Riley Payung

Professor Romanelli

CDS 290

November 14, 2020

Problem 1

I think that the shape of the distribution will be relatively normal and maybe even a little bit right-skewed, as with all things in the world. It is likely that the newer players will be paid less than the more veteran players, and the better players will get paid more than the worse players.

I think the mean will be $12-15 million.

Problem 2

The shape of the distribution is right-skewed.

There are 877 salaries in the dataset.

The mean salary is $4,509,878. The median is $1,400,000.

Yes it does make sense that the mean is larger, the maximum value is nearly 76x greater than the lowest value.

The standard deviation is $6,334,235.74. This standard deviation is extremely large. It gives us an idea of just how much more the better/veteran players are paid than the worse/newer players.

Problem 3

Its partial because its not making use of all of the data. Its only using a subset of 100 samples out of the 877 in the population.

Both samples are right-skewed.

Summary statistics:

Samples	n	Mean	Std. dev.	Median
Sample1(SALARY)	100	$3,007,772.20	$4,723,289.80	$677,050.00
Sample2(SALARY)	100	$4,926,585.40	$6,297,945.40	$1,558,333.00

The means compare relatively well in the second sample; however, the first mean is not compared too well against the entire population.

The standard deviation compares well for the second sample; however, the standard deviation of the first sample is not very similar; it is much lower. This is probably just due to the random nature of sampling.

If we look at the median for both samples, the second sample is closer to the real thing even though its off by a couple hundred thousand dollars.

Problem 4

A. I think that the sampling distribution will be similarly shaped (right skewed) to the original, regardless of the number of samples taken.
B. I think the mean will juggle between 3 million and 8 million.
C. I think the standard deviation will juggle between 3 million and 7 million.

Problem 5

Sample Size	Shape	Mean	Standard Deviation
2	Right skewed	3,220,076	3,617,389
4	Right skewed	4,474,390	2,904,796
5	Right skewed	4,538,538	3,075,360
8	Right skewed	4,594,899	2,207,257
10	Normal	4,662,121	1,926,168
12	Normal	4,430,715	2,077,628
16	Normal	4,334,874	1,378,739
18	Normal	4,520,041	1,632,822
20	Normal	4,547,735	1,582,289
22	Normal	4,578,251	1,448,958
25	Normal	4,492,987	1,107,605
28	Normal	4,593,892	1,183,905
30	Normal	4,559,390	1,259,600
32	Normal	4,644,339	1,114,390

Problem 6

The shape becomes more normal as the number of random samples increases. The center (or median), tends to become more centered based on the number of random samples, and the spread (or std. dev), tends to become more normal (e.g. closer to the confidence intervals) as the number of random samples increases. Basically, the entire distribution becomes more normal and natural as the number of random samples increases.

Summary

If we take a sampling distribution of size n from a non-normal population with mean, and standard, and n is sufficiently large ($n \geq 30$), then the shape of the sampling distribution is **Normal**.
This amazing fact is commonly known as **The Central Limit Theorem** or **T-Distribution**.

If we take a sampling distribution of size n from a normal population with mean, , and standard, , then the shape of the sampling distribution is **normal** for any sample size.

Part 2 (Parents and Gender)

500 participants, 60% agree.

300 people agreed that genders should have equal opportunities to participate in sports, 200 did not.

Lets say that we took a sample of size (n=15) of size (100), p-hat would be the distribution of the samples. What we get should is the mean gets closer and closer to 60% as we increase the number of samples (n) and the standard deviation should drop. The shape of this distribution would be slightly left skewed.