MATH 562: STATISTICAL INFERENCE

Homework 1

Rafael Barroso
Ingenierie Mathematique
École Polytechnique Fédérale de Lausanne
September 18, 2025

Contents

1 Exercise

1.1 Defining the problem

For $X_1, \ldots, X_n \sim^{i.i.d} \mathcal{N}(\mu, \sigma^2)$, derive the limiting distribution of $Y = \frac{1}{\bar{X}}$ as $n \to \infty$. Then, answer the following questions.

1. Why can the event $\bar{X}$ be neglected.
2. What does the result tell us in practice.

1.2 Solution

We start by exploring some properties of the sample mean $\bar{X}$ of a sequence of idependent, identically distributed ($i.i.d$) random variables from a population with mean μ and variance σ^2.

Theorem 1.1. Central Limit Theorem (C.L.T.): as $n \to \inf$ the distribution of $\bar{X} \approx \mathcal{N}(\mu, \sigma^2)$. More formally,

$$\sqrt{n}\left(\bar{X} - \mu\right) \to^d \mathcal{N}(0, \sigma^2)$$

where $\to^d$ denotes the convergence in distribution.

Also, in order to derive the limiting distribution, we make use of the *Delta Method*. This is a powerfull statistical tool that provides a way to find asymptotic distribution of a function $g(\bar{X})$.

Definition 1.2. Let g be our function in question. The delta method states that for a differentiable funtion with a non-zero evaluation of it's mean at it's derivative i.e. $g\prime(\mu) \neq 0$,

$$\sqrt{n}\left(g(\bar{X}) - g(\mu)\right) \to^d \mathcal{N}(0, \sigma^2 \left[g\prime(\mu)\right]^2)$$

Since our function $g(Y) = \frac{1}{X}$, we must assme that $\mu \neq 0$ in order for it to be defined at $g(\mu)$. Now, calculating the derivative of our function:

$$g\prime(x) = -\frac{1}{x^2} \Rightarrow g\prime(\mu) = -\frac{1}{\mu^2}$$

and so, $\left[g\prime(\mu)\right]^2 = \left[-\frac{1}{\mu^2}\right]^2 = \frac{1}{\mu^4}$. We may now plug this value into the delta method, and obtain the folowing result.

$$\text{Asymptotic variance: } \sigma^2 \left[g\prime(\mu)\right]^2 = \frac{\sigma^2}{\mu^4}$$

Hence, for a sample size of n, we get that the 'limiting distribution' is:

$$Y_n = \frac{1}{\bar{X}_n} \approx \mathcal{N}\left(\frac{1}{\mu}, \frac{\sigma^2}{n\mu^4}\right)$$

Why can the event $\bar{X} = 0$ be neglected?: Since $g(Y) = \frac{1}{X}$ is undefined at $X = 0$, we must consider the case when $\bar{X} = 0$. However, note that by definition $\bar{X} = \frac{1}{n}\sum X_i$ (it is a linear combination of normal random variables).

Remark 1.3. Funamental property of random variables: $\mathbb{P}(\bar{X} = c) = 0 \ \forall c \in \mathbb{C}$ i.e. there is zero probability that a random variable takes the value of a constant.

$$\therefore \ \mathbb{P}(\bar{X} = 0) = 0$$

Thus, our argument concludes.

What does the result tell us in practice? Our result gives a practical way to estimate the behaviour of the *reciprocal* of the sample mean.

- The distribution $Y = \frac{1}{X}$ is centered at $\frac{1}{\mu}$ (true value we're trying to estimate), hence, our estimator is asymptotically unbiased.
- The variance $\frac{\sigma^2}{\mu^4}$ gives us insight into reliability of our estimator:
  1. variance decreases as n increases i.e. larger samples imply better estimations
  2. variance increases as σ^2 (population variance) increases i.e. more variable data implies less precise estimates.

And so, we present the final verdict. We can estimate $\frac{1}{\mu}$ using $\frac{1}{X}$ BUT (big but) we should be extremely cautious if we believe that μ (true population mean) might be close to zero. If this is the case, our estimate might be unreliable even if we have a very large sample. e.g. if $\mu = 0.1 \Rightarrow \frac{1}{(0.1)^4} = 10000$!!! (no bueno amigo).