Statistical Modeling - 24DS636

Abhijith M S

2025-01-23

Table of contents

1 24DS636 Statistical Modeling: 2024-25 (Even Semester)

The web page contains course contents of the course titled "Statistical Modeling" offered by Abhijith M S, PhD to Masters students pursuing M.Tech in Data Science, during the even semester of the Academic year 2024-25.

Contact: ms_abhijith@cb.amrita.edu

2 Parameter Estimation

2.1 Point Wise Estimation: Maximum Likelihood Estimators

2.2 Interval Estimates

- Consider a sample $X_1, X_2, \ldots, X_n$ drawn from a known distribution with an *unknown* mean μ.
- It is established that the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ serves as the maximum likelihood estimator for μ.
- However, the sample mean $\bar{X}$ is not expected to be exactly equal to μ, but rather close to it.
- Therefore, instead of providing a single point estimate, it is often more useful to specify **an interval** within which we are confident that μ lies.
- To determine such an interval estimator, we utilize the probability distribution of the point estimator.

2.2.1 Normal Population with unknown Mean and known Variance

- Consider a sample $X_1, X_2, \ldots, X_n$ drawn from a normal distribution with an **unknown** mean μ and a **known** variance σ^2.
- The point estimator $\bar{X}$ is normal with mean μ and variance σ^2/n.
- Therefore, $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ follows a standard normal distribution.

> **Note**
>
> Consider that I want to find an interval around $\bar{X}$ such that the actual population mean μ falls within the interval, say 95 % of the times.

> **Tip**
>
> - For finding such an interval, I can use the Z-table. From the Z-table I can find:
>
> $$P\left(-1.96 < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < 1.96\right) = 0.9750 - 0.0250 = 0.95$$
>
> - Rewriting the above equation:
>
> $$P\left(-1.96\frac{\sigma}{\sqrt{n}} < \bar{X}-\mu < 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$
> $$P\left(1.96\frac{\sigma}{\sqrt{n}} > \mu-\bar{X} > -1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$
> $$P\left(-1.96\frac{\sigma}{\sqrt{n}} < \mu-\bar{X} < 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$
> $$P\left(\bar{X}-1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X}+1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$
>
> - We have P(Z < -1.96) = 0.025, similarly P(Z > 1.96) = 0.025. Usually 1.96 is represented generally as $z_{0.025}$. Thus, P(Z < -$z_{0.025}$) = 0.025 and P(Z > $z_{0.025}$) = 0.025.
>
> - So, I can re write the equation as:
>
> $$P\left(\bar{X}-z_{0.025}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X}+z_{0.025}\frac{\sigma}{\sqrt{n}}\right) = 0.95$$
> $$P\left(\bar{X}-z_{0.025}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X}+z_{0.025}\frac{\sigma}{\sqrt{n}}\right) = 1-2*(0.025)$$

- For a confidence level of $100(1-\alpha)$ percent, the corresponding critical value from the standard normal distribution is $z_{\alpha/2}$.
- The $100(1-\alpha)$ percent confidence interval for μ is given by:

$$\left(\bar{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \tag{2.1}$$

- The interval as given in Equation 2.1 is called a **two-sided confidence interval**.

What if we are interested in one sided confidence intervals !!

Tip

- To determine such an interval, for a standard normal random variable Z, we have;

$$P\left(Z < 1.645\right) = 0.95$$

- Thus,

$$P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < 1.645\right) = 0.95$$

$$P\left(\mu - \bar{X} > -1.645\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

$$P\left(\mu > \bar{X} - 1.645\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

- Thus a **95 percent one-sided upper confidence interval for** μ is

$$\left(\bar{X} - 1.645\frac{\sigma}{\sqrt{n}}, \infty\right)$$

or in other words; **100(1-0.05) percent one-sided upper confidence interval for** μ is

$$\left(\bar{X} - z_{0.05}\frac{\sigma}{\sqrt{n}}, \infty\right)$$

Warning

Can you think of another one sided confidence interval?

Note

- We have

$$P\left(Z > -1.645\right) = 0.95$$

- Proceed just like in the previous case and you will find a **100(1-0.05) percent one-sided lower confidence interval** for μ as;

$$\left(-\infty, \bar{X} + z_{0.05}\frac{\sigma}{\sqrt{n}}\right)$$

- In general, **100(1-α) percent one-sided upper confidence interval for** μ is given in Equation 2.2.

$$\left(\bar{X} - z_{\alpha}\frac{\sigma}{\sqrt{n}}, \infty\right) \tag{2.2}$$

- Also, **100(1-α)percent one-sided lower confidence interval for** μ is given in Equation 2.3.

$$\left(-\infty, \bar{X} + z_{\alpha}\frac{\sigma}{\sqrt{n}}\right) \tag{2.3}$$

3 Summary

In summary, this book has no content whatsoever.

References