MA 3140: Statistical Inference

Dr. Sameen Naqvi
Department of Mathematics, IIT Hyderabad
Email id: sameen@math.iith.ac.in

Logistics

- **Lecture Timings**

  Slot W: Monday (6 - 7.25 pm) and Thursday (6 - 7.25 pm).

- **Where?**

  Google classroom code **hyfrup6**.

  Meet link shared in the classroom.

- **Prerequisite:**

  MA 2110/MA4040 and MA 2140: Introduction to Probability and Introduction to Statistics

Logistics

- **Grading Scheme:** Relative Grading
  - Assignments (**30%**),
  - Mid-term Exam (**30%**), and
  - Final Exam (**40%**).

Mid-term Exam: October 29, 2020 (Thursday).

Final Exam: December 17, 2020 (Thursday).

Reference books

- Rohatgi, V.K. and Saleh, A.M.E., 2015. *An introduction to probability and statistics*. John Wiley & Sons.
- Casella, G. and Berger, R.L., 2002. *Statistical inference*. Pacific Grove, CA: Duxbury.
- Lehmann, E.L. and Casella, G., 2006. *Theory of point estimation*. Springer Science & Business Media.
- Lehmann, E.L. and Romano, J.P., 2006. *Testing statistical hypotheses*. Springer Science & Business Media.
- Kale, B.K., 2005. *A first course on parametric inference*. Alpha Science Int'l Ltd..

Course Structure

(I.) Methods of finding and evaluating estimators

(II.) Principles of Data Reduction

(III.) Testing of Hypotheses

(IV) Bayesian Estimation

Why study Statistical Inference?

Motivating examples

- The amount of rainfall in various parts of the country during monsoon.
- The number of individuals in a service queue at a ticket counter/petrol pump/theatre, and so on.
- The time taken by patients to get cured by a disease while undergoing a particular treatment.
- Deciding on the basis of clinical trials whether or not a new drug to be approved for human use of a particular disease.

What is Statistical Inference?

What is Statistical Inference?

- These problems are formulated mathematically by assuming the observed data to be values of a random variable X.

$$X \sim F(x, \theta)$$

where θ denotes some characteristic of the population.

> *The distribution of X is completely specified when the value of θ is known.*

- In SI, we make suitable statements or assertions about the unknown parameter θ based on the collected data, $X_1, \ldots, X_n$.

What is Statistical Inference?

- **Estimation:**
  - $g(\theta)$: parametric function
    An unknown feature of the population in which we are interested
  - $T(\underline{X})$: function of the random sample
    Used to estimate the unknown value of $g(\theta)$.
  - $\underline{x} = (x_1, \ldots, x_n)$: a realization of $\underline{X}$.
  - Then $T(\underline{X})$ is an **estimator** and $T(\underline{x})$ is an **estimate**.

What is Statistical Inference?

- One may either come up with an exact estimate (point estimation) or may give an interval where the unknown characteristic may lie (interval estimation).
- **PE**: Identify a statistic say $T(\underline{X})$ to estimate the parametric function $g(\theta)$.
- **IE**: Use two statistics say, $T_1(\underline{X})$ and $T_2(\underline{X})$, so that we make a statement, with certain level of confidence, regarding $g(\theta)$ lying in the interval $(T_1(\underline{X}), T_2(\underline{X}))$.

What is Statistical Inference?

- **Hypothesis testing:**
- Some information regarding the unknown feature of the population may be available.

  In the problem of HT, one would like to check whether the information is tenable in the light of the random sample drawn from the population.

- Use a statistic say $\phi(\underline{X})$ for taking a decision to accept/ reject a hypothesis. Here $\phi(\underline{X})$ is termed as a test function or test statistic.

What is Statistical Inference?

- Thus, in parametric inference, we make some assumptions about the data - for example, "the data is normally distributed", or the "population variances are equal".

- *But, what happens if the assumptions on which our methods are based don't hold?*

- In such cases, we use Nonparametric methods which require very few assumptions about the underlying distribution and can be used when the underlying distribution is unspecified.

Methods of finding estimators

Methods of finding estimators

(i) Methods of Moments

(ii) Method of Maximum Likelihood

Method of Moments

- Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with pdf or pmf $f(x; \theta_1, \ldots, \theta_k)$.
- Method of moments estimators (MME) are found by equating the first k sample moments with the corresponding k population moments, and solving the resulting system of equations.
- The k^{th} **population moment** is

$$M_k = E(X^k),\ k = 1, 2, \ldots,$$

and the corresponding k^{th} **sample moment** is

$$m_k = \frac{1}{n} \sum_{i=1}^{n} X_i^k,\ k = 1, 2, \ldots.$$

Method of Maximum Likelihood

- Suppose $X_1, X_2, \ldots, X_n$ is a random sample of size n from a population having p.d.f. (or p.m.f.) $f(x;\theta)$, where θ is an unknown parameter. Let $x_1, x_2, \ldots, x_n$ be the observed values of the random sample. Then, the likelihood function of the sample is

$$L(\theta) = f(x_1;\theta)f(x_2;\theta)\ldots f(x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta).$$

- Note that the likelihood function is a function of only the unknown parameter θ.
- The **maximum likelihood estimator** (MLE) of θ is that value of θ which maximizes the likelihood function $L(\theta)$.

Some natural questions

- Do they always give the same estimator? If not, then which one should be chosen or preferred?

- If MLE is chosen, then is it unique? How to find the unique estimator?

- Do they always exist? If not, what to do in such a situation?

Example 1

Let $X \sim \mathrm{U}(0, \theta)$, where $\theta \in \Theta = (0, \infty)$ is unknown. Check whether MME and MLE of θ are different.

Solution:

MME

Since $E(X) = \frac{\theta}{2}$, it follows that M.M.E. of θ is $2X$.

MLE

Also, for a fixed realization $x > 0$,

$$L(\theta) = \begin{cases} \frac{1}{\theta}, & \text{if } \theta > x \\ 0, & \text{if } 0 < \theta \leq x \end{cases}.$$

Clearly $L(\theta)$ is maximized at $\theta = x$. Thus the M.L.E. of θ is X.

Example 2

Let $X_1, \ldots, X_n$ be a random sample from a Poisson(λ) distribution, where $\lambda > 0$ is unknown. Find MME and MLE of λ.

Solution:

MME

Equate population moment $E(X)$ to the sample moment $\overline{X}$.

Since $E(X) = \lambda$ for Poisson, we get

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} X_i = \overline{X}$$

as the MME of λ.

Example 2 contd.

MLE

$$L(\lambda, \boldsymbol{x}) = \prod_{i=1}^{n} f(x_i, \lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda}\lambda^{\sum x_i}}{\prod_{i=1}^{n} x_i!}$$

$$l(\lambda) = \log L(\lambda, \boldsymbol{x}) = -n\lambda + \sum x_i \log \lambda - \log \left\{ \prod_{i=1}^{n} x_i! \right\}$$

$$\frac{dl}{d\lambda} = -n + \frac{\sum x_i}{\lambda} = \frac{\sum x_i - n\lambda}{\lambda} \begin{array}{ll} > 0 & \text{if } \lambda < \overline{x} \\ < 0 & \text{if } \lambda > \overline{x} \end{array}$$

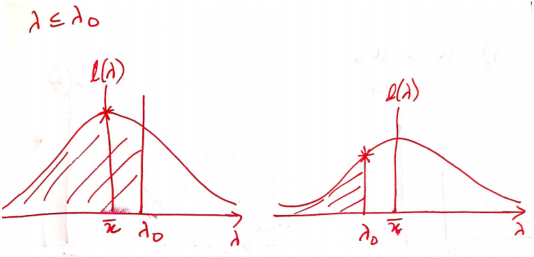
So, $\hat{\lambda} = \overline{X}$ is the MLE of λ.

Example 2 contd.

Suppose $\lambda \leq \lambda_0$

In this case, the MLE of λ is:

$$\hat{\lambda}_{RML} = \begin{cases} \overline{X}, & \text{if} \quad \overline{X} \leq \lambda_0 \\ \lambda_0, & \text{if} \quad \overline{X} > \lambda_0 \end{cases}$$
$$= \min\{\overline{X}, \lambda_0\}$$



Example 3

Let $X_1, \ldots, X_n$ be a random sample from a $N(\mu, \sigma^2)$ distribution, where σ^2 is known. Find MME and MLE of μ.

Solution:

MME

Equate population moment $E(X)$ to the sample moment $\overline{X}$.

Since $E(X) = \mu$ for Normal, we get

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

as the MME of μ.

Example 3 contd.

MLE

$$
\begin{aligned}
L(\mu, \boldsymbol{x}) &= \prod_{i=1}^{n} f(x_i, \mu) \\
&= \prod_{i=1}^{n} \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i-\mu)^2} \right] \\
&= \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\sum(x_i-\mu)^2} \\
l(\mu) = \log L(\mu, \boldsymbol{x}) &= -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum(x_i-\mu)^2 \\
\frac{dl}{d\mu} &= \sum(x_i-\mu) \Longrightarrow \hat{\mu} = \overline{X}
\end{aligned}
$$

So, $\hat{\mu} = \overline{X}$ is the MLE of μ.

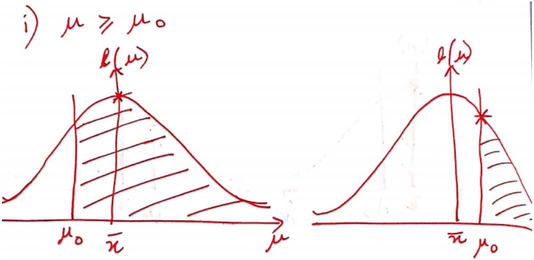
Example 3 contd.

(i) **Suppose $\mu \geq \mu_0$**

In this case, the MLE of μ is:

$$\hat{\mu}_{RML} = \begin{cases} \overline{X}, & \text{if} \quad \overline{X} \geq \mu_0 \\ \mu_0, & \text{if} \quad \overline{X} < \mu_0 \end{cases}$$

$$= \max\{\overline{X}, \mu_0\}$$

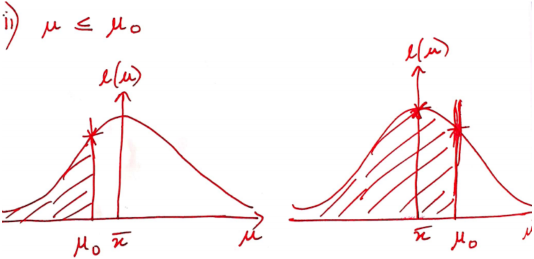


Example 3 contd.

(ii) **Suppose** $\mu \leq \mu_0$

In this case, the MLE of μ is:

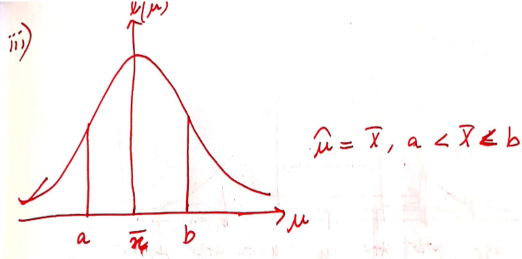
$$\hat{\mu}_{RML} = \begin{cases} \mu_0, & \text{if } \overline{X} \geq \mu_0 \\ \overline{X}, & \text{if } \overline{X} < \mu_0 \end{cases}$$
$$= \min\{\overline{X}, \mu_0\}$$



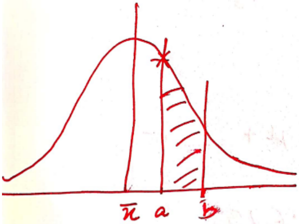
(iii) **Suppose $a \leq \mu \leq b$**

In this case, the MLE of μ is:

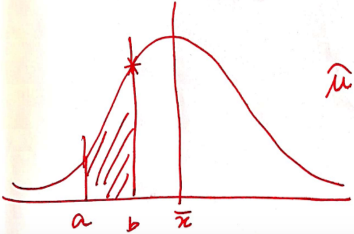
$$\hat{\mu}_{RML} = \begin{cases} a, & \text{if } \overline{X} \leq a \\ \overline{X}, & \text{if } a < \overline{X} < b \\ b, & \text{if } \overline{X} \geq b \end{cases}$$



Example 3 contd.



$\hat{\mu} = a$, if $\bar{X} \leq a$



$\hat{\mu} = b$, if $\bar{X} \geq b$

Example 4

Let $X \sim Bin(n, p)$. One observation on X is available, and it is known that n is either 2 or 3 and $p = \frac{1}{2}$ or $\frac{1}{3}$. Find the MLE for (n, p).

Solution:

The table below gives the probability that $X = x$ for each possible pair (n, p):

x	$(2, \frac{1}{2})$	$(2, \frac{1}{3})$	$(3, \frac{1}{2})$	$(3, \frac{1}{3})$	Maximum Probability
0	$\frac{1}{4}$	$\frac{4}{9}$	$\frac{1}{8}$	$\frac{8}{27}$	$\frac{4}{9}$
1	$\frac{1}{2}$	$\frac{4}{9}$	$\frac{3}{8}$	$\frac{12}{27}$	$\frac{1}{2}$
2	$\frac{1}{4}$	$\frac{1}{9}$	$\frac{3}{8}$	$\frac{6}{27}$	$\frac{3}{8}$
3	0	0	$\frac{1}{8}$	$\frac{1}{27}$	$\frac{1}{8}$

Example 4 contd.

Thus, the estimator is:

$$(\hat{n}, \hat{p})(x) = \begin{cases} (2, \frac{1}{3}), & \text{if } x = 0, \\ (2, \frac{1}{2}), & \text{if } x = 1, \\ (3, \frac{1}{2}), & \text{if } x = 2, \\ (3, \frac{1}{2}), & \text{if } x = 3. \end{cases}$$

Thanks for your patience!