8 Central Limit Theorem

Demonstration: The Penny Experiment

We calculated sample means in *The Penny Experiment*, and we saw that the shape of the "histogram" changed dramatically when we went from the population of individual penny years to the population of average penny years in samples of size 5.

For large sample sizes, the sample means have a very convenient property:

> **The Central Limit Theorem**
> For a fixed sample size n, the set of all possible $\bar{x}$-values can be approximated by a normal distribution when n is sufficiently large, regardless of the shape of the population distribution.

Rule of thumb: the Central Limit Theorem can be used if the sample size is $n \geq 30$.

The mean of all $\bar{x}$-values for a fixed sample size is the same as the mean of the population that the samples were taken from. However, the standard deviation of the $\bar{x}$-values is smaller than the standard deviation of the original population.

> **Definition:** For a fixed sample size n, the standard deviation of the population of all $\bar{x}$-values is called the standard error of the mean, and is given by
>
> $$SE = \frac{\sigma}{\sqrt{n}},$$
>
> where σ is the SD of the population that the samples were taken from.

As a result, for samples of size $n \geq 30$, the z-score corresponding to a given $\bar{x}$-value is

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}},$$

and we use the *Standard Normal Distribution Table* to find probabilities involving sample means.

Example 8.1. A large class has a test average of 72 with a SD of 8. Take a random sample of n tests. Find the probability that the average of the n tests is more than 75 if:

(a) $n = 30$.

(b) $n = 80$.

Example 8.2. Suppose that checked baggage has a mean mass of 21 kg with a SD of 4 kg.

(a) If 40 bags are randomly selected, find the probability that their average mass is between 20 and 23 kg.

(b) Find the probability that the total mass of 50 randomly selected bags is greater than 1130 kg.

Example 8.3. [1, p. 241] Suppose that the sediment density (in g/cm^3) of specimens from a certain region is normally distributed with a mean of 2.65 and a standard deviation of 0.85.

(a) If a single specimen is randomly selected, what is the probability that its sediment density is at most 2.75?

(b) If a random sample of 35 such specimens is selected, what is the probability that the sample average sediment density is at most 2.75?

(c) How large a sample would be required to ensure that $P(\bar{x} \leq 2.75) \geq 0.99$?

(d) For a sample size of $n = 35$, find c such that $P(\bar{x} \geq c) = 0.95$?

Additional Notes

Additional Notes