Geneva Graduate Institute, Econometrics I

Problem Set 1 Solutions

Francesco Casalena

Fall 2024

Problem 1

Suppose you have data on the height of n female adults living in Switzerland – $\{x_i\}_{i=1}^n$ – whereby the observations in your sample are independent. Based on that, you want to estimate the average height of female adults in the whole population (i.e. the whole of Switzerland). Let this parameter of interest be denoted by θ. You can write your observations as

$$x_i = \theta + u_i \ , \quad \text{with} \quad \mathbb{E}[u_i|\theta] = 0 \ ,$$

i.e. the height of an individual i, x_i, is given by the true average height θ plus some noise u_i around it. Note that this is just another way of writing $\mathbb{E}[x_i|\theta] = \theta$.

1. Find a point estimator for θ using the Least Squares (LS) estimation method, $\hat{\theta}$.

Solution:

The LS estimator minimizes the total sum of squares, defined as

$$\sum_{i=1}^n (x_i - \mathbb{E}[x_i|\theta])^2 = \sum_{i=1}^n (x_i - \hat{\theta})^2 \ .$$

Since this function is globally concave, we can compute the LS estimator by taking the first order condition:

$$\begin{aligned} \frac{\partial}{\partial \hat{\theta}} \sum_{i=1}^n (x_i - \hat{\theta})^2 &= -2 \sum_{i=1}^n (x_i - \hat{\theta}) \\ &= -2 \left(\sum_{i=1}^n x_i - \sum_{i=1}^n \hat{\theta} \right) \\ &= -2 \left(\sum_{i=1}^n x_i - n\hat{\theta} \right) = 0 \end{aligned}$$

We can then solve for $\hat{\theta}$ to find the LS estimator: $\hat{\theta} = \frac{1}{n} \sum\limits_{i=1}^n x_i = \bar{X}$.

2. What is the mean of $\hat{\theta}$? Is your $\hat{\theta}$ unbiased? Besides assuming $\mathbb{E}[x_i|\theta] = \theta$, is there any other assumption on the pdf of $x_i|\theta$ involved in finding this quantity? Are any assumptions regarding your sample $\{x_i\}_{i=1}^n$ involved?

Solution:

$$\begin{aligned} \mathbb{E}[\hat{\theta}|\theta] &= \mathbb{E}[\bar{X}|\theta] \\ &= \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n x_i|\theta \right] \\ &= \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^n x_i|\theta \right] \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[x_i|\theta] \\ &= \frac{1}{n} \sum_{i=1}^n \theta \\ &= \frac{1}{n} n\theta = \theta \end{aligned}$$

$\hat{\theta}$ is therefore unbiased. We do not need any additional assumptions on the distribution of X, other than knowing the mean of X. We do not need to assume anything about our sample either.

3. What is the variance of $\hat{\theta}$? Besides assuming $\mathbb{E}[x_i|\theta] = \theta$, is there any other assumption on the pdf of $x_i|\theta$ involved in finding this quantity? Are any assumptions regarding your sample $\{x_i\}_{i=1}^n$ involved?

Solution:

$$
\begin{aligned}
\mathbb{V}[\hat{\theta}|\theta] &= \mathbb{V}[\bar{X}|\theta] \\
&= \mathbb{V}\left[\frac{1}{n}\sum_{i=1}^{n} x_i|\theta\right] \\
&= \frac{1}{n^2}\mathbb{V}\left[\sum_{i=1}^{n} x_i|\theta\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{V}\left[x_i|\theta\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sigma_U^2 \\
&= \frac{1}{n^2}n\sigma_U^2 \\
&= \frac{\sigma_U^2}{n}
\end{aligned}
$$

We only need to know the first and second moments of X (i.e. θ and σ_X^2), not its whole pdf. Also, in this case, we need to assume a random sample to compute the variance of the estimator $\hat{\theta}$. We made use of this assumption in the fourth passage above, by showing that the variance of the sum of random independent variables equals the sum of the variances.

4. Suppose $u_i \sim U(-5, 5)$, i.e. u_i is distributed Uniformly between -5 and 5. Using a statistical software of your choice, write a program that, given a choice of n and θ simulates a dataset $\{x_i\}_{i=1}^n$. Fix $\theta = 175$ and $n = 10$. Compute $\hat{\theta}$ using this simulated data. Is your estimate close to the true value of $\theta = 175$? What happens under a dataset with $n = 100$ observations? What happens if you take $n = 1000$?

Solution:

```
rm(list = ls())

# Set random number seed
# (it ensures that we always get the same quasi-random results, making our analysis replicable)
set.seed(2024)

# Set sample size
n <- 10

# Set population mean
theta <- 175

# Generate n-sized random sample from Uniform distribution on [-5,5] interval
u <- runif(n, -5, 5)

# Generate random sample for x_i
```

```
x <- theta + u

# Compute sample mean
mean(x)
```

```
## [1] 175.4093
```

The sample mean is rather close to the population mean.

```
# Set sample size to 100
n <- 100

# Draw Uniform random sample
u <- runif(n,-5,5)

# Generate random sample for x_i
x <- theta + u

# Compute mean
mean(x)
```

```
## [1] 174.847
```

The sample gets closer to the population mean.

```
# Repeat everything with sample size = 1000
n <- 1000
u <- runif(n,-5,5)
x <- theta + u
mean(x)
```

```
## [1] 175.0886
```

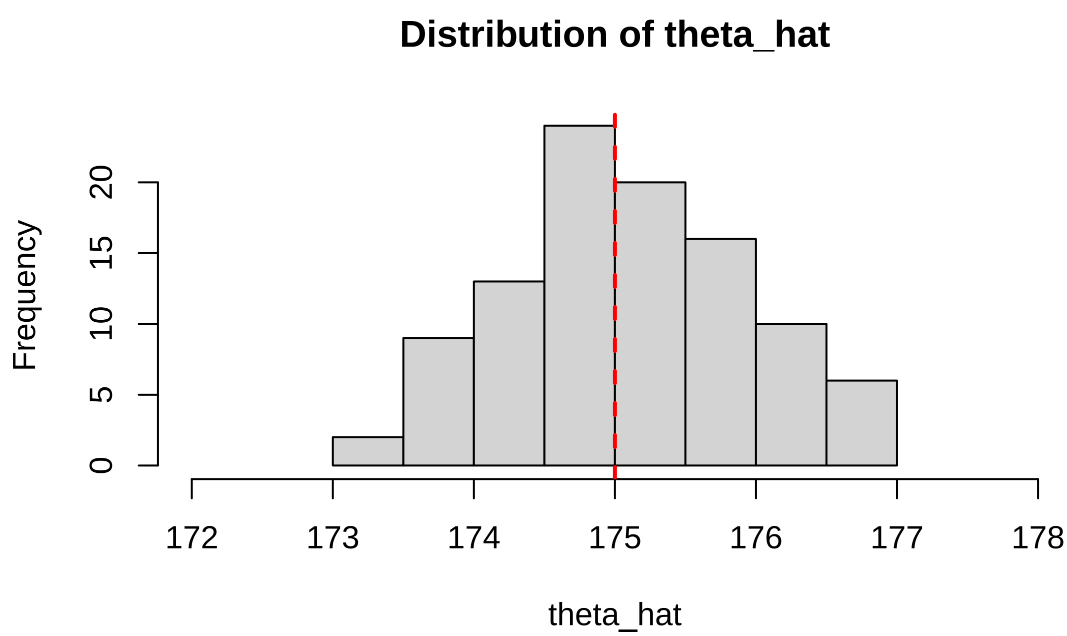
The sample mean gets even closer to the population mean. We can see that, as the sample size grows to infinity, the estimator $\hat{\theta}$ converges in probability to the population mean.

5. Now let's use the program you wrote to analyze the behavior of $\hat{\theta}$ in repeated sampling.
   (a) simulate $M = 100$ different datasets of size $n = 10$: $\{\{x_i^m\}_{i=1}^n\}_{m=1}^M$
   (b) for each dataset $\{x_i^m\}_{i=1}^n$, compue the LS-point estimator $\hat{\theta}^m$
   (c) plot a histogram of $\{\hat{\theta}^m\}_{m=1}^M$
   Comment on the histogram (distribution) of $\hat{\theta}$-values. Is it in line with your expectations, based on the calculations you did for the questions above?

Solution:

```
rm(list = ls())
set.seed(2024)

# Set the number of replications
M <- 100

# Set sample size for every replication
n <- 10

# Set population mean
```

```r
theta <- 175

# Generate empty vector of length M
# (Here we store the replications)
v <- numeric(M)

# For each replication:
for (i in 1:M) {
  # Generate random sample of x_i
  u <- runif(n,-5,5)
  x <- theta + u
  # Compute sample mean and store it
  v[i] <- mean(x)
}

hist(v, main = "Distribution of theta_hat", xlab = "theta_hat", xlim = c(172,178))
abline(v = 175, col = "red", lwd = 2, lty = 2)
```

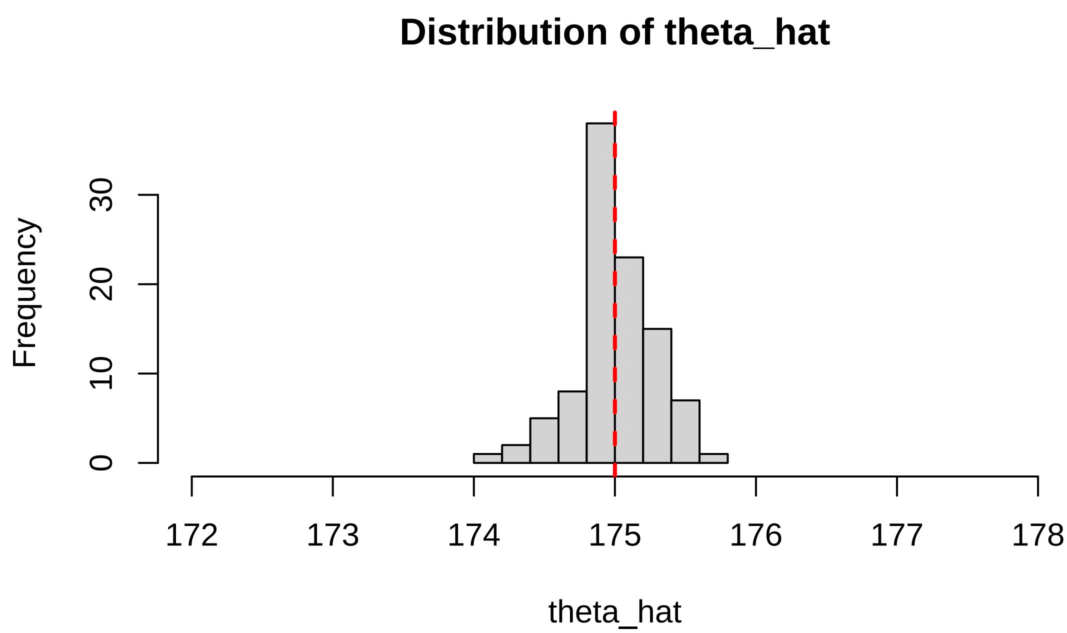


6. Redo the previous exercise using $n = 100$ as well as $n = 1000$. How does the histogram (distribution) of $\hat{\theta}$ change? Relate this to the theoretical analysis we conducted in class.

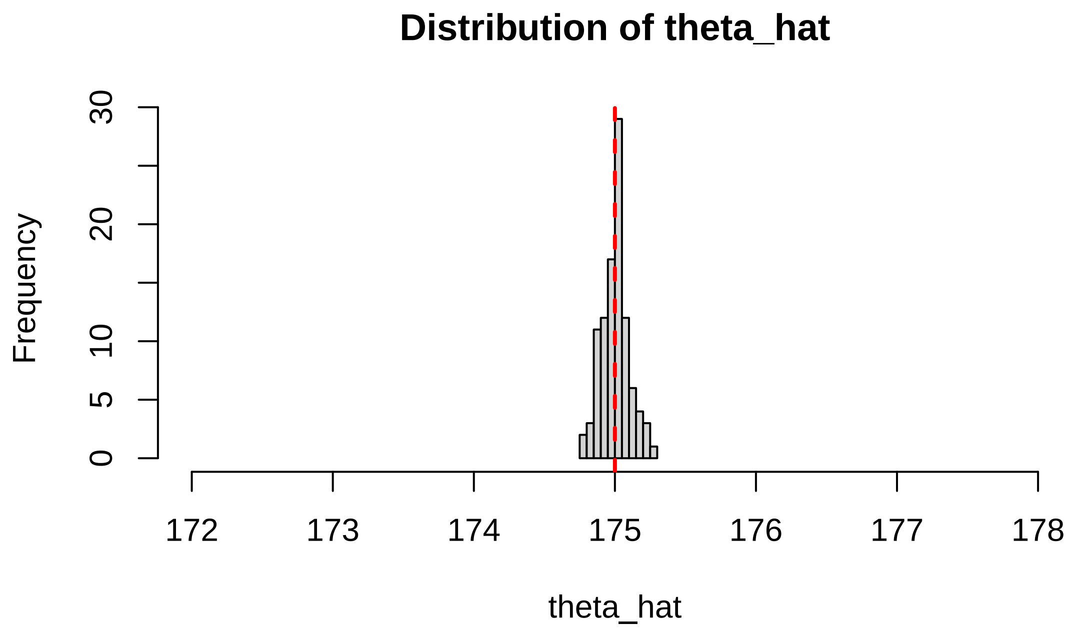
Solution:

```r
# Redo with sample size n = 100
n <- 100

for (i in 1:M) {
  u <- runif(n,-5,5)
  x <- theta + u
  v[i] <- mean(x)
}
```

```r
hist(v, main = "Distribution of theta_hat", xlab = "theta_hat", xlim = c(172,178))
abline(v = 175, col = "red", lwd = 2, lty = 2)
```




```r
# Redo with sample size n = 1000
n <- 1000

for (i in 1:M) {
  u <- runif(n,-5,5)
  x <- theta + u
  v[i] <- mean(x)
}

hist(v, main = "Distribution of theta_hat", xlab = "theta_hat", xlim = c(172,178))
abline(v = 175, col = "red", lwd = 2, lty = 2)
```



As the sample size increases, the mean of the estimator $\hat{\theta}$ gets more tightly distributed around the true mean θ. As the sample size approaches infinity, the estimator eventually collapses on a spike corresponding to the population mean.