

PH 712 Probability and Statistical Inference

Part VI: Properties of Estimators II: Large-sample/asymptotic Properties

Zhiyang Zhou (zhou67@uwm.edu, zhiyanggeezhou.github.io)

2025/11/02 20:02:49

Well-known (but NOT required) identities for large samples

- Laws of large numbers (LLN): if X_1, \dots, X_n are iid with finite mean μ , then $\bar{X} \approx \mu$ as $n \rightarrow \infty$.
 - The “ \approx ” notation is abused here and, is supposed to be “ \xrightarrow{P} ” (convergence in probability):
 $\bar{X} \xrightarrow{P} \mu \Leftrightarrow$ for each $\varepsilon > 0$, $\lim_{n \rightarrow \infty} \Pr(|\bar{X} - \mu| > \varepsilon) = 0$;
 - A sufficient condition for $\bar{X} \xrightarrow{P} \mu$: as $n \rightarrow \infty$, $E(\bar{X}) \rightarrow \mu$ and $\text{var}(\bar{X}) \rightarrow 0$.

```
set.seed(1)
B = 1e4L # Number of simulations
n_values = c(10, 100, 1000)
population_mean = 50 # True population mean
population_sd = 10 # Population standard deviation
sample_means = numeric()
par(mfrow = c(1, 3)) # Set up a 1x3 plotting area
for (n in n_values) {
  for (i in 1:B) {
    # Generate a random sample and compute its mean
    sample = runif(
      n,
      min = (2*population_mean-(population_sd^2*12)^.5)/2,
      max = (2*population_mean+(population_sd^2*12)^.5)/2
    )
    sample_means[i] <- mean(sample)
  }

  # Plot the distribution of sample means
  hist(sample_means,
    main = paste("Sample size n =", n),
    xlab = "Sample Mean",
    xlim = c(population_mean - 3, population_mean + 3),
    col = "lightblue",
    border = "black",
    freq = T)
  abline(v = population_mean, col = "red", lwd = 2) # Add a line for the population mean
}
par(mfrow = c(1, 1)) # Reset plotting layout
```

- Central limit theorem (CLT): if X_1, \dots, X_n are iid with finite mean μ and finite variance σ^2 , then as

$n \rightarrow \infty$,

$$\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \approx \mathcal{N}(0, 1).$$

- A normal approximation to the distribution of \bar{X} (regardless the distribution of each X_i): $\bar{X} \approx \mathcal{N}(\mu, \sigma^2/n)$
- The “ \approx ” notation is abused too and is supposed to be “ \xrightarrow{d} ” (convergence in distribution): $\sqrt{n}(\bar{X} - \mu)/\sigma \xrightarrow{d} \mathcal{N}(0, 1)$ means that the limiting distribution of $\sqrt{n}(\bar{X} - \mu)/\sigma$ is $\mathcal{N}(0, 1)$.

```
set.seed(1)
B = 1e4L # Number of simulations
n_values = c(10, 100, 1000)
population_mean = 50 # True population mean
population_sd = 10 # Population standard deviation
std_sample_means = numeric()
par(mfrow = c(1, 3)) # Set up a 1x3 plotting area
for (n in n_values) {
  for (i in 1:B) {
    # Generate a random sample and compute its mean
    sample = runif(
      n,
      min = (2*population_mean-(population_sd^2*12)^.5)/2,
      max = (2*population_mean+(population_sd^2*12)^.5)/2
    )
    std_sample_means[i] <- n^.5*(mean(sample)-population_mean)/population_sd
  }

  # Plot the distribution of sample means
  hist(std_sample_means,
       main = paste("Sample size n =", n),
       xlab = "Standardized Sample Mean",
       xlim = c(-3, 3),
       col = "lightblue",
       border = "black",
       freq = T)
  abline(v = population_mean, col = "red", lwd = 2) # Add a line for the population mean
}
par(mfrow = c(1, 1)) # Reset plotting layout
```

Consistency/consistence

- An estimator $\hat{\theta}$ is consistent for θ if and only if $\hat{\theta} \approx \theta$ as $n \rightarrow \infty$.

Example Lec6.1

- Suppose $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$ with given μ and unknown σ^2 . Please check the consistency of the following estimators of σ^2 .
 1. $T_n = n^{-1} \sum_i (X_i - \mu)^2$
 2. $W_n = (n-1)^{-1} \sum_i (X_i - \mu)^2$

Asymptotic efficiency

- An estimator $\hat{\theta}$ is *asymptotically efficient* for θ if and only if $\sqrt{n}(\hat{\theta} - \theta) \approx \mathcal{N}(0, I_1^{-1}(\theta))$
 - Where $I_1(\theta)$ is the Fisher information for $n = 1$

- * For an iid sample, $I_1(\theta) = n^{-1}I_n(\theta)$, no longer a function of n
- Roughly speaking, when n is large enough, an asymptotically efficient $\hat{\theta}$ is following $\mathcal{N}(\theta, I_n^{-1}(\theta))$.