

7.2: The Sum Distribution

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

$$\sum X \sim N((n)(\mu_x), (\sqrt{n})(\sigma_x)). \quad (7.2.1)$$

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 $\sum X$

Example 7.2.1

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 = $\sum X = 7,500$

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

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

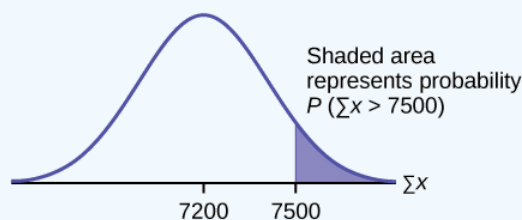


Figure 7.2.1.

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

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

$$\text{normalcdf} \quad (7500, 1E99, (80)(90), (\sqrt{80})(15)) = 0.0127$$

REMINDER

$$1E99 = 10^{99}.$$

Press the EE key for E.

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

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

Example 7.2.2

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 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

- $\mu_x = 34$ and $n\mu_x = 1,700$ and $\sigma_{\sum X} = \sqrt{n}\sigma_x = (\sqrt{50})(15) = 106.01$
The distribution is normal for sums by the central limit theorem.
- $P(1500 < \sum X < 1800) = P(1,500, 1,800, (50)(34), (\sqrt{50})(15)) = 0.7974$
- Let k = the 80th percentile.
 $k = (0.80, (50)(34), (\sqrt{50})(15)) = 1,789.3$

Example 7.2.3

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
 $P(\sum X \geq 600) = \text{normalcdf}(600, E99, (70)(8.2), (\sqrt{70})(1)) = 0.0009$

WeBWork Problems

References

- 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: The Sum Distribution](#) 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.

- **7.3: The Central Limit Theorem for Sums** by [OpenStax](#) is licensed [CC BY 4.0](#). Original source: <https://openstax.org/details/books/introductory-statistics>.