Point and Interval Estimation: Some Motivation; I

In point estimation, the goal is to find an estimator $\widehat{\theta}_n$ for θ that has nice properties such as low MSE, low bias and consistency: $\Pr\big(|\widehat{\theta}_n - \theta| \le \delta\big)$ is large.

However, no matter how large n is, usually $\Pr\big(\widehat{\theta}_n = \theta\big) = 0$, e.g., when $\widehat{\theta}_n$ is a continuous rv.

Suppose instead of "hitting" θ exactly, we constructed a random set (e.g., a *confidence interval*)

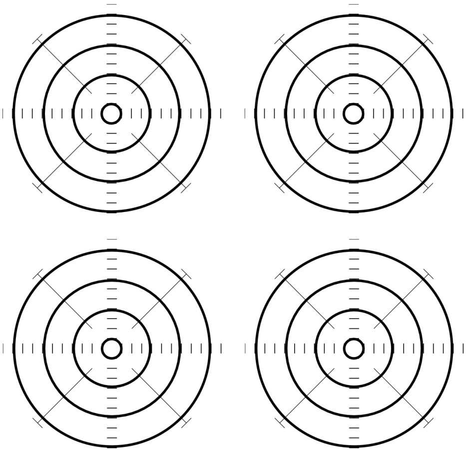
$$[L(X_1, \ldots, X_n|\alpha), U(X_1, \ldots, X_n|\alpha)]$$

such that

$$\Pr\left(\theta \in [L(X_1, \ldots, X_n|\alpha), U(X_1, \ldots, X_n|\alpha)]\right) = 1 - \alpha.$$

Here, α is some small number, e.g., 0.05 or 0.01.

Point and Interval Estimation: Some Motivation; II



Ingredients of the Confidence Intervals (CIs)

- L : lower bound, U : upper bound.

- $1-\alpha$: *confidence coefficient*
  $\equiv$ probability of coverage $\equiv$ level of the CI.

- If L and U are both finite, $[L, U]$ is called a 2-sided interval.

- If $|L|$ or U (but not both) is infinity, the CI is called one-sided. If $L = -\infty$, then the CI is called left-sided. If $U = \infty$, then the CI is called right-sided.

- Note that $[L(X_1, \ldots, X_n|\alpha), U(X_1, \ldots, X_n|\alpha)]$ is a random interval; this is an *interval estimatOR* of θ.

- When $(X_1, \ldots, X_n) = (x_1, \ldots, x_n)$,
  $[L(x_1, \ldots, x_n|\alpha), U(x_1, \ldots, x_n|\alpha)]$ is the *interval estimatE* of θ (a particular realization of the random interval estimator).

Large-sample justification of a level $(1-\alpha)$ CI

Let $X_{i,j}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$ be iid from F_θ. Let $L_i = L\left(X_{i,1}, X_{i,2}, \ldots, X_{i,n}\right)$ and $U_i = U\left(X_{i,1}, X_{i,2}, \ldots, X_{i,n}\right)$.

Write the $X_{i,j}$'s in a matrix and compute $[L_i, U_i]$ for each row:

$$\begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,n} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m,1} & X_{m,2} & \cdots & X_{m,n} \end{bmatrix}$$

Interpretation: If the "experiment" $(X_1, X_2, \ldots, X_n)$ is independently replicated m times and the CI is computed each time, then the frequency of coverage $\frac{S_m}{m}$ of θ by the random intervals $[L_1, U_1], \ldots, [L_m, U_m]$ tends to $1-\alpha$ as $m \to \infty$.