2.2 Properties of Point Estimators - Sufficiency and Consistency

Steven Cheung

15 February 2022

Desired Properties of Point Estimators

- unbiasedness
- efficiency (minimal variance)
- sufficiency
- consistency

Recap: Conditional Probability
$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad P(B) \neq 0$
If $P(B) > 0$, then $P(A \cap B) = P(B)P(A|B)$
$P(A) > 0$, then $P(A \cap B) = P(A)P(B|A)$

Definition:

Let $X_1, \ldots, X_n$ be a random sample from a probability distribution with unknow parameter θ. Then the statistics $U = u(X_1, \ldots, X_n)$ is said to be sufficient for θ if the conditional $p.d.f$ $f_X(X_1, \ldots, X_n | U = u(x_1, \ldots, x_n))$ (or pmf $p_X(X_1, \ldots, X_n | U = u(x_1, \ldots, x_n))$) does not depend on θ for any value of $u(x_1, \ldots, x_n)$.

An estimator $\hat{\theta}$ that is a function of a sufficient statistic for θ is said to be a sufficience estimator of θ.

Theorem (Neyman-Fischer Factorization Criteria):

Let $\hat{\theta} = u(X_1, \ldots, X_n)$ be a statistic based on the random sample $X_1, \ldots, X_n$.

$\hat{\theta}$, sufficient statistic for $\theta \Leftrightarrow$ discrete joint pmf $p_X(x_1, \ldots, x_n, \theta)$ can be factored into two non-negative functions.

$$L(\hat{\theta}) = p_X(x_1, \ldots, x_n, \theta) = g(u(x_1, \ldots, x_n), \theta) \cdot h(x_1, \ldots, x_n)$$
$$\text{for all } x_1, \ldots, x_n$$

Definition:

A sequence of random variables $X_1, \ldots, X_n$, converges in probability to a random variable X if for every $\epsilon > 0$

$$lim_{n \to \infty} P(|X_n - X| < \epsilon) = 1 \Leftrightarrow lim_{n \to \infty} P(|X_n - X| \geq \epsilon) = 0$$
$$\text{denoted as } X_n \to^p X$$

Definition:

A sequence $\hat{\theta_n} = u(X_1, \ldots, X_n)$, $n = 1, 2, 3, \ldots$ is said to be consistent sequence of estimators for θ if it converges in probability to θ
i.e for $\epsilon > 0$

$$lim_{n \to \infty} P(|\hat{\theta}_n - \theta| < \epsilon) = 1$$

Consistency means that the probability of our estimator being with some small ϵ-interval of θ can be made as close to one as we liked making sure the sample size of n sufficiently large.

Theorem (Weak Law of Large Numbers):

Let $X_1, \ldots, X_n$ be i.i.d random variables with $E(X_i) = \mu$ and $var(X_i) = \sigma^2 < \infty$. Define $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$. Then for every $\epsilon > 0$

$$lim_{n \to \infty} P(|\bar{X}_n - \mu| < \epsilon) = 1$$

that is, $\bar{X}_n$ converges in probability to μ.