SOLUTIONS

(1) Which of the following is a continuous probability distribution? **(D)**

(2) In measuring the strength of a positive linear relationship between X and Y, which of the following affect the magnitude of the coefficient of correlation? **(C)**

(3) Which kind of data often causes an endogeneity bias? **(C)**

(4) For a sample taken from negatively skewed population, which of the following statements about the sample mean would be expected? **(D)**

(5) For a sample of size 100 taken from a Poisson population with $\lambda = 1$, what is the distribution of the sample mean? **(D)**

(6) Considering a bell shaped population with a mean of \$41,201 and a standard deviation of \$10,779, what is the population 51st percentile? **(A)**

(7) Which of the following describe observational data? **(B)**

(8) X measures the negotiated selling price in dollars for a new Toyota Camry in Toronto. The population distribution of X is normal with a mean of \$23,000. What is the probability that for a randomly selected deal the buyer pays a price more than one standard deviation above average? **(A)**

(9) What is the point estimate of the population mean? **(B)**

(10) What is the population standard deviation? **(C)**

(11) For the empty box next to "Std. Dev." what is the <u>expected value</u> of what belongs in that box? **(C)**

(12) Which of the following would be expected to decrease the value of "Std. Dev."? **(B)**

(13) These Monte Carlo simulation results support which conclusion? **(A)**

(14) In words, the meaning of $E[\overline{X}_1 - \overline{X}_2] = \mu_1 - \mu_2$ is that the difference between the sample means is ________________________ the difference between the population means. **(B)**

(15) Which linear transformation of X yields Y? **(D)**

(16) The width of the confidence interval estimator of the population mean will increase with an increase in which of the following? **(A)**

(17) What percent of the sample is within one standard deviation of the mean? **(A)**

(18) Which is the approximate 95% confidence interval estimator of the population mean? **(D)**

(19) What is the 95% confidence interval estimator of the population mean? **(A)**

(20) Supposing that the sample standard deviation did not change, doubling the sample size would do what to the overall width of the 95% confidence interval estimator? **(C)**

Last Name:	S	O	L	U	T	I	O	N	S											

(1) $\overline{X} = 0.2359*0 + 0.32.04*1 + 0.2254*2 + 0.1444*3 + 0.0458*4 + 0.0246*5 + 0.0035*8 = 1.54$
Let's see what the probability distribution would be if λ = 1.54 (by using Poisson probability formula):

X	P(X)
0	0.214
1	0.330
2	0.254
3	0.130
4	0.050
5	0.015
6	0.004
7	0.001
8	0.0002

It seems reasonable to infer that the sample came from a Poisson population with $\lambda \approx 1.54$. The probabilities are not exactly equal to the relative frequencies in the sample, but they are close. We know that there will be sampling noise in the sample so it is not reasonable to expect equality. (Note: Some students may calculate the sample variance, which is 1.73, and note that it is close to the sample mean of 1.54. However, this is an incomplete answer. Just because the mean and variance are close does <u>not</u> necessarily mean that the shape is Poisson: it is necessary but not sufficient. Hence, substantial points should be deducted if the above table is missing.)

(2) This can be solved by recognizing that the mean and median (because normal population, which is symmetric) are 6 (center line of box) and any of the following facts from the box plot (student must understand the meaning of the box edges, 25th and 75th percentiles):

$$P(6 < X < 7) = 0.25$$
$$P(5 < X < 6) = 0.25$$
$$P(X < 5) = 0.25$$
$$P(X > 7) = 0.25$$

Here is one sample solution making use of the fact that P(X > 7) = 0.25 and the fact that the mean is about 6.

$$P(X > 7) = 0.25$$
$$P(Z > ?) = 0.25$$
$$? = 0.675$$
$$\frac{7 - \overline{X}}{s} = ?$$
$$\frac{7 - 6}{s} = 0.675$$
$$s \approx 1.48$$

(Note: The value from the standard normal table, 0.675, does not need to be extrapolated for full marks: 0.68 or 0.67 is good enough.)

(3) We have a sufficiently large sample size (35 > 30) to employ the Central Limit Theorem (CLT), which says that the sampling distribution of the sample mean will be approximately normal. The fact that the population is positively skewed is hence irrelevant. To find the probability, standardize.

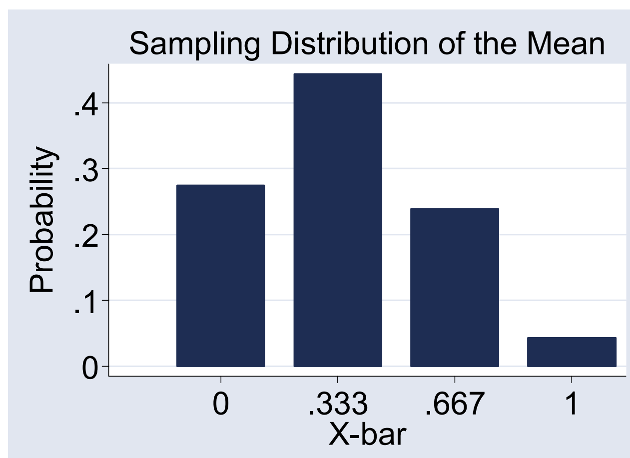
$$
\begin{aligned}
& P(\overline{X} < 18 \mid \mu = 20, \sigma^2 = 100) \\
&= P\left(Z < \frac{18 - 20}{10 / \sqrt{35}} \right) \\
&= P(Z < -1.18) \\
&= 0.5 - 0.3810 \\
&= 0.119
\end{aligned}
$$

This means that there is about an 11.9% chance that we got such a low sample mean due to pure chance: sampling error. Hence, we would not feel confident in saying that the Public Relations person is lying. You could recommend that the union boss collect a larger sample size to see if convincing evidence of lying can be found.

(4)

(a) One approach is to recognize that we can use the Binomial probability formula. Another is to do this by hand (as we did in Lecture 18), which is the solution given here:

Sample	Probability	X-bar
0,0,0	$(0.65)^3 = 0.2746$	0
0,0,1	$(0.65)^2(0.35)^1 = 0.1479$	1/3
0,1,0	$(0.65)^2(0.35)^1 = 0.1479$	1/3
0,1,1	$(0.65)^1(0.35)^2 = 0.0796$	2/3
1,0,0	$(0.65)^2(0.35)^1 = 0.1479$	1/3
1,0,1	$(0.65)^1(0.35)^2 = 0.0796$	2/3
1,1,0	$(0.65)^1(0.35)^2 = 0.0796$	2/3
1,1,1	$(0.35)^3 = 0.0429$	1



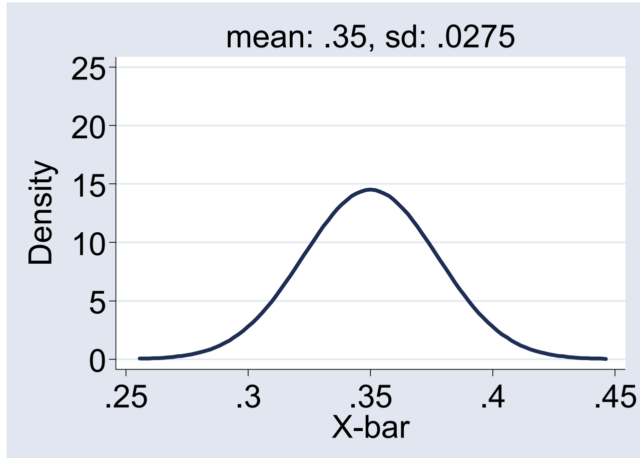
We can calculate the requested probability by using the sampling distribution we just found:

$$P(\overline{X} > 0.25) = 1 - 0.2746 = 0.7254$$

(b) A sample size of 300 is sufficiently large (300 > 30) so that we can use the Central Limit Theorem (CLT) to figure out that the sampling distribution of the sample mean is bell shaped (normal). All that remains is to find the mean and standard deviation of the distribution.

$$E[\overline{X}] = \mu$$
$$\mu = 0.35$$
$$E[\overline{X}] = 0.35$$

$$V[\overline{X}] = \frac{\sigma^2}{n}$$
$$V[\overline{X}] = \frac{0.35 * 0.65}{300}$$
$$V[\overline{X}] = 0.000758$$
$$SD[\overline{X}] = 0.0275$$



We can calculate the requested probability by using the sampling distribution we just found:

$$P(\overline{X} > 0.25) = P\left(Z > \frac{0.25 - 0.35}{0.0275}\right) = P(Z > -3.64) \approx 1$$