The Central Limit Theorem

David Armstrong

UCI

Central Limit Theorem

Suppose we are interested in the sample mean $\bar{X}$, and we know that X has some distribution with $E[X] = \mu$ and $SD[X] = \sigma$.

- The MEAN of the SAMPLE AVERAGE is:

  $E\left[\bar{X}\right] = \mu_{\bar{X}}$

  As our sample size increases the mean of our sample will stay the same.

  PROOF:

- The STANDARD DEVIATION of the SAMPLE AVERAGE is:

  $SD\left[\bar{X}\right] = \sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}}$

  In general, taking the average of larger sample sizes gives a more precise estimate of the true mean. Thus, the spread around the center gets smaller.

  PROOF:

Central Limit Theorem

- Suppose $X \sim Normal\,(\mu, \sigma)$ and $n \geq 1$

  Then $\bar{X} \sim Normal\left(\mu_{\bar{X}} = \mu, \sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}}\right)$

  $Z = \dfrac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \dfrac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$

- Suppose $X \sim ApproximatelyNormal\,(\mu, \sigma)$ and $n \geq 1$

  Then $\bar{X} \sim ApproximatelyNormal\left(\mu_{\bar{X}} = \mu, \sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}}\right)$

  $Z = \dfrac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \dfrac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$

- The Central Limit Theorem (CLT):

  Draw a Simple Random Sample (SRS) of size $n \geq 30$ from any non-normal population with $E[X] = \mu$ and $SD[X] = \sigma$, then the sample mean has a sampling distribution that is approximately normal.

  Suppose $X \sim NonNormal\,(\mu, \sigma)$ and $n \geq 30$

  Then $\bar{X} \sim ApproximatelyNormal\left(\mu_{\bar{X}} = \mu, \sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}}\right)$

  $Z = \dfrac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \dfrac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$

Example: In engineering, weights of people are considered so that airplanes and elevators aren't overloaded, chairs won't break. Men's weights are normally distributed with a mean of 173 lbs., and a standard deviation of 30 lbs.

a. What is the distribution for a one randomly selected man's weight?

b. What is the probability a randomly selected man weighs more than 180 lbs.?

c. What is the distribution of the average men's weight if we are considering a SRS of 9 men?

d. If 9 men are randomly selected (say to be in an elevator), what is the probability that their average weight is more than 180 lbs.

Example: A rental car company has noticed that the distribution of the number of miles customers put on rental cars per day is right skewed. The distribution has a mean of 60 miles and a standard deviation of 25 miles. A random sample of 120 rental cars is selected.

a. Describe the sampling distribution of the average number of miles driven per day for the sample of 120 rental cars. Use the appropriate notation.

b. What is the probability that the mean number of miles driven per day for the sample of 120 cars is less than 54?

c. What is the probability that the total number of miles driven per day in the sample of 120 cars exceeds 7400?

Example: Which statement is correct regarding the Central Limit Theorem:

A. All variables have approximately normal shaped distributions if a random sample contains at least 30 observations.
B. Population distributions are normal whenever the population size is large.
C. For non-normal populations, the sampling distribution of the sample mean is approximately normal with a sufficiently large random sample.
D. The sampling distribution of the sample mean looks identical to the population distribution with a large sample size.

Inverse Calculations: This Z-score calculation can also be rearranged to solve for a sample mean:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \qquad \longrightarrow \qquad \bar{X} = Z\frac{\sigma}{\sqrt{n}} + \mu$$

Example: The amounts of telephone bills for all households in a large city have a distribution that is not normal with a mean of \$75 and a standard deviation of \$27. A random sample of 90 households is selected from this city.

a. What is the probability that the sample average telephone bill will be less than \$70?

b. What is the average telephone bill cost representing the 25th percentile?