swissQuant
Mastering Complexity

Mathematical Challenge December 2019

Unbiased estimators

References

[1] N. Mukhopadhyay. *Probability and statistical inference.* CRC Press, 2000.

[2] G. Casella and R. L. Berger. *Statistical inference.* Cengage Learning, 2001.

Description

Motivation

The theory of point estimators (see for instance Ch.7 of [1] or Ch.7 of [2]) is one of the foundations of classical statistics. Given some model assumptions, the theory explores the estimators properties in that model, and suggests a common ground of reference and modus operandi in order to better approach the derivation of an estimator which ought to be optimal in some sense (e.g. unbiasedness). The problem that estimation theory tries to solve can seem trivial, yet it constitutes the foundation of more general and powerful frameworks, such as statistical learning and bayesian inference. Within point estimation, it is not a rare occurrence to see an apparently simple problem turn into a complex endeavor - solving this challenge is an interesting example of such occurrence.

Technical Details

A *statistical model* is a statement regarding the outcomes of an experiment. Any experiment relies on sampling, and most of the times assuming a model translates into assuming a sample joint distribution. An example would be that of a normal i.i.d. (independent and identically distributed) sample:

$$\{X_i \sim N(\mu, \sigma^2) \quad \forall i = 1, \ldots, n\}.$$

More in general, we can consider

$$\{X_i \sim f_\theta(x_i), \quad \theta \in \Theta, \quad \forall i = 1, \ldots, n\},$$

where θ, whose possible value belongs must lie within Θ, is the set of parameters that uniquely specify the density f. Within this statistical model, we assume θ to be unknown, and we call a statistic any function of the sample T(X) which does not depend on θ. We call an estimator of θ any statistic T(X) chosen in order to make inference on θ; the estimator is said unbiased if

$$\mathbb{E}_f[T(X)] = \mathbb{E}_\theta[T(X)] = \theta. \tag{1}$$

Note that this expression must hold over the entire parametric space, i.e. $\forall \theta \in \Theta$.
Before proceeding to the questions section, we here recall some distributions and their first two central moments.

Bernoulli distribution

$$X \sim Bern(p) \sim f(x) = p^x(1-p)^{1-x} \quad , \quad x \in \{0, 1\}$$
$$\mathbb{E}(X) = p \qquad \mathbb{V}(X) = p(1-p)$$

Binomial distribution

$$X \sim Bin(n, p) \sim f(x) = \binom{n}{x} p^x(1-p)^{n-x} \quad , \quad x \in \{0, n\}$$
$$\mathbb{E}(X) = np \qquad \mathbb{V}(X) = np(1-p)$$

Geometric distribution

$$X \sim Geo(p) \sim f(x) = p(1-p)^{x-1} \quad , \quad x \in \mathbb{N}$$
$$\mathbb{E}(X) = \frac{1}{p} \qquad \mathbb{V}(X) = \frac{1-p}{p^2}$$

Questions

Consider conducting an experiment to estimate the proportion p of people with blue eyes. Observations over a group of n people are collected as a series of 0 and 1. Each observation can be considered as an independent Bernoulli realization, that is $Xi \overset{i.i.d}{\sim} Bern(p)$.

- **Q1:** Using definition 1, derive an unbiased estimator of p for sample size $n = 2$. Do not use the fact that the sample mean $\Sigma X/n$ is an unbiased estimator of p. (Hint: $T(x1, x2)$ is just a number, one for each possible realization of the sample).
- **Q2:** Generalize the result of the previous question by proving that, for any sample size n, an unbiased estimator of p exists. (Hint: Compare the implicit *degrees of freedom* in $T(X)$ and the degree of the p-polynomial resulting from equation 1).
- **Q3:** Prove that an unbiased estimator for $1/p$ over $p \in (0, 1)$ does not exist.
- **Q4:** Suppose you know that p is either equal to $2/3$ or $1/3$. Is it possible to derive an unbiased estimator for $1/p$ (assume $n > 3$)?
- **Q5:** Suppose you are allowed to change the experiment design (e.g. sampling method, way data are counted, etc.). Is there a way to find an unbiased estimator of $1/p$ for any value of $p \in (0, 1)$?

We look forward to your opinions and insights.

Best Quant Regards,

swissQuant Group Leadership Team