TA’S NAME:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Problem Set #3**

*Researchers at Johns Hopkins University recently determined that the mean batting average of all Major Leagues Baseball players in 1954 was 0.265, with a standard deviation of 0.035. In 2004, they noted, the mean batting average of all Major Leagues Baseball players was higher --- 0.275 --- with a somewhat smaller standard deviation of 0.030. In both years, the distribution of batting averages followed a standard normal curve. To summarize:*

*Year Mean Batting Average Standard Deviation of Batting Averages*

1954 0.265 0.035

2004 0.275 0.030

1. In which year was there more variability in the distribution of batting averages? Explain the reasoning behind your answer.

ANSWER: There was more variability in 1954; this is simply because the standard deviation of batting averages was higher in that year.

1. In 1954, Ted Williams had a batting average of 0.345. Roughly what proportion of players had batting averages as high or higher than Williams in that year? *Show your work*.

ANSWER: To find out, compute the *Z*-score corresponding to a batting average of 0.345. Since $Z=\frac{Y-\overline{Y}}{s\_{Y}}$, then in our example $Z=\frac{0.345-0.265}{0.035}$=2.29. According to Z tables, the probability of obtaining a *Z*-score of 2.29 or higher is 0.011.

1. Approximately what proportion of batters had batting averages between 0.200 and 0.300 in 1954? *Show your work*.

ANSWER: An average of 0.200 corresponds to a *Z*-score of $\frac{0.200-0.265}{0.035}$= -1.86. An average of 0.300 corresponds to a *Z*-score of $\frac{0.300-0.265}{0.035}$= 1. Since 0.0314 of the curve lies below *Z*=-1.86 and 0.1587 of the curve lies above *Z*=1, then 1-(0.0314+0.1587)=0.8099 (or 80.99%) of batters batted between 0.200 and 0.300 in 1954.

1. In 2004, Torii Hunter had a batting average of 0.271. Roughly what proportion of players had batting averages as low or lower than Hunter that year? *Show your work*.

ANSWER: To find out, we compute the Z-score corresponding to a batting average of 0.271 in 2004. Since$Z=\frac{Y-\overline{Y}}{s\_{Y}}$, then in our example$Z=\frac{0.271-0.275}{0.030}$= -0.13. According to Z tables, the probability of obtaining a *Z*-score as low or lower than -0.13 is 0.4483.

1. Approximately what proportion of batters had batting averages between 0.200 and 0.300 in 2004? *Show your work*.

ANSWER: An average of 0.200 corresponds to a *Z*-score of $\frac{0.200-0.275}{0.030}$= -2.5. An average of 0.300 corresponds to a *Z*-score of = 0.83. Since 0.0062 of the curve lies below *Z*=-2.5 and 0.2033 of the curve lies above *Z*=0.83, then 1-(0.0062+0.2033)=0.7905 (or 79.05%) of batters batted between 0.200 and 0.300 in 2004.

1. Given these data, in which year would it be *more* surprising to observe a batter with a batting average at or above 0.400? Explain the reasoning behind your answer.

A batter with a 0.400 average in 1954 would be $Z=\frac{Y-\overline{Y}}{s\_{Y}}$=$\frac{0.400-0.265}{0.035}$=3.86 standard deviations above the mean for that year. On the other hand, a batter with a 0.400 average in 2004 would be $Z=\frac{Y-\overline{Y}}{s\_{Y}}$ = $\frac{0.400-0.275}{0.030}$=4.17 standard deviations above the mean for that year. In which year is the area under the curve to the right of 0.400 larger? 1954. So even though mean batting averages went up between 1954 and 2004, we were less likely to observe a batter with a 0.400 average in 2004 than we were in 1954.

1. Psychologists know that 10% of Americans are left-handed. What is the probability of drawing a random sample of 2,000 Americans and observing that 12% or more of the sample members are left handed?

ANSWER: $Z=\frac{\hat{p}-p}{sd\_{\hat{p}}}=\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}=\frac{0.12-0.10}{\sqrt{\frac{(0.1)(0.9)}{2,000}}}=2.98$

Using a Z table, we see that the probability of obtaining a Z score of 2.98 or higher is 0.00144. That is, the probability of drawing a random sample of 2,000 Americans and observing that 12% or more of the sample members are left handed is very low.

1. I learned at the 2019 Minnesota State Fair that pregnant sows typically give birth to 10 piglets, with a standard deviation of 2 piglets. If we randomly selected 100 pregnant sows from the population for which these parameters hold, then what is the probability that the mean number of piglets born to these sows would 10.5 or lower?

ANSWER: $Z=\frac{\overline{Y}-μ\_{Y}}{sd\_{\overline{Y}}}=\frac{\overline{Y}-μ\_{Y}}{{σ}/{\sqrt{N}}}=\frac{10.5-10.0}{2/\sqrt{100}}=2.5$.

Using a Z table we see that the probability of obtaining a *Z* score of 2.5 or higher is 0.0062. Thus, the probability of obtaining a *Z* score of 2.5 or lower is 1–0.0062=0.9938. Thus the probability that the mean number of piglets born to these sows would be 10.5 or lower is very high.