x)}$, $Log(e^{f(x)}) = f(x)$.

Whiteboard: Example 9

Unbiasedness

Another good property for an estimator $\hat{\theta} = g(\underline{X})$ of θ is

$$E[\hat{\theta}] \quad = \quad \theta.$$

That is, that if we took many random samples $\underline{x}$ and calculate the corresponding estimates $\hat{\theta}$, in average we would get the real value θ. An estimator verifying this is called an **Unbiased Estimator**.

If it is biased (Whiteboard: Example 10)

$$Bias(\hat{\theta}) \quad = \quad E[\hat{\theta}] - \theta.$$

- $Bias(\hat{\theta}) > 0$: **Positively biased**, tends to overestimate the true θ,
- $Bias(\hat{\theta}) < 0$: **Negatively biased**, tends to underestimate the true θ.

Measuring the goodness of our estimator

Mean Squared Error

For a given estimator $\hat{\theta} = g(\underline{X})$, we define its **Mean Squared Error** as

$$MSE[\hat{\theta}] \quad = \quad E[(\hat{\theta} - \theta)^2],$$

which is a measure of accuracy.

For a general estimator $\hat{\theta}$,

$$MSE[\hat{\theta}] \quad = \quad Var[\hat{\theta}] + Bias(\hat{\theta})^2,$$

so that $MSE[\hat{\theta}] = Var[\hat{\theta}]$ for unbiased estimators.

Whiteboard: Prove the equation above and Example 11.