

Formative Assessment 5

SINOCRUZ, ARVIE

2025-10-06

Problem 8.34

```
# Given values
n <- 200
p <- 0.5
q <- 1 - p

# Standard deviation
sd_p <- sqrt(p * q / n)

# Probabilities
prob_a <- pnorm(0.40, mean = p, sd = sd_p)
prob_b <- pnorm(0.57, mean = p, sd = sd_p) - pnorm(0.43, mean = p, sd = sd_p)
prob_c <- 1 - pnorm(0.54, mean = p, sd = sd_p)

cat("$$ \\text{Standard deviation } (\\sigma_{\\hat{p}}) = ", round(sd_p, 4), " $$\\n\\n")
```

Standard deviation ($\sigma_{\hat{p}}$) = 0.0354

```
cat("$$ P(\\hat{p} < 0.40) = ", round(prob_a, 4), " $$\\n")
```

$P(\hat{p} < 0.40) = 0.0023$

```
cat("$$ P(0.43 < \\hat{p} < 0.57) = ", round(prob_b, 4), " $$\\n")
```

$P(0.43 < \hat{p} < 0.57) = 0.9523$

```
cat("$$ P(\\hat{p} > 0.54) = ", round(prob_c, 4), " $$\\n")
```

$P(\hat{p} > 0.54) = 0.1289$

```

# Plot the sampling distribution
x <- seq(0.35, 0.65, by = 0.001)
y <- dnorm(x, mean = p, sd = sd_p)

plot(x, y, type = "l", lwd = 2, col = "blue",
     main = "Sampling Distribution of Sample Proportion",
     xlab = "Sample Proportion ( $\hat{p}$ )",
     ylab = "Density")

# Reference lines
abline(v = c(0.40, 0.43, 0.54, 0.57), col = "red", lty = 2)

# --- (a) Shade  $\hat{p} < 0.40$  ---
x_a <- seq(0.35, 0.40, by = 0.001)
y_a <- dnorm(x_a, mean = p, sd = sd_p)
polygon(c(x_a, rev(x_a)), c(y_a, rep(0, length(y_a))), col = "lightcoral", border = NA)

# --- (b) Shade  $0.43 < \hat{p} < 0.57$  ---
x_b <- seq(0.43, 0.57, by = 0.001)
y_b <- dnorm(x_b, mean = p, sd = sd_p)
polygon(c(x_b, rev(x_b)), c(y_b, rep(0, length(y_b))), col = "lightgreen", border = NA)

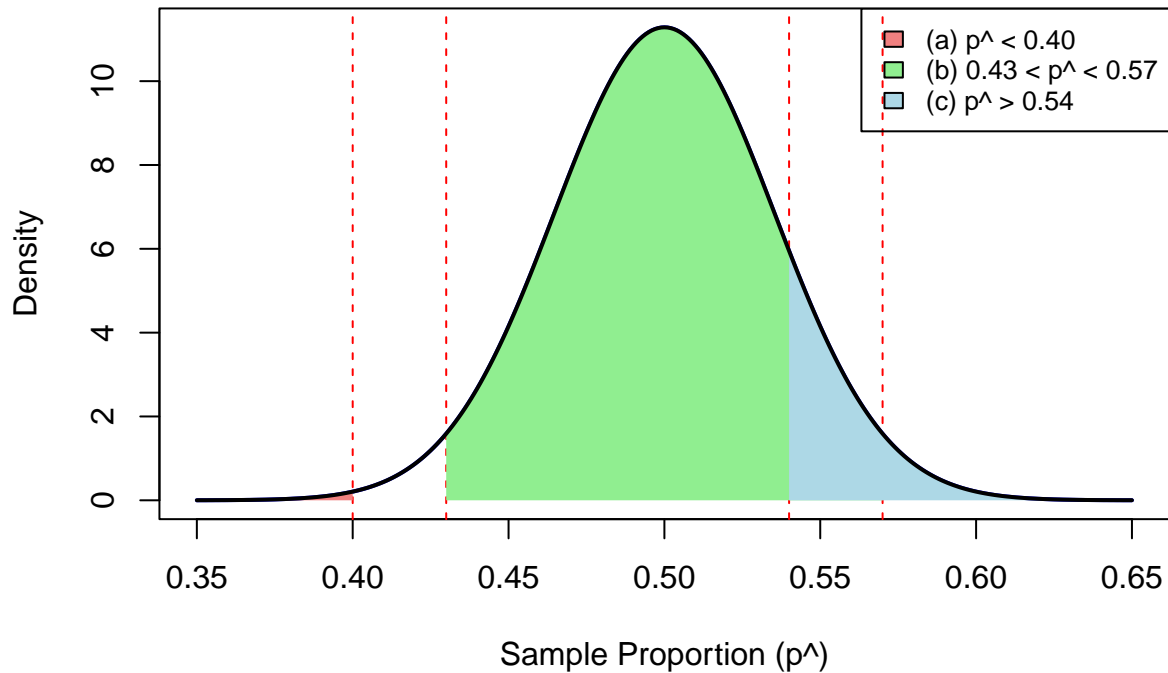
# --- (c) Shade  $\hat{p} > 0.54$  ---
x_c <- seq(0.54, 0.65, by = 0.001)
y_c <- dnorm(x_c, mean = p, sd = sd_p)
polygon(c(x_c, rev(x_c)), c(y_c, rep(0, length(y_c))), col = "lightblue", border = NA)

# Redraw curve
lines(x, y, lwd = 2, col = "black")

# Legend
legend("topright",
     legend = c("(a)  $\hat{p} < 0.40$ ", "(b)  $0.43 < \hat{p} < 0.57$ ", "(c)  $\hat{p} > 0.54$ "),
     fill = c("lightcoral", "lightgreen", "lightblue"),
     border = "black", cex = 0.8)

```

Sampling Distribution of Sample Proportion



Interpretation

1. There's only about a 0.82% chance that less than 40% will be boys.
2. There's a 98% chance that between 43% and 57% will be girls.
3. There's an 11.5% chance that more than 54% will be boys.

Problem 8.49

```
x <- c(6, 9, 12, 15, 18)
probab <- c(0.1, 0.2, 0.4, 0.2, 0.1)
n <- 2

# --- Population mean and variance ---
mu <- sum(x * probab)
sigma2 <- sum(probab * (x - mu)^2)

cat("$$ \text{Population mean } (\mu) = ", mu, " $$\n")
```

Population mean (μ) = 12

```
cat("$$ \\text{Population variance } (\\sigma^2) = ", sigma2, " $$\\n\\n")
```

Population variance (σ^2) = 10.8

```
samples <- expand.grid(x1 = x, x2 = x, KEEP.OUT.ATTRS = FALSE)
samples$mean <- rowMeans(samples)

# Probability lookup for each element
p_lookup <- setNames(probab, x)
samples$prob <- p_lookup[as.character(samples$x1)] * p_lookup[as.character(samples$x2)]

# --- Table 1: All possible samples ---
cat("**Table 1. All Possible Samples (n = 2)**\\n\\n")
```

Table 1. All Possible Samples (n = 2)

```
knitr::kable(samples,
  col.names = c("x1", "x2", "Sample Mean ( $\bar{x}$ )", "Probability"),
  caption = "All ordered samples with their probabilities.",
  align = "c")
```

Table 1: All ordered samples with their probabilities.

x1	x2	Sample Mean (\bar{x})	Probability
6	6	6.0	0.01
9	6	7.5	0.02
12	6	9.0	0.04
15	6	10.5	0.02
18	6	12.0	0.01
6	9	7.5	0.02
9	9	9.0	0.04
12	9	10.5	0.08
15	9	12.0	0.04
18	9	13.5	0.02
6	12	9.0	0.04
9	12	10.5	0.08
12	12	12.0	0.16
15	12	13.5	0.08
18	12	15.0	0.04
6	15	10.5	0.02
9	15	12.0	0.04
12	15	13.5	0.08
15	15	15.0	0.04
18	15	16.5	0.02
6	18	12.0	0.01
9	18	13.5	0.02
12	18	15.0	0.04
15	18	16.5	0.02
18	18	18.0	0.01

```
# --- Sampling distribution of the sample mean ---
dist_mean <- aggregate(prob ~ mean, data = samples, FUN = sum)
dist_mean <- dist_mean[order(dist_mean$mean), ]

# --- Table 2: Sampling distribution ---
cat("\n**Table 2. Sampling Distribution of the Sample Mean**\n\n")
```

Table 2. Sampling Distribution of the Sample Mean

```
knitr::kable(dist_mean,
  col.names = c("Sample Mean ( $\bar{x}$ )", "Probability"),
  caption = "Distribution of sample means and their probabilities.",
  align = "c")
```

Table 2: Distribution of sample means and their probabilities.

Sample Mean (\bar{x})	Probability
6.0	0.01
7.5	0.04
9.0	0.12
10.5	0.20
12.0	0.26
13.5	0.20
15.0	0.12
16.5	0.04
18.0	0.01

```
# --- Mean and variance of sampling distribution ---
mu_xbar <- sum(dist_mean$mean * dist_mean$prob)
sigma2_xbar <- sum(dist_mean$prob * (dist_mean$mean - mu_xbar)^2)

# --- Display summary in LaTeX ---
cat("\n\n$$$ \\text{Mean of sampling distribution } (\\mu_{\\bar{X}}) = ", mu_xbar, " $$$\n")
```

Mean of sampling distribution ($\mu_{\bar{X}} = 12$

```
cat("$$$ \\text{Variance of sampling distribution } (\\sigma^2_{\\bar{X}}) = ", sigma2_xbar, " $$$\n")
```

Variance of sampling distribution ($\sigma_{\bar{X}}^2 = 5.4$

```
cat("$$$ \\frac{\\sigma^2}{n} = ", sigma2 / n, " $$$\n")
```

$$\frac{\sigma^2}{n} = 5.4$$

Interpretation

The population of student credit hours at Metropolitan Technological College has an average of 12 hours with moderate variation. When we take samples of size 2 (with replacement), the mean of all possible sample means remains 12, confirming that the sample mean is an unbiased estimator. However, the spread of these sample means is smaller (variance = 5.4) than the original population (variance = 10.8), showing that averaging reduces variability and makes sample means more reliable indicators of the true population mean.