Point Estimation

Yanrong Yang

Research School of Finance, Actuarial Studies and Statistics
The Australian National University

April 24, 2017

Content

Statistical Inference

With the knowledge of **statistic**, **sampling distribution** and **central limit theorem**, what should we do next? **Statistical Inference**!

1. **Estimation**: Point Estimation and Interval Estimation
2. **Hypothesis Test**

Point Estimator

Let $X_1, \ldots, X_n$ be sampled from a population described by a pdf or pmf $f(x|\theta)$. The parameter θ is unknown. The goal is to find a good estimator for θ.

Definition (Point Estimator)

A point estimator is any function $W(X_1, \ldots, X_n; C)$ of a sample; that is, any statistic a point estimator.

Two jobs in point estimation:

1. how to find a point estimator?
2. how to evaluate a candidate point estimator?

Example 1

Question: We have a bent coin and are interested in p, the probability of heads coming up on a single toss. How can we estimate p?

Analysis:

1. X_i denotes the ith toss result, i.e. 1 if head and 0 if tail. $i = 1, \ldots, n$.
2. $\hat{p} = \frac{1}{n} \sum_{i=1}^{n} X_i$.

Remark

Some terminology is needed to be attention.

1. p: target parameter or estimand;
2. $\hat{p}$: an estimator or a point estimator for p;
3. the value of $\hat{p}$: an estimate or a point estimate.

Criteria to Judge an Estimator

Two important concepts

1. **Bias**: The bias of an estimator $\hat{\theta}$ of θ is $B\left(\hat{\theta}\right) = \mathbb{E}\left(\hat{\theta}\right) - \theta$.
2. **Mean Square Error (MSE)**: The MSE of an estimator $\hat{\theta}$ of θ is $MSE\left(\hat{\theta}\right) = \mathbb{E}\left(\hat{\theta} - \theta\right)^2$.

Remark

Attention to these two criteria from the following points.

1. If $B\left(\hat{\theta}\right) = 0$, $\hat{\theta}$ is **unbiased** for θ.
2. $MSE\left(\hat{\theta}\right) = Var\left(\hat{\theta}\right) + B\left(\hat{\theta}\right)^2$.
3. **Bias** is used to assess accuracy while **MSE** can also measure precision.

Example 2

Question: Two numbers are randomly chosen between 0 and c. They are $x = 3.6$ and $y = 5.4$. Consider the two following estimates of c:
$u = x + y = 3.6 + 5.4 = 9.0$, $v = \max(x, y) = max(3.6, 5.4) = 5.4$.
Which estimate should we choose? Why?

Analysis:

1. $X, Y \sim i.i.d. U(0, c)$.
2. $U = X + Y$ and $V = \max(X, Y)$.
3. $\mathbb{E}(U) = \mathbb{E}(X + Y) = c$; $B(U) = \mathbb{E}(U) - c = 0$;
   $MSE(U) = Var(U) = Var(X) + Var(Y) = c^2/6$.
4. CDF for V: $F(v) = P(V < v) = P(X < v, Y < v) = P(X < v)P(Y < v) = (v/c)^2$ and then its pdf is $f(v) = F^{'}(v) = 2v/c^2$. So $\mathbb{E}V = \int_0^c v\frac{2v}{c^2}dv = \frac{2c}{3}$, $\mathbb{E}V^2 = \int_0^c v^2\frac{2v}{c^2}dv = \frac{c^2}{2}$, $B(V) = -\frac{c}{3}$,
   $MSE(V) = \frac{c^2}{6}$.

Modification for an Estimator

Conclusion: the estimator U is unbiased but has larger MSE than the estimator V.
A modified estimator $W = \frac{3V}{2}$ is unbiased since $\mathbb{E}(V) = \frac{2c}{3}$. For this new estimator W, $MSE(W) = \frac{c^2}{8}$. In view of this, W is a better estimator than V and U.

Assess Two Important Estimators

Unbiased Specific Estimators:
Suppose that $X_1, \ldots, X_n \sim i.i.d.(\mu, \sigma^2)$. Consider the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ and the sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$. Then $\mathbb{E}(\bar{X}) = \mu$ and $\mathbb{E}(S^2) = \sigma^2$.

1. $\mathbb{E}\left(\bar{X}\right) = \frac{1}{n}\sum_{i=1}^n \mathbb{E}(X_i) = \mu$.
2. $\mathbb{E}\left(S^2\right) = \frac{1}{n-1}\sum_{i=1}^n \mathbb{E}(X_i - \bar{X})^2 = \frac{n}{n-1}\mathbb{E}(X_1 - \bar{X})^2 =$ $\frac{n}{n-1}\left(Var(X_1) + Var(\bar{X}) - 2Cov(X_1, \bar{X})\right) =$ $\frac{n}{n-1}\left(\sigma^2 + \frac{\sigma^2}{n} - 2\frac{\sigma^2}{n}\right) = \sigma^2$.

Example 3

Question: A bottling machine dispenses volumes independently with mean μ and variance σ^2. $n = 3$ bottles are randomly sampled from the output of the machine, and their volumes are $1.9, 1.4, 1.8$ litres. Find unbiased estimates of μ and σ^2.

Analysis:

1. $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 1.7;$
2. $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}) = 0.07.$

Summary

1. What is a point estimator?
2. How to find a point estimator?
3. How to evaluate a point estimator? Unbiased, smaller MSE.