# ACTL2131 2.2 - Point and Interval Estimation Techniques

Tadhg Xu-Glassop

2025T1

# What is an Estimator?

Suppose you have some observations, $X_1, X_2, \ldots, X_n$ and you know they come from some distribution with parameter(s) $\theta$.
An **estimator** bridges the gap between these two - it is a formula(s) to *estimate* the parameter(s) that (typically) uses the data.

**Notation**

We denote the actual, unobservable parameter $\theta$, and we add a hat to this, $\hat{\theta}$, to indicate that this is an *estimator* for $\theta$.

# Method of Moments (MME)

Suppose we have $k$ parameters to estimate, so $\tilde{\theta} = [\theta_1, \ldots, \theta_k]^\intercal$. MME estimates all as follows:

**MME**

1. Equate the first $k$ sample moments to the first $k$ population moments, where the latter is written in terms of $\tilde{\theta}$.
2. Solve for $\tilde{\theta}$.

Recall we define the $k$th population and sample moments as $\mathbb{E}[X^k]$ and $\frac{1}{n}\sum_{i=1}^{n} X_i^k$ respectively.

# Likelihood Functions

So far, we have observed $X$ as being something *given* we know $\tilde{\theta}$ (any parameters), that is, we have spent all our time so far discussing $\mathbb{P}(\tilde{X} = \tilde{x} \mid \tilde{\Theta} = \tilde{\theta})$.

But what if it's the other way around - we have our $X$ (a sample), but we don't know $\tilde{\theta}$? Then we ought to have some way of expressing $\mathbb{P}(\tilde{\Theta} = \tilde{\theta} \mid \tilde{X} = \tilde{x})$...

This is what the likelihood function $\mathcal{L}$ is - the likelihood that our parameter values are something **given** that we have some observations!

# Likelihood Functions (cont.)

> Q: Suppose you observe 1, 2 and 3 from an exponential distribution with unknown parameter $\lambda$. What is the likelihood function?

A: $\mathcal{L}(\lambda) = \lambda e^{-\lambda} \cdot \lambda e^{-2\lambda} \cdot \lambda e^{-3\lambda} = \lambda^3 e^{-6\lambda}$.

Given this, what would you decide is the best value (estimator) for $\lambda$ ?

The most likely value, that is, the **maximum** value

# Maximum Likelihood Estimation

Again, suppose we have $k$ parameters to estimate, so $\tilde{\theta} = [\theta_1, \ldots, \theta_k]^\intercal$. MLE estimates all as follows:

**MLE**

1. Determine the likelihood function $\mathcal{L}$,
2. Determine the log-likelihood function $\ell$,
3. Use first-order condition to find critical points,
4. Check second-order condition to ensure it is a maximum.

# Confidence Intervals - Motivation

Suppose the random variable $X \sim \mathcal{N}(0, 1)$ is of great interest to me, particularly when the magnitude of it exceeds 2. What is the probability this will happen?

Using our Orange Book and considering both tails, we find it is 0.0455. So, 95.45% of the time, $X$ will fall in the interval $(-2, 2)$, i.e., I am 95.45% sure that before I even see $X$, it is in the interval $(-2, 2)$.

# Confidence Intervals - Motivation (cont.)

Thus far, we have seen estimators that gives us **one** point - but this tells us nothing about how *confident* we are in this estimate!

A lot of the time, statistics we get from data, such as sample mean $\overline{X}$ or sample variance $s^2$ follows a known distribution - so we can use the approach from the previous slide to construct a **confidence interval** - an interval where we are $x\%$ sure our parameter falls into.

# Confidence Intervals - Pivots

Such a statistic is called a **pivot** - the distribution of a pivot is free of the unknown parameter, and so we can "pivot" from our known distribution to a confidence interval of an unknown parameter.

> The point of this is that given a random sample from a distribution we don't know a parameter about, we can find an interval that the unknown parameter probably lies in.

## Steps to find a Pivot

1. What random variable am I working with? - Normal, binomial, Gamma...0
2. What am I trying to construct a CI about? - $\mu$, $\sigma^2$...
3. What information can I get from the question? - Population/sample mean, population/sample variance...
4. Looking at my formulae, CLT, methods from wk9/10, which one links these three together?

After this, draw a graph of the distribution to find the critical points of the distribution.