Curtin University, Semester 2, 2026

ECOM5005

Tutorial 6 Detailed Solutions

Sampling Distributions - classroom-ready worked solutions

Learning Outcomes

1. Interpret the concept of the sampling distribution.
2. Calculate probabilities related to the sample mean and the sample proportion.
3. Recognise the importance of the Central Limit Theorem.

PART 1 - Detailed worked solutions

Question 1. Means in quality control

Question

An auto-maker tests paint thickness on a part. The population distribution is right-skewed with $\mu = 2$ mm and $\sigma = 0.5$ mm. A random sample of $n = 100$ points is taken. Find the probability that the sample mean is within 0.1 mm of the target value 2 mm, working through the normality, mean, standard error and probability steps.

Detailed Solution

Step 1: Establish the shape of the sampling distribution. The population is right-skewed, but $n = 100 \geq 30$. By the Central Limit Theorem, the sampling distribution of $\bar{x}$ is approximately normal. Therefore, answer (1) is B: Approximately normal.

Step 2: Find the mean of the sampling distribution. The sample mean is an unbiased estimator of the population mean, so $\mu_{\bar{x}} = \mu = 2$ mm.

Step 3: Find the standard error. $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 0.5/\sqrt{100} = 0.05$ mm.

Step 4: Translate "within 0.1 mm" into an interval: $1.9 < \bar{x} < 2.1$.

Step 5: Standardise the two limits. For 1.9, $Z = (1.9 - 2)/0.05 = -2$. For 2.1, $Z = (2.1 - 2)/0.05 = +2$.

Step 6: Find the probability. $P(1.9 < \bar{x} < 2.1) = P(-2 < Z < 2) = 0.9545$, or about 95.45%. Hence the multiple-choice answer is C: 95%.

Excel check: =NORM.DIST(2.1,2,0.05,TRUE)-NORM.DIST(1.9,2,0.05,TRUE) → 0.9544997.

Interpretation

If the process really has mean thickness 2 mm and standard deviation 0.5 mm, then about 95% of random samples of 100 points will have an average thickness between 1.9 mm and 2.1 mm. The sample average is much more stable than an individual thickness measurement because averaging reduces variability.

Question 2. Proportions in polling results

Question

The population proportion of Americans aged over 25 with a high school diploma is $\pi = 0.87$. A random sample of $n = 200$ is selected. Find $P(\hat{p} < 0.85)$, after checking the normal approximation and finding the mean and standard error of $\hat{p}$.

Detailed Solution

Step 1: Check the expected counts. $n\pi = 200(0.87) = 174$ and $n(1 - \pi) = 200(0.13) = 26$. Both are at least 5, so the normal approximation is appropriate.

Step 2: Mean of the sampling distribution: $\mu_{\hat{p}} = \pi = 0.87$.

Step 3: Standard error: $\sigma_{\hat{p}} = \sqrt{[\pi(1-\pi)/n]} = \sqrt{[0.87(0.13)/200]} = 0.023780 \approx 0.024$.

Step 4: Standardise 0.85. $Z = (0.85 - 0.87)/0.023780 = -0.8410$.

Step 5: $P(\hat{p} < 0.85) = P(Z < -0.8410) = 0.20016 \approx 0.20$. Therefore the correct multiple-choice answer is B.

Excel check using the unrounded standard error: =NORM.DIST(0.85,0.87,SQRT(0.87*(1-0.87)/200),TRUE) → 0.200164.

Interpretation

Even though the true population proportion is 87%, random sampling variation means that about 20% of samples of 200 people would produce a sample proportion below 85%.

Extension problems - detailed solutions

Extension Problem A - Airline turnaround-time reliability

Detailed Solution

Given $\mu = 42$ minutes, $\sigma = 15$ minutes and $n = 64$.

Step 1: Normality. The underlying distribution is strongly right-skewed, but $n = 64 \geq 30$. The Central Limit Theorem therefore makes the sampling distribution of $\bar{x}$ approximately normal.

Step 2: $\mu_{\bar{x}} = 42$ and $\sigma_{\bar{x}} = 15/\sqrt{64} = 15/8 = 1.875$ minutes.

Step 3: $P(\bar{x} > 45)$. $Z = (45 - 42)/1.875 = 1.60$. Hence $P(\bar{x} > 45) = 1 - \Phi(1.60) = 0.0548$.

Step 4: $P(40 < \bar{x} < 44)$. The Z limits are $(40-42)/1.875 = -1.0667$ and $(44-42)/1.875 = 1.0667$. Thus the probability is approximately 0.7139.

Excel: =1-NORM.DIST(45,42,1.875,TRUE) → 0.0548; and =NORM.DIST(44,42,1.875,TRUE)-NORM.DIST(40,42,1.875,TRUE) → 0.7139.

Business interpretation

If the process is stable at the stated population mean and standard deviation, only about 5.5% of weekly samples of 64 flights would have an average turnaround time above 45 minutes. A week above 45 minutes is therefore relatively unusual and may justify operational investigation.

Extension Problem B - Streaming-platform conversion-rate monitoring

Detailed Solution

Given $\pi = 0.36$ and $n = 400$.

Step 1: Normality check. $n\pi = 144$ and $n(1-\pi) = 256$; both are much greater than 5.

Step 2: $\mu_{\hat{p}} = 0.36$ and $\sigma_{\hat{p}} = \sqrt{[0.36(0.64)/400]} = 0.024$.

Step 3: $P(\hat{p} < 0.33)$. $Z = (0.33 - 0.36)/0.024 = -1.25$, so $P = 0.1056$.

Step 4: $P(0.34 < \hat{p} < 0.39)$. The Z limits are -0.8333 and 1.25. The probability is 0.6920.

Excel: =NORM.DIST(0.33,0.36,0.024,TRUE) → 0.1056; and =NORM.DIST(0.39,0.36,0.024,TRUE)-NORM.DIST(0.34,0.36,0.024,TRUE) → 0.6920.

Business interpretation

A weekly conversion rate below 33% is not impossible even when the true rate is 36%; it occurs about 10.6% of the time from sampling variation alone. Analysts should therefore avoid treating every low weekly sample proportion as evidence that the underlying conversion rate has changed.

Extension Problem C - Hotel check-in times and the effect of sample size

Detailed Solution

Given $\mu = 7.5$ minutes and $\sigma = 4.8$ minutes. The population is strongly right-skewed.

Step 1: Normal approximation. For $n = 16$, the lecture's general $n \geq 30$ rule is not met, so a normal approximation is not automatically justified. For $n = 64$, the Central Limit Theorem gives an approximately normal sampling distribution.

Step 2: Standard errors. Manager A: $\sigma_{\bar{x}} = 4.8/\sqrt{16} = 1.2$ minutes. Manager B: $\sigma_{\bar{x}} = 4.8/\sqrt{64} = 0.6$ minutes.

Step 3: For Manager B, $P(6.5 < \bar{x} < 8.5)$. The Z limits are $(6.5-7.5)/0.6 = -1.6667$ and $(8.5-7.5)/0.6 = 1.6667$. Therefore $P \approx 0.9044$.

Step 4: Increasing n from 16 to 64 divides the standard error by 2. Larger samples therefore produce sample means that cluster more tightly around μ and provide more precise estimates.

PART 2 - Data exercise: Restaurant_Times.xlsx

What the supplied data contain

The workbook contains 100 observations and three variables: Waiting Time, Serving Time, and Complimentary Drink (Yes = 1, No = 0). The worked values below were independently checked against the uploaded workbook and the supplied Excel solution.

Step-by-step Excel instructions

Step 1 - Open and inspect the data

Open Restaurant_Times.xlsx and select the worksheet named "Restaurant Times". Row 1 contains the three variable names. The 100 observations are in rows 2 to 101.

Step 2 - Turn on the Analysis ToolPak if "Data Analysis" is missing

Windows: File > Options > Add-ins > Manage Excel Add-ins > Go > tick Analysis ToolPak > OK. Mac: Tools > Excel Add-ins > tick Analysis ToolPak > OK. Then return to the Data tab.

Step 3 - Descriptive Statistics

Go to Data > Data Analysis > Descriptive Statistics. Input Range: A1:C101. Choose Grouped By: Columns, tick Labels in first row, choose an Output Range or New Worksheet Ply, and tick Summary statistics. Click OK.

Step 4 - Read the output

For Waiting Time and Serving Time, the mean and standard deviation are in minutes. For Complimentary Drink, the mean is a proportion because the variable is coded 1 for Yes and 0 for No. A mean of 0.34 therefore means 34% of restaurants provided a complimentary drink.

Step 5 - Create a waiting-time histogram

A simple approach is to use bins 0-<5, 5-<10, 10-<15, 15-<20, 20-<25 and 25-30. Use COUNTIFS to count observations in each bin, then Insert > Column Chart. Reduce Gap Width if you want it to look more like a histogram.

Step 6 - Generate Sample 1 of size 30 with replacement

In a blank cell (for example T32), enter =INDEX($A$2:$A$101,RANDBETWEEN(1,100)). Fill the formula down through T61 so that you have 30 sampled waiting times. Because RANDBETWEEN recalculates, immediately copy T32:T61 and use Paste Special > Values to freeze the sample.

Step 7 - Calculate Sample 1 mean

In T30 (or another labelled cell), enter =AVERAGE(T32:T61). Label the column “Sample 1”.

Step 8 - Repeat for Samples 2 to 7

Repeat the same procedure in six new columns. Each sample must contain 30 waiting times. Freeze each sample with Paste Special > Values before moving on.

Step 9 - Compare the sample means

Calculate the overall mean of all 100 Waiting Time observations using =AVERAGE(A2:A101). Then compare every sample mean with this overall mean.

Step 10 - Mean and standard deviation of the sample means

If the seven sample means are in T30:Z30, calculate =AVERAGE(T30:Z30) and =STDEV.S(T30:Z30). The average of repeated sample means should tend toward the overall mean as the number of repeated samples becomes large.

Step 11 - Estimate the theoretical standard error

Use =STDEV.S(A2:A101)/SQRT(30). This gives an estimate of the standard deviation of the sampling distribution of $\bar{x}$ under the large-population/with-replacement framework used in the lecture.

Step 12 - Create the sample-means comparison chart

Create a small table with Sample 1 to Sample 7 and their means, plus a final row labelled “Overall mean” with 6.10. Select the table and choose Insert > Column Chart. The bars should cluster much more tightly than the individual waiting-time observations.

Final descriptive-statistics results

Statistic	Waiting Time	Serving Time	Complimentary Drink
Mean	6.100	60.000	0.340
Standard error	0.565	1.337	0.0476
Median	3.850	60.000	0
Mode(s)	2.9, 3.7	43.5, 59.4, 59.8, 65.3, 65.6, 75.8	0
Standard deviation	5.651	13.370	0.4761
Sample variance	31.932	178.749	0.2267
Skewness	1.589	-0.048	0.686
Kurtosis (excess)	2.736	0.021	-1.561
Range	27.700	72.500	1
Minimum	0.300	22.500	0
Maximum	28.000	95.000	1
Sum	610.000	6000.000	34
Count	100	100	100

Interpretation of the descriptive statistics

Waiting Time is strongly right-skewed (skewness ≈ 1.589) with a long upper tail and a maximum of 28 minutes. Serving Time is almost symmetric (skewness ≈ -0.048). The complimentary-drink mean of 0.34 means 34 of the 100 restaurants provided a complimentary drink.

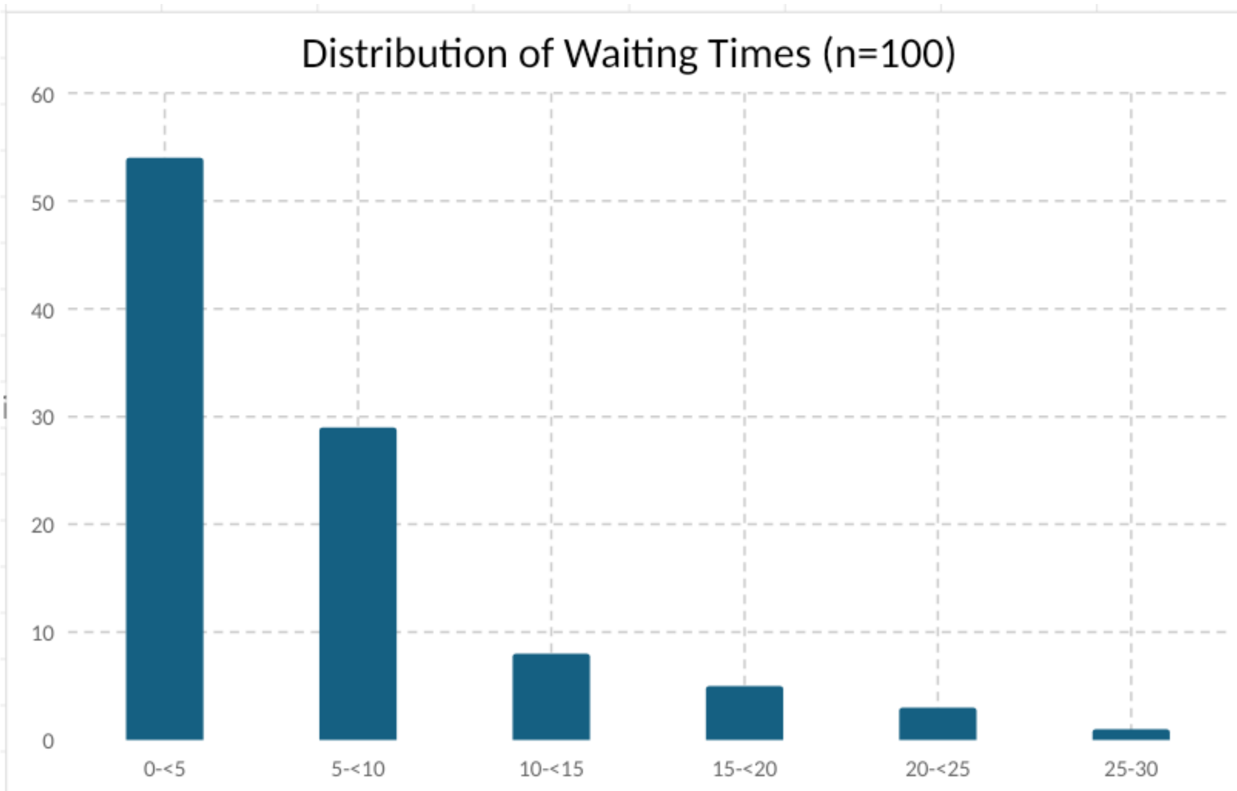

Figure 1: Waiting times are strongly right-skewed: 54 of the 100 observations are below 5 minutes, while a small number extend above 20 minutes

Worked repeated-sampling results from the supplied Excel solution

Sample	Mean waiting time (min)	Difference from 6.10
Sample 1	4.3500	-1.7500
Sample 2	6.3833	+0.2833
Sample 3	5.7167	-0.3833
Sample 4	6.2033	+0.1033
Sample 5	4.3833	-1.7167
Sample 6	5.6967	-0.4033
Sample 7	5.7267	-0.3733

Summary of the seven sample means

Overall mean waiting time for all 100 restaurants = 6.1000 minutes.

Mean of the seven worked sample means = 5.4943 minutes.

Observed standard deviation of the seven sample means = 0.8146 minutes.

Estimated standard error for n = 30 using $s/\sqrt{30} = 5.65085/\sqrt{30} = 1.0317$ minutes.

The mean of only seven sample means does not have to equal 6.10 exactly. Here it is lower because two of the seven samples have relatively low means (4.35 and 4.3833). With many more repeated random samples, the average of the sample means would tend to move closer to the population mean.

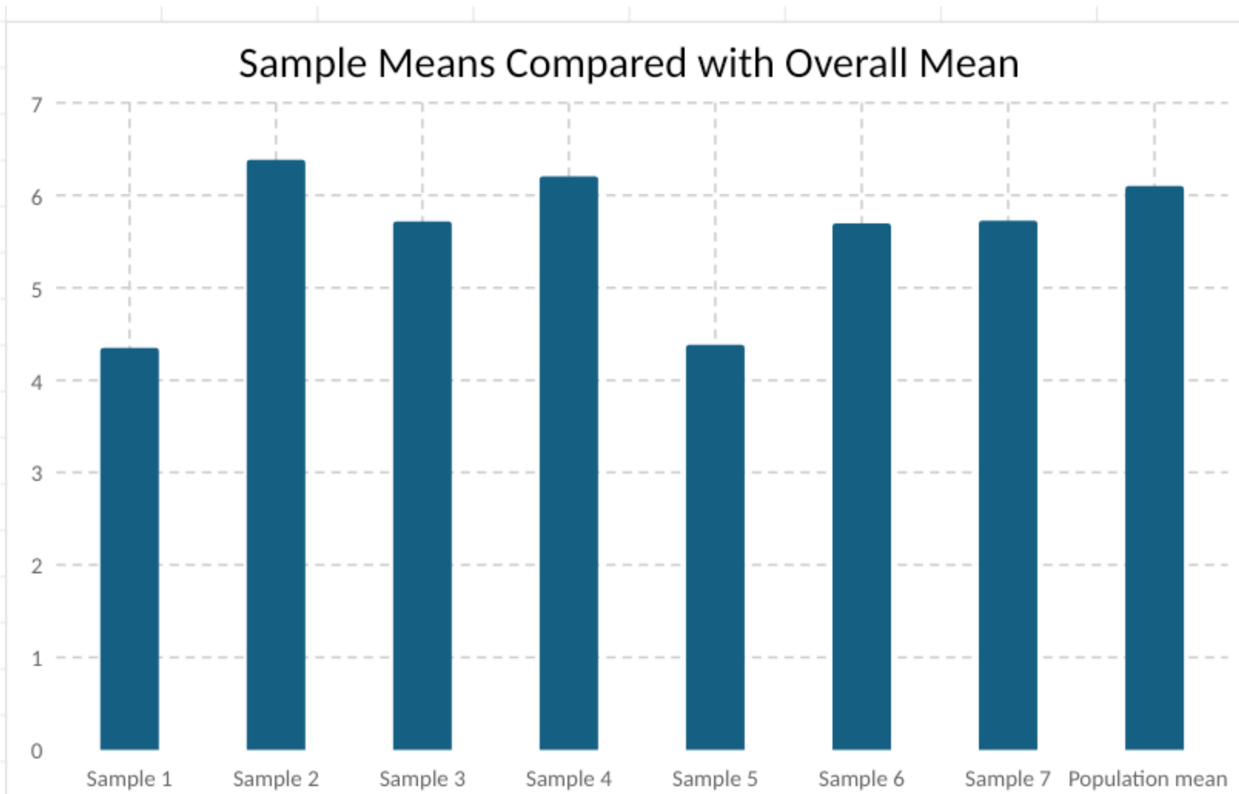

Figure 2: The seven sample means vary around the overall mean of 6.10 minutes. Their spread is much smaller than the 27.7-minute range of individual waiting times

How Part 2 demonstrates the Central Limit Theorem

Key lesson

The original Waiting Time distribution is strongly right-skewed, but averages based on samples of 30 observations are much less variable than individual waiting times. The Central Limit Theorem says that, for sufficiently large n, the sampling distribution of $\bar{x}$ becomes approximately normal even when the population itself is not normal. The standard error $s/\sqrt{n}$ quantifies how much sample means typically vary from sample to sample.

PART 3 - Pearson MyLab

1. Complete the Homework #3 questions on MyLab.

Note: If you are having issues with obtaining access to MyLab, please re-read the instructions (available in the welcome topic). If you are still unsuccessful, please send an email to the teaching team explaining the issue. It is important that you complete MyLab tasks on a weekly basis since the final e-test in this unit will be completed using MyLab.