

Normal Distribution Worksheet/Sampling Distributions and the Central Limit Theorem

1. a. Draw two normal curves on the same axis below that have the same mean but different standard deviations. Describe the similarities and differences in their graphs.
b. Draw two normal curves on the same axis below that have the same standard deviation but different means. Describe the similarities and differences in their graphs.
2. Explain what a z-score of -1.8 means.

Steps to Find Percentile Corresponding to a Given Data Value

1. Sketch related normal distribution graph shading area.
2. Standardize the data value: Compute $z = (\text{Data Value} - \text{Mean}) / (\text{Standard Deviation})$
3. Find the corresponding probability in the z-table OR using TI-83/84: DISTR → 2:normalcdf(

Steps to Find Data Value Corresponding to a Given Percentile from a Table of Values

1. Sketch related normal distribution graph shading area.
2. Look up the given probability—the lower tail (the closest number to it) INSIDE the z-table OR using TI-83/84: DISTR → 3:invNorm(enter the lower tail area here)
3. Find the corresponding z-score.
4. Compute $x = (z\text{-score}) * \text{standard deviation} + \text{mean}$

3. Weights of Monkeys

The weights adult male rhesus monkeys are normally distributed with a mean of 15 pounds and a standard deviation of 3 pounds.

- a. What's the probability that a randomly selected monkey weighs less than 15 pounds?
- b. What proportion of monkeys weigh less than 13 pounds?
- c. What's the probability that a randomly selected monkey weighs more than 20 pounds?
- d. What proportion of monkeys weigh more than 22 pounds?
- e. What proportion of monkeys weigh between 11 and 19 pounds?
- f. What's the probability that a randomly selected monkey weighs between 8 and 23 pounds?
- g. What are the weights of the monkeys in the lowest 10%?
- h. What are the weights of the monkeys in the highest 20%?
- i. What are the weights of the monkeys in the lowest 15%?
- j. What are the weights of the monkeys in the highest 25%?

The Central Limit Theorem:

Given these conditions -

- The random variable X has a distribution with a mean of μ and standard deviation of σ .
- All samples taken are of size n .

We can conclude -

- If X was *normally distributed*, and $n > 0$, then:
The distribution of the sample mean \bar{X} will be *exactly* normal with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$.
- If X was *not* normally distributed, and n is large*, then:
The distribution of the sample mean \bar{X} will be *approximately* normal with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$.
- If X was *not* normally distributed, and n is small*, then:
No conclusions can be drawn about the distribution of \bar{X} .

*For convention, "n is large" usually means $n \geq 30$, and "n is small" means $n < 30$.

Note: n is the sample size, not the population size. The size of the population does not play any role in our inference.

4. For the following reports identify (i) the population, (ii) the population parameter of interest, (iii) the sample, and (iv) the sample statistic.
 - a. A company packaging snack foods maintains quality control by randomly selecting 10 cases from each day's production and weighing the bags, and then inspecting the contents. The weight of a case should be 2 lbs. One day they found that the weight of the 10 cases was 20.5 lbs.
 - b. State police set up a roadblock to estimate the percentage of cars with up-to-date registration and insurance. They found problems with 10% of the cars they stopped.
 - c. A magazine asked all subscribers whether they had used alternative medical treatments and, if so, whether they had benefited from them. For almost all of the treatments, approximately 20% of those responding reported cures or substantial improvement in their condition.
5. For each of the following indicate whether what is described is a *parameter* or a *statistic*:
 - a. The fraction of all Americans who have never seen an ocean in person.
 - b. The mean number of spots in a sample of 100 ladybugs have.
 - c. The proportion of 100 randomly chosen single-family houses in Orange County with a swimming pool
 - d. The percent of all defective iPods made by Apple.
 - e. The mean height of all kindergarten kids in California.

6. For each of the following, indicate whether what is described is a *parameter* or a *population*:

- a. All four-inch ham sandwiches sold at Quizno's.
- b. The average weight in ounces of all four-inch ham sandwiches sold at Quizno's.
- c. The proportion of registered drivers in California who had an accident in 2008.
- d. All apartment units in New York that are larger than 2000 square feet.
- e. The percentage of dogs and cats in Los Angeles that have been spayed or neutered.
- f. All 100 members of the United States Senate.

7. Suppose that studies were made to estimate the mean number of televisions owned by families in various California cities. Indicate which of the following would be likely to give the most accurate estimate of the city's mean, and which would probably give the least accurate estimate:

- a. A random sample of 500 families from Bakersfield (population 330,000)
- b. A random sample of 500 families from San Diego (population 1,350,000)
- c. A survey form distributed in various locations (shopping malls, libraries, post offices, etc.) of Santa Barbara (population 95,000), with 7,285 responses received

8. A jar of Jiffy Peanut Butter (labeled as a 32 oz. jar) is selected randomly off the end of the assembly line. The weight of the jar is measured. If X measures this weight then X has the $N(32.3, 0.40)$ distribution (the filling machine is calibrated this way). That is, the mean fill is 32.3 oz. and the standard deviation of the fills is 0.40 oz.

- a) Identify the variable being measured. Is it categorical or quantitative?
- b) Identify the population and parameters.
- c) What is the average fill of all Jiffy jars? Is this value a statistic?
- d) What's the probability a randomly selected jar is filled with less than 32.0 fl oz? Such a jar would be underweight.
- e) What proportion of all Jiffy jars are underweight?
- f) Consider a simple random sample of 6 jars. Let \bar{x} be the sample mean weight for these 6 jars. Is \bar{x} a parameter or a statistic?
- g) What are the mean and standard deviation of the (sampling) distribution for the sample mean?
- h) What is the shape of the sample means' distribution?
- i) Find the probability that the sample mean fill of the six jars is less than 32.0 fl oz.
- j) Now take a simple random sample of 24 jars. (The sample size has been multiplied by 4.) What are the mean and standard deviation of the distribution for \bar{x} ? How has the standard deviation of changed with the quadrupling of the sample size?
- k) Find the probability the sample mean fill of the 24 jars is less than 32.0 fl.oz.

9. The per capita consumption of processed fruits by people in the U.S. in a recent year was normally distributed with a mean of 152.7 lb and a standard deviation of 51.6 lb.
- Suppose many random samples of 9 people each are drawn from the population, and the mean of each sample is determined. Use the Central Limit Theorem (CLT) to find the mean and standard error of the sampling distribution of the sample means. Would it be unusual for a random group of 9 people to have a sample mean as high as 175 lb?
 - Now suppose many random samples of 36 people each are drawn from the population, and the mean of each sample is determined. Use the CLT to find the mean and standard error of the sampling distribution of the sample means. Would it be unusual for a random group of 36 people to have a sample mean as high as 175 lb?
 - Finally, suppose many random samples of 64 people each are drawn from the population and the mean of each sample is determined. Use the CLT to find the mean and standard error of the sampling distribution of the sample means. Would it be unusual for a random group of 64 people to have a sample mean as high as 175 lb?
 - Sketch the sampling distributions from parts a, b, and c on the same scale.
10. Weekly salaries of a certain big company have a mean of \$564 and an s.d. of \$105.
- If one employee is randomly selected from this company, find the probability that his/her salary is for more than \$750 per week.
 - If 16 employees are randomly selected, what is the probability that their mean income per week is less than \$400?
 - If 40 employees are randomly selected, find the probability that their mean income per week is between \$525 and \$600.
11. A brake pad manufacturer claims its brake pads will last for 38,000 miles. You work for a consumer protection agency and you are testing this manufacturer's brake pads. Assume that the life spans of the brake pads are normally distributed. You randomly select 30 brake pads. In your tests, the mean life span of the brake pads you sample is 35,700 miles. Past analyses indicate that $\sigma = 5000$ miles.
- Assuming the manufacturer's claim is correct, what is the probability that the mean life span of the sample is as low as 35,700 miles? (That is, 35,700 miles or lower.)
 - Using your answer from (a), what do you think of the manufacturer's claim?
 - Would it be unusual to have an *individual* brake pad last for 35,700 miles? Why, or why not?