

Stat 321 - Day 33
Central Limit Theorem Calculations

You have previously discovered the single most important result in statistics, known as the *Central Limit Theorem*:

- The mean of the sampling distribution of \bar{X} equals the population mean μ , regardless of the sample size or the population distribution.
- The standard deviation of the sampling distribution of \bar{X} equals the population standard deviation σ divided by the square root of the sample size, regardless of the population distribution.
- The shape of the sampling distribution of \bar{X} is approximately normal for large sample sizes, regardless of the population distribution, and it is normal for any sample size when the population distribution is normal.

Example: Grade Point Averages (cont.)

Return to Oliver's grade point distribution for a single course:

x	2	3	4	$E(X) = \mu = 3.3$
$p(x)$.2	.3	.5	$V(X) = \sigma^2 = 0.61$ $SD(X) = \sigma = .781$

Suppose that Oliver takes $n=40$ courses in his college career. Let \bar{X} represent his GPA.

(a) In describing the distribution of \bar{X} , would the Central Limit Theorem apply here? Explain.

(b) What does the CLT say about the probability distribution of his GPA? [Be sure to mention shape, center, and spread.] Also draw a sketch of this sampling distribution, and be sure to indicate a label and numerical scale on the horizontal axis.

Shape:

Center:

Spread:

Sketch:

- (c) Use the CLT to approximate the probability that Oliver's GPA (from his 40-course career) is 3.5 or higher. [*Hints*: Standardize the value of 3.5 and use the normal probability table, or use Minitab, or use the Java applet. Be sure to use proper notation to express this probability as $P(\bar{X} \geq 3.5)$, and shade the corresponding area in the above graph.]
- (d) Would this probability increase or decrease (or stay the same) if the number of courses were to increase? Explain.
- (e) Use the CLT to determine the (approximate) probability that Oliver's GPA is 3.0 or lower.
- (f) Would this probability increase or decrease (or stay the same) if the number of courses were to increase? Explain.
- (g) Use the CLT to determine the (approximate) probability that his GPA is between 3.2 and 3.4.
- (h) Would this probability increase or decrease (or stay the same) if the number of courses were to increase? Explain.

Example: Potato Chip Assembly Line

Suppose that the weights of bags of potato chips coming off an assembly line are normally distributed with mean $\mu=12$ ounces and standard deviation $\sigma=0.4$ ounces.

- (a) What is the probability that one randomly selected bag weighs less than 11.9 ounces? [*Hint*: First draw and label a sketch of the probability distribution and shade the region whose area corresponds to this probability.]
- (b) If you took a random sample of ten bags, would you expect the probability of their sample mean weight being less than 11.9 ounces to be greater or less than the probability found in (a)? Explain, without performing the calculation.
- (c) Calculate the probability asked about in the previous question. [*Hint*: Draw and label a sketch of the sampling distribution and shade the region whose area corresponds to this probability.] Does this probability indicate that a sample mean as small as 11.9 ounces would be surprising if the population mean were really 12 ounces?
- (d) Repeat this analysis, for a sample of 100 randomly selected bags.
- (e) Repeat this analysis for a random sample of 1000 bags.

- (f) What is the smallest sample size for which the probability of the sample mean being less than 11.9 ounces is less than .01? [Hints: Find the first percentile of the standard normal distribution as the value z such that $P(Z \leq z) = .01$. Set this percentile equal to the z -score from standardizing 11.9 and solve for n . Alternatively, you can use Minitab's Calc > Probability Distributions > Normal command, choose "inverse cumulative probability," and use .01 as the input constant, so that Minitab will report the value x such that $P(X \leq x) = .01$.]
- (g) If you were told that a consumer group had weighed randomly selected bags and found a sample mean weight of 11.9 ounces, would you doubt the claim that the true mean weight of all of the potato chip bags is 12 ounces? On what unspecified information does your answer depend? Explain.
- (h) Which of your above answers would be affected if the distribution of the weights of the bags was not normal but was rather skewed?
- (i) What is the probability that the sample mean weight from a sample of 10 randomly selected bags would be between 11.75 and 12.25 ounces?
- (j) What is the probability that the sample mean weight from a sample of 100 randomly selected bags would be between 11.92 and 12.08 ounces?
- (k) Find a value k such that the probability of the sample mean weight of 1000 randomly selected bags being between $12-k$ and $12+k$ is roughly 0.95. In other words, between what two \bar{x} values do the middle 95% of the \bar{x} values fall?
- (l) What is the smallest sample size for which the probability is .95 that the sample mean falls within $\pm .05$ of 12 ounces (i.e., between 11.95 and 12.05)?