

7.2.1: The Central Limit Theorem for Sums

Suppose X is a random variable with a distribution that may be known or unknown (it can be any distribution) and suppose:

- μ_x = the mean of X
- σ_x = the standard deviation of X

If you draw random samples of size n , then as n increases, the random variable $\sum X$ consisting of sums tends to be normally distributed and

$$\sum X \sim N((n)(\mu_x), (\sqrt{n})(\sigma_x)). \quad (7.2.1.1)$$

The central limit theorem for sums says that if you keep drawing larger and larger samples and taking their sums, the sums form their own normal distribution (the sampling distribution), which approaches a normal distribution as the sample size increases. The normal distribution has a mean equal to the original mean multiplied by the sample size and a standard deviation equal to the original standard deviation multiplied by the square root of the sample size.

The random variable $\sum X$ has the following z-score associated with it:

- $\sum x$ is one sum.
- $z = \frac{\sum x - (n)(\mu_x)}{(\sqrt{n})(\sigma_x)}$
 - $(n)(\mu_x)$ = the mean of $\sum X$
 - $(\sqrt{n})(\sigma_x)$ = standard deviation of $\sum X$

Calculator

To find probabilities for sums on the calculator, follow these steps.

2nd DISTR

2: normalcdf

normalcdf (lower value of the area, upper value of the area, $(n)(\text{mean})$, $(\sqrt{n})(\text{standard deviation})$)

where:

- *mean* is the mean of the original distribution
- *standard deviation* is the standard deviation of the original distribution
- *sample size* = n

Example 7.2.1.1

An unknown distribution has a mean of 90 and a standard deviation of 15. A sample of size 80 is drawn randomly from the population.

- Find the probability that the sum of the 80 values (or the total of the 80 values) is more than 7,500.
- Find the sum that is 1.5 standard deviations above the mean of the sums.

Answer

Let X = one value from the original unknown population. The probability question asks you to find a probability for the sum (or total of) 80 values.

$\sum X$ = the sum or total of 80 values. Since $\mu_x = 90$, $\sigma_x = 15$, and $n = 80$, $\sum X \sim N((80)(90), (\sqrt{80})(15))$

- mean of the sums = $(n)(\mu_x) = (80)(90) = 7,200$
- standard deviation of the sums = $(\sqrt{n})(\sigma_x) = (\sqrt{80})(15) = (80)(15)$
- sum of 80 values = $\sum X = 7,500$

a. Find $P(\sum X > 7,500)$

$$P(\sum X > 7,500) = 0.0127$$

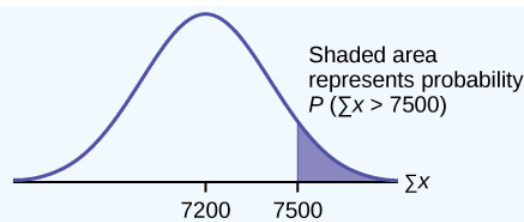


Figure 7.2.1.1.

`normalcdf` (lower value, upper value, mean of sums, `stdev` of sums)

The parameter list is abbreviated (*lower, upper, (n)(μ_x, (√n)(σ_x))*)

`normalcdf` (7500, 1E99, (80)(90), (√80)(15)) = 0.0127

REMINDER

1E99 = 10⁹⁹.

Press the **EE** key for E.

b. Find $\sum x$ where $z = 1.5$.

$$\sum x = (n)(\mu_x) + (z)(\sqrt{n})(\sigma_x) = (80)(90) + (1.5)(\sqrt{80})(15) = 7,401.2$$

? Exercise 7.2.1.1

An unknown distribution has a mean of 45 and a standard deviation of eight. A sample size of 50 is drawn randomly from the population. Find the probability that the sum of the 50 values is more than 2,400.

Answer

0.0040

Calculator

To find percentiles for sums on the calculator, follow these steps.

2nd DISTR

3:invNorm

$k = \text{invNorm}(\text{area to the left of } k, (n)(\text{mean}), (\sqrt{n})(\text{standard deviation}))$

where:

- k is the k^{th} **percentile**
- *mean* is the mean of the original distribution
- *standard deviation* is the standard deviation of the original distribution
- *sample size* = n

✓ Example 7.2.1.2

In a recent study reported Oct. 29, 2012 on the Flurry Blog, the mean age of tablet users is 34 years. Suppose the standard deviation is 15 years. The sample of size is 50.

- What are the mean and standard deviation for the sum of the ages of tablet users? What is the distribution?
- Find the probability that the sum of the ages is between 1,500 and 1,800 years.
- Find the 80th percentile for the sum of the 50 ages.

Answer

- $\mu_{\sum x} = n\mu_x = 1,700$ and $\sigma_{\sum x} = \sqrt{n}\sigma_x = (\sqrt{50})(15) = 106.01$
The distribution is normal for sums by the central limit theorem.

- b. $P(1500 < \sum X < 1800) = (1,500, 1,800, (50)(34), (\sqrt{50})(15)) = 0.7974$
 c. Let k = the 80th percentile.
 $k = (0.80, (50)(34), (\sqrt{50})(15)) = 1,789.3$

? Exercise 7.2.1.2

In a recent study reported Oct. 29, 2012 on the Flurry Blog, the mean age of tablet users is 35 years. Suppose the standard deviation is ten years. The sample size is 39.

- What are the mean and standard deviation for the sum of the ages of tablet users? What is the distribution?
- Find the probability that the sum of the ages is between 1,400 and 1,500 years.
- Find the 90th percentile for the sum of the 39 ages.

Answer

- $\mu_{\sum X} = n\mu_X = 1,365$ and $\sigma_{\sum X} = \sqrt{n}\sigma_x = 62.4$
 The distribution is normal for sums by the central limit theorem.
- $P(1400 < \sum X < 1500) = \text{normalcdf}(1400, 1500, (39)(35), (\sqrt{39})(10)) = 0.2723$
- Let k = the 90th percentile.
 $k = \text{invNorm}(0.90, (39)(35), (\sqrt{39})(10)) = 1445.0$

✓ Example 7.2.1.3

The mean number of minutes for app engagement by a tablet user is 8.2 minutes. Suppose the standard deviation is one minute. Take a sample of size 70.

- What are the mean and standard deviation for the sums?
- Find the 95th percentile for the sum of the sample. Interpret this value in a complete sentence.
- Find the probability that the sum of the sample is at least ten hours.

Answer

- $\mu_{\sum X} = n\mu_X = 70(8.2) = 574$ minutes and $\sigma_{\sum X}(\sqrt{n})(\sigma_x) = (\sqrt{70})(1) = 8.37$ minutes
- Let k = the 95th percentile.
 $k = \text{invNorm}(0.95, (70)(8.2), (\sqrt{70})(1)) = 587.76$ minutes
 Ninety five percent of the app engagement times are at most 587.76 minutes.
- ten hours = 600 minutes
 $P(\sum X \geq 600) = \text{normalcdf}(600, E99, (70)(8.2), (\sqrt{70})(1)) = 0.0009$

? Exercise 7.2.1.3

The mean number of minutes for app engagement by a table use is 8.2 minutes. Suppose the standard deviation is one minute. Take a sample size of 70.

- What is the probability that the sum of the sample is between seven hours and ten hours? What does this mean in context of the problem?
- Find the 84th and 16th percentiles for the sum of the sample. Interpret these values in context.

Answer

- 7 hours = 420 minutes
 10 hours = 600 minutes
 $\text{normalcdf}(420 \leq \sum X \leq 600) = \text{normalcdf}(420, 600, (70)(8.2), \sqrt{70}(1)) = 0.9991$
 This means that for this sample sums there is a 99.9% chance that the sums of usage minutes will be between 420 minutes and 600 minutes.
- $\text{invNorm}(0.84, (70)(8.2), \sqrt{70}(1)) = 582.32$
 $\text{invNorm}(0.16, (70)(8.2), (\sqrt{70})(1)) = 565.68$
 Since 84% of the app engagement times are at most 582.32 minutes and 16% of the app engagement times are at most

565.68 minutes, we may state that 68% of the app engagement times are between 565.68 minutes and 582.32 minutes.

References

1. Farago, Peter. "The Truth About Cats and Dogs: Smartphone vs Tablet Usage Differences." The Flurry Blog, 2013. Posted October 29, 2012. Available online at blog.flurry.com (accessed May 17, 2013).

Review

The central limit theorem tells us that for a population with any distribution, the distribution of the sums for the sample means approaches a normal distribution as the sample size increases. In other words, if the sample size is large enough, the distribution of the sums can be approximated by a normal distribution even if the original population is not normally distributed. Additionally, if the original population has a mean of μ_x and a standard deviation of σ_x , the mean of the sums is $n\mu_x$ and the standard deviation is $(\sqrt{n})(\sigma_x)$ where n is the sample size.

Formula Review

- The Central Limit Theorem for Sums: $\sum X \sim N[(n)(\mu_x), (\sqrt{n})(\sigma_x)]$
- Mean for Sums $(\sum X) : (n)(\mu_x)$
- The Central Limit Theorem for Sums z -score and standard deviation for sums: z for the sample mean = $\frac{\sum x - (n)(\mu_x)}{(\sqrt{n})(\sigma_x)}$
- Standard deviation for Sums $(\sum X) : (\sqrt{n})(\sigma_x)$

This page titled [7.2.1: The Central Limit Theorem for Sums](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [OpenStax](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.