

60080079 Introduction to Statistical Methods
Semester 2 2023-2024
Homework Assignment 4

21 CST H3Art

1. A few facts about the sampling distribution of the mean:

a) Its mean is equal to the mean of the population from which the sample came, as in

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

b) Its variance is equal to the original variance divided by the sample size, as in

$$\sigma_{\bar{X}}^2 = \sigma^2/n, \text{ or equivalently, } \sigma_{\bar{X}} = \sigma/\sqrt{n}.$$

c) Regardless of the shape of the original distribution, the sampling distribution will resemble the normal distribution more and more as the sample size gets larger.

Based on 1a) and 1b), complete the table below for $\mu = 0.10$ and $\sigma = 0.30$.

	$n = 10$	$n = 10,000$
$\mu_{\bar{X}}$	[a] 0._0	[c] 0._0
$\sigma_{\bar{X}}$	[b] 0.0_49	[d] 0.00_0

The four-digit number representing the missing numbers from [a] through [d] is ____.

Write your answer as a four-digit number.

Answer: 1913

2. Based on your answers above:

2.1 Let $\bar{X} \sim N([a],[b])$. It follows that $P(\bar{X} < 0.1_40) = 0.75$.

2.2 Let $\bar{X} \sim N([c],[d])$. Then $P(\bar{X} < 0.1_20) = 0.75$.

Note that these questions are equivalent to asking what the third quartiles (i.e., Q3) or the 75th percentiles are.

Write your answer as a two-digit number.

Answer: 60

3. As discussed in the class, when the normal distribution approximates another distribution relatively well, the area under the normal curve can be used to determine some aspects of the distribution.

For the problems below, use only $n = 10$ and 10,000 in **LLN.sav**.

In SPSS, determine the first quartile for $n = 10$. Do the same for $n = 10,000$. Use **Analyze** → **Descriptive Statistics** → **Explore**, then click the **Statistics** button, and then check the **Percentiles** option, before proceeding.

The Q3 for $n = 10$ is 0._000, whereas the Q3 for $n = 10,000$ is 0._019.

Write your answer as a two-digit number.

Answer: 21

4. In comparing the results in 2) and 3):

4.1 We can see that the approximation based on $n = 10$ is off by 0.0_60, whereas the approximation based on $n = 10,000$ is off by 0.0_01.

4.2 From these results, we can conclude that the sampling distribution is better approximated by the normal distribution when n is equal to (1) 10 / (2) 10,000.

Write your answer as a three-digit number.

Answer: 302

5. “How would you describe your own physical health at this time? Would you say your physical health is—excellent, good, only fair, or poor?” the Gallup Poll asked this question of 100 randomly selected adults, of whom 29% said “excellent.” Suppose that in fact the proportion of the adult population who say their health is excellent is $p = 0.29$.

5.1 Define $X = 1$ if physical health is excellent; $X = 0$ otherwise. Then X follows a ____ distribution.

- 1) normal
- 2) uniform
- 3) Bernoulli

5.2 Based on the distribution you selected, the mean of X is ____, and the standard deviation is ____.

- 1) 0.290
- 2) 0.710
- 3) 0.454
- 4) 0.206

5.3 Based on the Central Limit Theorem, the mean of the distribution approximately follows ____.

- 1) $N(0.29, 0.454)$
- 2) $Ber(0.29)$ with $SD = 0.454$

- 3) $Ber(0.29)$ with $SD = 0.0454$
- 4) $N(0.29, 0.0454)$

Write your answer as a four-digit number.

Answer: 3134

6. The scores of students on the ACT college entrance examination in 2001 had mean $\mu = 21.0$ and standard deviation $\sigma = 4.7$. The distribution of scores is only roughly normal.

6.1 “What is the approximate probability that a single student is randomly chosen from all those taking the test scores 23 or higher?” If we define X as the SAT score, this question is statistically equivalent to ____.

- 1) “ $P(X \leq 23) = ?$ ”
- 2) “ $P(\bar{X} \geq 23) = ?$ ”
- 3) “ $P(|X| = 23) = ?$ ”
- 4) “ $P(X \geq 23) = ?$ ”

6.2 The answer in 6. 1 is ____.

- 1) 0.3336
- 2) 0.6664
- 3) 0.4556
- 4) 0.6113

Write your answer as a two-digit number.

Answer: 41

7. Continuing Problem 6 now take an SRS of 50 students who took the test.

7.1 The mean and standard deviation of the sample mean score \bar{X} of these 50 students are ____ and ____, respectively.

- 1) 21.0
- 2) 2.97
- 3) 4.7
- 4) 0.66

7.2 The approximate probability that the mean score \bar{X} of these students is 23 or higher is ____.

- 1) 0.0223
- 2)
- 3) 0.0013
- 4) 0.6664

Write your answer as a three-digit number.

Answer: 143

8. The number of accidents per week, assumed to be X , at a hazardous intersection varies with mean 2.2 and standard deviation 1.4. This distribution takes only whole-number values, so it is certainly not normal. Let \bar{X} be the number of accidents per week at the intersection during a year (52 weeks).

8.1 What is the distribution of \bar{X} according to the Central Limit Theorem?

- 1) Approximately $N(2.2, 1.4)$
- 2) Approximately $N(2.2, 0.194)$
- 3) Exactly $N(2.2, 1.4)$
- 4) Exactly $N(2.2, 0.194)$

8.2 The approximate probability that \bar{X} is less than 2 is ____.

- 1) 0.2678
- 2) 0.7853
- 3) 0.0015
- 4) 0.1515

Write your answer as a two-digit number.

Answer: 24