1 The Central Limit Theorem and the Sampling Distribution of the Mean

The goal of inferential statistics is to use a sample to make an inference about a population.

A class of 50 students wants to study the average GPA at KSU.

Student number 1 collects a sample of 5 student GPA's.
$S_1 = \{3.01, 3.28, 2.97, 3.41, 3.21\}$, $\bar{x} = 3.176$
Student number 2 collects a sample of 5 student GPA's.
$S_2 = \{2.89, 3.33, 1.97, 2.59, 3.01\}$, $\bar{x} = 2.758$
Student number 3 collects a sample of 5 student GPA's.
$S_3 = \{2.93, 2.78, 3.41, 3.17, 2.81\}$, $\bar{x} = 3.02$
The remaining 47 students proceed in a similar fashion.

Problem 1 *Given 50 sample averages what might you do to estimate the true population average?*

Problem 2 *Are there differences in the variations in the single observations and the variations of the sample averages?*

Definition 1 *The **sampling distribution** of a sample statistic is the distribution of the values of the statistic created by repeated samples of n observations.*

Theorem 1 ***The Central Limit Theorem*** *(CLT for means) The mean of a random sample has a sampling distribution whose shape can be approximated by a normal model if $n \geq 30$. The larger the sample, the better the approximation will be. The center of the sampling distribution, μ_x, is the population mean μ. The spread of the sampling distribution, σ_x, is $\frac{\sigma}{\sqrt{n}}$. The value σ_x is called the **standard error**.*

Example 1 *When throwing three darts, the average score is 45 with a standard deviation of 11.5. What is the center and spread for the sampling distribution of the mean for the average score when each of the 40 students in class throws three darts? The center $\mu_x = \mu = 45$ and the spread is the standard error $\sigma_x = \frac{\sigma}{\sqrt{n}} = \frac{11.5}{\sqrt{40}} = 1.8183$*

Example 2 *What is the probability that the average score of the forty students in class will be less than 42 when throwing three darts? Using* $normalcdf(-\infty, 42, 45, 1.82)$ *we get* 0.0496.

Example 3 *What is the probability that the average score of the forty students in class will between 40 and 50 when throwing three darts? Using* $normalcdf(40, 50, 45, 1.82)$ *we get* 0.994.

Example 4 *What is the probability that Natasha will score between 40 and 50 when throwing three darts? Using* $normalcdf(40, 50, 45, 11.5)$ *we get* 0.3363.

Example 5 *What is the cutoff score between the worst 15% and the best 85% of groups of 40 students? This would be* P_{15} *in the sampling distribution of the mean. Using* $invNorm(.15, 45, 1.82)$ *we get* 43.11. *So, the worst 15% of all groups of 40 students have an average score less than 43.11.*

Example 6 *Is the normal model good for predicting the average score for a sample of 5 students? Explain. No, this sample size is too small to apply the Central Limit Theorem.*

Problem 3 *The time it takes students in a cooking school to learn to prepare seafood gumbo is a random variable with a normal distribution where the average is 3.2 hours with a standard deviation of 1.8 hours.*

Find the probability that the average time it will take a class of 36 students to learn to prepare seafood gumbo is less than 3.4 hours.

Find the probability that it takes one student between 3 and 4 hours to learn to prepare seafood gumbo.

Would it be unusual for the average time a group of 50 students needs to learn to prepare seafood gumbo is less than two hours and thirty minutes?

Problem 4 *Grade point averages at a particular school follow a normal distribution with* $\mu = 2.89$ *and* $\sigma = 0.63$.

i. Find the probability that the average GPA for a sample of 35 students is greater than 3.0.

ii. Find the probability that the average GPA for a sample of 40 students is between 2.0 and 2.75.

iii. Find the probability that Nathan's GPA is between 2.0 and 2.75.

Problem 5 *The time it takes a baseball player to learn the team's base-running signals is a random variable with a normal distribution where the average is 8.2 hours with a standard deviation of 2.2 hours.*

i. Find the probability that the average time it will take a roster of 36 players to learn the team's base-running signals is less than 9 hours.

ii. Find the probability that the average time it will take a roster of 45 players to learn the team's base-running signals is less than 9 hours.

iii. Find the probability that it takes Jesse less than 9 hours to learn the team's base-running signals.

Problem 6 *The lifespan of drummers in Spinal Tap follows a normal distribution with an average of 10 months and a standard deviation of 1.5 months. Neil signs on to be their next drummer. What is the probability that Neil survives to see his 1 year anniversary as the new drummer in Spinal Tap?*

2 Exercises

1. Navidi/Monk Section 6.3: 5, 6, 9-14, 17, 18, 20, 21, 23