AP Statistics Chapter 7 Practice Test
Sampling Distributions

Name Key
Per ___ Date ___

measure from sample

Part 1: sample **Multiple Choice.** *Circle the letter corresponding to the best answer.*

A study of voting chose 663 registered voters at random shortly after an election. Of these, 72% said they had voted in the election. Election records show that only 56% of registered voters voted in the election.

measure from population

a 1. Which of the following terms describes the number 72%?
(a) statistic. (b) sample. (c) sample parameter. (d) population parameter. (e) population.

d 2. Which of the following terms describes the number 56%?
(a) statistic. (b) sample. (c) sample parameter. (d) population parameter. (e) population.

c 3. The sampling distribution of a statistic is
(a) the probability that we obtain the statistic in repeated random samples.
(b) the mechanism that determines whether randomization was effective.
(c) the distribution of values taken by a statistic in all possible samples of the same sample size from the same population.
(d) the extent to which the sample results differ systematically from the truth.
(e) the distribution of values in a sample of size *n* from the population

d 4. If a statistic used to estimate a parameter is such that the mean of its sampling distribution is equal to the true value of the parameter being estimated, what is the statistic said to be?
(a) random (b) biased (c) a proportion (d) unbiased (e) non-varying.

c 5. In order to use the formula $\sigma_x = \dfrac{\sigma}{\sqrt{n}}$ to calculate the standard deviation of the sampling distribution of the sample mean, which of the following conditions must be met?

I. ~~$n \geq 30$~~
II. ~~The~~ population's distribution is approximately Normal.
III. The sample size is less than 10% of the population size. ✓

(a) I only (b) II only (c) III only (d) III and *either* I or II
(e) All three conditions must be met.

b 6. Which of the following distributions has a mean that varies from sample to sample?

~~I. The population distribution~~
II. The distribution of sample data
III. The sampling distribution

(a) I only (b) II only (c) III only (d) II and III
(e) all three distributions

→ used when not normal

a 7. The central limit theorem is important in statistics because it allows us to use the Normal Distribution to make inferences concerning the population mean
(a) if the sample size is reasonably large no matter the shape of the population distribution
~~(b)~~ if the population is Normally distributed and the sample size is reasonably large.
~~(c)~~ if the population is Normally distributed no matter the sample size
~~(d)~~ if the population is Normally distributed and the population variance is known
(e) if the ~~population size~~ is reasonably large no matter the shape of the population distribution

d 8. A simple random sample of 4000 Alaskan voters showed that 62% were unhappy with our Governor. A simple random sample of 4000 Californians found that 59% were unhappy with their Governor. The sampling variability associated with these statistics is

(a) smaller for the sample of Californians because the percent unhappy was smaller than that for the Alaskans.
(b) larger for the Alaskans because Alaskans are more widely dispersed throughout the state than are Californians, hence they have more variable views.
(c) much smaller for the sample of Alaskans because the population of Alaska is much smaller than that of the California, hence the sample is a larger proportion of the population.
(d) not exactly the same, but very close.
(e) exactly the same.

C 9. A machine is designed to fill 16-ounce bottles of shampoo. When the machine is working properly, the mean amount poured into the bottles is 16.05 ounces with a standard deviation of 0.1 ounce. Assume the machine is working properly. If 12 bottles are randomly selected, which of the following best describes what we know about the sampling distribution of means for the ounces of shampoo for the sample?

(a) $\mu_{\bar{x}}=16.05$; $\sigma_{\bar{x}}=0.1$; shape of distribution unknown
(b) $\mu_{\bar{x}}=16.05$; $\sigma_{\bar{x}}=0.1$; distribution approximately Normal
(c) $\mu_{\bar{x}}=16.05$; $\sigma_{\bar{x}}=0.03$; shape of distribution unknown
(d) $\mu_{\bar{x}}=16.05$; $\sigma_{\bar{x}}=0.03$; distribution approximately Normal
(e) $\mu_{\bar{x}}=16.05$; $\sigma_{\bar{x}}=$ unknown; shape of distribution unknown

$\sigma = \frac{.1}{\sqrt{12}} = 0.0288675135$

C 10. You take a sample of size 35 from a very large population in which the true proportion is $p = 0.2$. Which statement below best describes what you know about the sampling distribution of $\hat{p}$?

(a) We do not know $\mu_{\hat{p}}$; $\sigma_{\hat{p}}=\sqrt{\frac{(0.2)(0.8)}{35}}$; the distribution is approximately Normal.

(b) $\mu_{\hat{p}}=0.2$; $\sigma_{\hat{p}}=\sqrt{\frac{(0.2)(0.8)}{35}}$; the distribution is approximately Normal.

(c) $\mu_{\hat{p}}=0.2$; $\sigma_{\hat{p}}=\sqrt{\frac{(0.2)(0.8)}{35}}$; the distribution is not approximately normal

(d) $\mu_{\hat{p}}=0.2$; we cannot use the formula $\sigma_{\hat{p}}=\sqrt{\frac{(0.2)(0.8)}{35}}$; the distribution is not approximately Normal

(e) $\mu_{\hat{p}}=0.2$; we cannot use the formula $\sigma_{\hat{p}}=\sqrt{\frac{(0.2)(0.8)}{35}}$; the distribution is approximately Normal.

normal condition $\rightarrow 35(.2) \ngeq 10$

b 11. The Wechsler Adult Intelligence Scale is a common "IQ test" for adults. The distribution of WAIS scores for persons over 16 years of age is approximately Normal with mean 100 and standard deviation of 15. Which of the following represents the probability that the mean score of a random sample of 60 adults from this population is greater than 105?

(a) *normalcdf (0, 105, 100, 1.9)*
(b) *normalcdf (105, 1000, 100, 1.9)*
(c) *normalcdf (0, 100, 105, 0.25)*
(d) *normalcdf (100, 1000, 105, 0.25)*
(e) *normalcdf (105, 1000, 100, 0.25)*

$\mu = 100$

$\sigma = \frac{15}{\sqrt{60}} = 1.9365$

LB 105

c 12. In a large population, suppose 15% of all adults jog. A simple random sample of 1540 adults from this population is to be contacted and the sample proportion computed. Which of the following expressions represents the probability that less than 13% of random sample of adults are actually joggers?

~~(a) *normalcdf (0.13, 1, 0.15, 0.01)*~~
~~(b) *normalcdf (0.15, 1, 0.13, 0.01)*~~
(c) *normalcdf (0, 0.13, 0.15, 0.01)*
~~(d) *normalcdf (0, 0.15, 0.13, 0.01)*~~
(e) *It doesn't meet the normal condition requirement.*

$\mu = .15$

$\sigma = \sqrt{\frac{.15(.85)}{1540}} = 0.009099$

$p = 0.15$

LB
UB .13

normal ✓
$.15(1540)$
$231 \geq 10$

Part 2: Free Response

Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

13. The amount of households pay service providers for access to the Internet varies quite a bit, but the mean monthly fee is $28 and the standard deviation is $10.

(a) Can you calculate the probability that a randomly chosen household pays for access to the internet exceeds $29? If so, do it. If not, explain why not.

No, we do not know if the population distribution is normal.

(b) What is the *shape* of the sampling distribution of the mean $\bar{x}$ monthly fee for samples of 55 randomly selected households? Justify your answer.

Since our sample size is greater than 30, the sampling distribution is approximately normal.

(c) What are the mean and standard deviation for the mean monthly fee $\bar{x}$ for internet access by an SRS of 55 households?

$\mu_{\bar{x}} = 28$

$\sigma_{\bar{x}} = \frac{10}{\sqrt{55}} = 1.348399725$

(d) Find the probability that the average monthly fee for internet access by an SRS of 55 households is greater than $29. Show your work.

normalcdf(lower bound 29, upper bound 1000, $\mu_{\bar{x}} = 28, \sigma_{\bar{x}} = 1.348...$)

≈ 0.2292

There is a 22.92% chance that a sample of 55 households will have an average monthly fee greater than $29.

14. Suppose you are going to roll a six-sided die 60 times and record $\hat{p}$, the proportion of times that a 1 or 2 is showing. 1/3

(a) Assume for the moment that the die is fair. What are the mean and standard deviation of the sampling distribution of $\hat{p}$?

$\mu_{\hat{p}} = 1/3 = 0.\overline{33}$

$\sigma_{\hat{p}} = \sqrt{\frac{1/3(2/3)}{60}} = 0.0608580619$

(b) Explain why you can use the formula for the standard deviation of $\hat{p}$ in this setting.

Each roll of the die is independent from each other.

(c) You roll the die 60 times and you get 30 rolls of 1 or 2, for a $\hat{p}$ of 0.5. Are you suspicious about whether the die is fair? Support your answer with an appropriate probability calculation.

normalcdf (lower bound .5, upper bound 1, $\mu_{\hat{p}} = 1/3, \sigma_{\hat{p}} = 0.0608580619$)

$= 0.003085$

normal condition

$1/3(60) = 20 \geq 10$

$2/3(60) = 40 \geq 10$

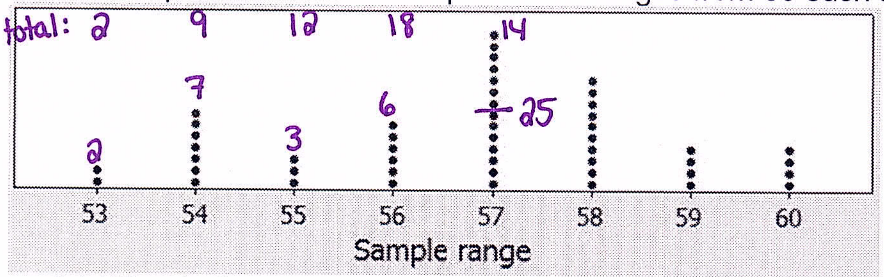
✓

Yes, since there is only a 0.3% chance of getting 30 out of 60 rolls resulting in a 1 or 2, I am suspicious that the die is not fair.

15. A small internet mail-order company keeps track of the number of orders it fills per day for many years and determines that the distribution of the variable "orders filled per day" is roughly symmetric and has the following five-number summary:

Minimum = 20 Quartile 1 = 32 Median = 46 Quartile 3 = 63 Maximum = 80

Suppose we take random samples of size 40 from this distribution and calculate the range for each of our samples. Below is a dotplot of the ranges from 50 such samples.



50 samples
median after
25 samples

(a) What is the parameter of interest and its value?

Range (max-min) of the population distribution of "orders per day" for a small internet mail-order company = 60 orders

(b) Briefly explain what the dot at 53 represents.

53 is the range (max-min) of orders per day of a random sample of 40 days.

(c) What is the median of the "sampling distribution"?

The median range is 57 orders

(d) Is the sample range an unbiased estimator of the population range? Use the dotplot above to justify your answer.

No, $57 \neq 60$

This makes sense since there is no way to get a range larger than the true range AND the only way to get the true range is if your sample happens to have both the max and min. All other samples will have a smaller range.