Formative Assessment 5

Borromeo_Mayo_Mercado

2025-09-30

[*https://github.com/chieelo/STATS/tree/main/FA5_GROUP3_BORROMEO_MAYO_MERCADO*]

Elementary Sampling Theory

Problem 8.18

List all samples of size n=2 that are possible (with replacement) from the population in Problem 8.17. Use the chart wizard of EXCEL to plot the sampling distribution of the mean to show that

$$\mu_{\bar{X}} = \mu$$

and show that

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{2}$$

Population (from 8.17):

$$\{9, 12, 15\}, \quad p(x) = \frac{1}{3}$$

```r
pop <- c(9, 12, 15)

samples <- expand.grid(pop, pop)
names(samples) <- c("x1", "x2")

samples$xbar <- rowMeans(samples)

dist_table <- as.data.frame(table(xbar = samples$xbar))
dist_table$xbar <- as.numeric(as.character(dist_table$xbar))
dist_table$prob <- dist_table$Freq / nrow(samples)

mu_pop <- mean(pop)
var_pop <- sum((pop - mu_pop)^2) / length(pop)

mu_xbar <- sum(dist_table$xbar * dist_table$prob)
E_xbar2 <- sum((dist_table$xbar^2) * dist_table$prob)
var_xbar <- E_xbar2 - mu_xbar^2

theoretical_var_xbar <- var_pop / 2

print(samples)
```

```
##   x1 x2 xbar
## 1  9  9  9.0
## 2 12  9 10.5
## 3 15  9 12.0
## 4  9 12 10.5
## 5 12 12 12.0
## 6 15 12 13.5
## 7  9 15 12.0
## 8 12 15 13.5
## 9 15 15 15.0
```

```r
print(dist_table)
```

```
##   xbar Freq      prob
## 1  9.0    1 0.1111111
## 2 10.5    2 0.2222222
## 3 12.0    3 0.3333333
## 4 13.5    2 0.2222222
## 5 15.0    1 0.1111111
```

```r
cat("Population variance (sigma^2) =", var_pop, "\n")
```

```
## Population variance (sigma^2) = 6
```

```r
cat("Population mean (mu) =", mu_pop, "\n")
```

```
## Population mean (mu) = 12
```

```r
cat("Mean of sampling distribution (mu_xbar) =", mu_xbar, "\n")
```

```
## Mean of sampling distribution (mu_xbar) = 12
```

```r
cat("Variance of sampling distribution (sigma^2_xbar) =", var_xbar, "\n")
```

```
## Variance of sampling distribution (sigma^2_xbar) = 3
```

```r
cat("Theoretical sigma^2/n =", theoretical_var_xbar, "\n")
```

```
## Theoretical sigma^2/n = 3
```

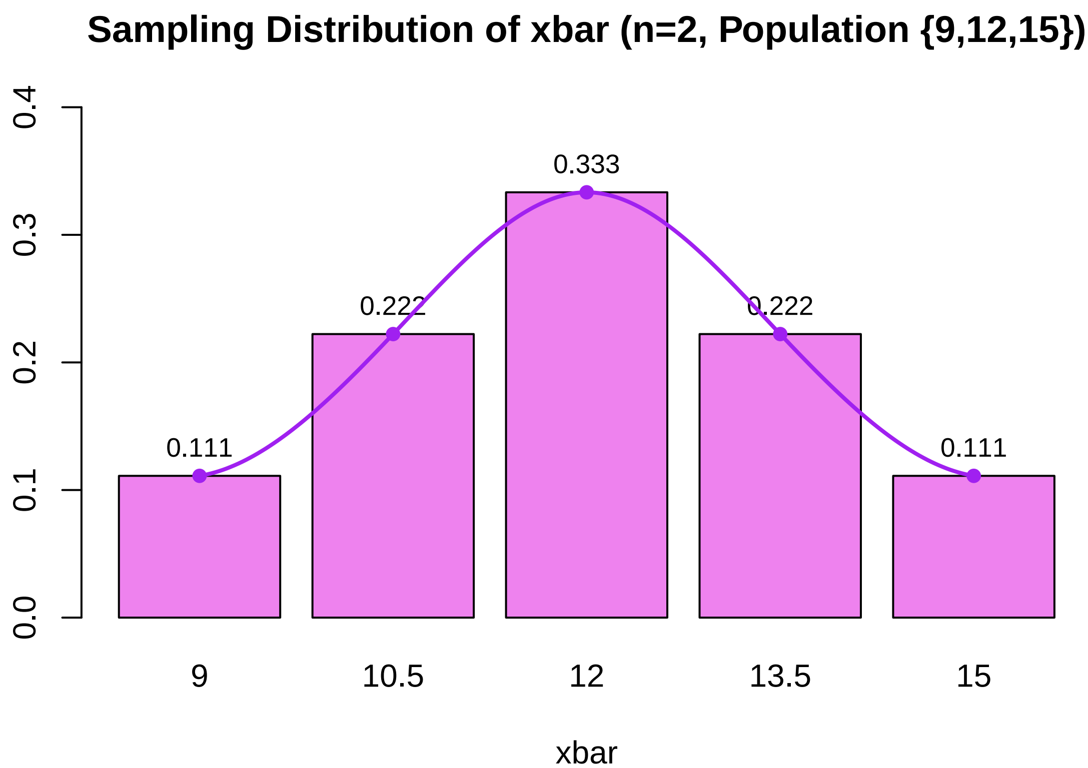
```r
mp <- barplot(height = dist_table$prob,
              names.arg = dist_table$xbar,
              main = "Sampling Distribution of xbar (n=2, Population {9,12,15})",
              xlab = "xbar",
              ylab = "",
              col = "violet",
              ylim = c(0, 0.4))

text(x = mp,
```

```r
      y = dist_table$prob,
      labels = round(dist_table$prob, 3),
      pos = 3, cex = 0.8)

smooth_line <- spline(x = mp, y = dist_table$prob, n = 200)

lines(smooth_line, col = "purple", lwd = 2)
points(mp, dist_table$prob, pch = 16, col = "purple")
```



As we can see,

$$\mu_{\bar{X}} = \mu$$
$$12 = 12$$
$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{2}$$
$$3 = 3$$

Problem 8.21

A population consists of the four numbers 3, 7, 11, and 15. Consider all possible samples of size 2 that can be drawn with replacement from this population. Find (a) the population mean, (b) the population standard deviation, (c) the mean of the sampling distribution of means, and (d) the standard deviation of the sampling distribution of means. Verify parts (c) and (d) directly from (a) and (b) by using suitable formulas.

```
pop <- c(3, 7, 11, 15)

mu_pop <- mean(pop)
var_pop <- sum((pop - mu_pop)^2) / length(pop)
sd_pop <- sqrt(var_pop)

samples <- expand.grid(pop, pop)
samples$xbar <- rowMeans(samples)

dist_table <- as.data.frame(table(xbar = samples$xbar))
dist_table$xbar <- as.numeric(as.character(dist_table$xbar))
dist_table$prob <- dist_table$Freq / nrow(samples)

mu_xbar <- sum(dist_table$xbar * dist_table$prob)

E_xbar2 <- sum((dist_table$xbar^2) * dist_table$prob)
var_xbar <- E_xbar2 - mu_xbar^2
sd_xbar <- sqrt(var_xbar)

theoretical_sd_xbar <- sqrt(var_pop / 2)

cat("Population values:", pop, "\n\n")
```

```
## Population values: 3 7 11 15
```

```
cat("(a) Population mean (mu) =", mu_pop, "\n")
```

```
## (a) Population mean (mu) = 9
```

```
cat("(b) Population standard deviation (sigma) =", sd_pop, "\n\n")
```

```
## (b) Population standard deviation (sigma) = 4.472136
```

```
cat("(c) Mean of sampling distribution (mu_xbar) =", mu_xbar, "\n")
```

```
## (c) Mean of sampling distribution (mu_xbar) = 9
```

```
cat("(d) Standard deviation of sampling distribution (sigma_xbar) =", sd_xbar, "\n\n")
```

```
## (d) Standard deviation of sampling distribution (sigma_xbar) = 3.162278
```

$$\text{From the population } \{3, 7, 11, 15\}, \text{ we found } \mu = 9, \quad \sigma = \sqrt{20} \approx 4.472.$$

$$\text{From the sampling distribution of sample means with } n = 2: \quad \mu_{\bar{X}} = 9, \quad \sigma_{\bar{X}} = \sqrt{10} \approx 3.162.$$

$$\mu_{\bar{X}} = \mu,$$

Therefore, the mean of the sampling distribution equals the population mean.

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{4.472}{\sqrt{2}} = 3.162,$$

This agrees with the computed value.

$$\therefore E[\bar{X}] = \mu, \quad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}.$$

Problem 8.34

Find the probability that of the next 200 children born, (a) less than 40% will be boys, (b) between 43% and 57% will be girls, and (c) more than 54% will be boys. Assume equal probabilities for the births of boys and girls.

```r
n <- 200
p <- 0.5
phat_mean <- p
phat_sd <- sqrt(p * (1 - p) / n)

z_a <- (0.40 - phat_mean) / phat_sd
p_a <- pnorm(z_a)

z1_b <- (0.43 - phat_mean) / phat_sd
z2_b <- (0.57 - phat_mean) / phat_sd
p_b <- pnorm(z2_b) - pnorm(z1_b)

z_c <- (0.54 - phat_mean) / phat_sd
p_c <- 1 - pnorm(z_c)

cat("(a) =", p_a, "\n")
```

```
## (a) = 0.002338867
```

```r
cat("(b) =", p_b, "\n")
```

```
## (b) = 0.9522851
```

```r
cat("(c) =", p_c, "\n")
```

```
## (c) = 0.1289495
```

$$(a)\ P(\hat{p} < 0.40) \approx 0.00234$$

$$(b)\ P(0.43 \le \hat{p} \le 0.57) \approx 0.95229$$

$$(c)\ P(\hat{p} > 0.54) \approx 0.12895$$

—

Problem 8.49

The credit hour distribution at Metropolitan Technological College is as follows:

x	6	9	12	15	18
p(x)	0.1	0.2	0.4	0.2	0.1

Find mu and sigma^2. Give the 25 (with replacement) possible samples of size 2, their means, and their probabilities.

```r
x <- c(6, 9, 12, 15, 18)
px <- c(0.1, 0.2, 0.4, 0.2, 0.1)

mu <- sum(x * px)
E_x2 <- sum(x^2 * px)
var_pop <- E_x2 - mu^2

cat("Population mean =", mu, "\n")
```

```
## Population mean = 12
```

```r
cat("Population variance =", var_pop, "\n\n")
```

```
## Population variance = 10.8
```

```r
samples <- expand.grid(x, x)
colnames(samples) <- c("x1", "x2")
samples$mean <- rowMeans(samples)
samples$prob <- px[match(samples$x1, x)] * px[match(samples$x2, x)]
print(samples)
```

```
##    x1 x2 mean prob
## 1   6  6  6.0 0.01
## 2   9  6  7.5 0.02
## 3  12  6  9.0 0.04
## 4  15  6 10.5 0.02
## 5  18  6 12.0 0.01
## 6   6  9  7.5 0.02
## 7   9  9  9.0 0.04
```

```
## 8   12  9 10.5 0.08
## 9   15  9 12.0 0.04
## 10  18  9 13.5 0.02
## 11   6 12  9.0 0.04
## 12   9 12 10.5 0.08
## 13  12 12 12.0 0.16
## 14  15 12 13.5 0.08
## 15  18 12 15.0 0.04
## 16   6 15 10.5 0.02
## 17   9 15 12.0 0.04
## 18  12 15 13.5 0.08
## 19  15 15 15.0 0.04
## 20  18 15 16.5 0.02
## 21   6 18 12.0 0.01
## 22   9 18 13.5 0.02
## 23  12 18 15.0 0.04
## 24  15 18 16.5 0.02
## 25  18 18 18.0 0.01
```

Therefore,

$$\mu = \sum xp(x) = 12$$
$$\sigma^2 = \sum x^2 p(x) - \mu^2 = 10.8$$

And all 25 (with replacement) possible samples can easily be seen above.