Math 10B with Professor Stankova
Worksheet, Discussion #16; Tuesday, 3/19/2019
GSI name: Roy Zhao

1 Z-Scores, CLT, LoLN

1.1 Concepts

1. In order to compute the probability $P(a \leq X \leq b)$ for a normal distribution, we need to take an integral $\int_a^b \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2/\sigma^2}$ and this integral is almost impossible to do without a calculator. So, what we do is have a table of values for this integral and look up the value that we need. Given a z score such as 1.5, when we look it up in the table, $z(1.5) = P(0 \leq Z \leq 1.5)$, where Z is the standard normal distribution; the bell curve with mean $\mu = 0$ and standard deviation $\sigma = 1$. To find how many standard deviations a value a is away from the mean, you can use the formula $\frac{|a-\mu|}{\sigma}$.

   One key area these pop up in is when taking the average of a bunch of trials. For X_i independent and identically distributed (i.i.d.) (e.g. rolling a die multiple times) with $E[X_i] = \mu$ and $Var(X_i) = \sigma^2$, the **Central Limit Theorem (CLT)** tells us that the average that we get (e.g. the average number that we roll)

   - is **approximately** normal distributed (we can approximate probabilities with z-scores)
   - $\bar{\mu} = E[\bar{X}] = \mu$,
   - $\bar{\sigma} = SE(\bar{X}) = \sigma/\sqrt{n}$.

   So
   $$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$
   is approximately normally distributed with $E[\bar{X}] = \mu$ and $Var(\bar{X}) = \sigma^2/n$.

   In general, you can think of it in this table:

X	The test score of a single student in $10B$.
$\mu = E[X]$	The average test score of the students in $10B$.
$\sigma = SE(X) = \sqrt{Var(X)}$	The likelihood of a random student having a test score far away from the average μ.
$\bar{X} = \frac{X_1 + \cdots + X_n}{n}$	The average test score in a particular discussion section.
$\bar{\mu} = E[\bar{X}]$	The average over all sections of the average test score of the section.
$\bar{\sigma} = SE(\bar{X}) = \sqrt{Var(\bar{X})}$	The likelihood of a random discussion section having an average test score far away from the average $\bar{\mu}$.

Note that putting a bar over the constants μ, σ, X just tells us that we are considering the μ, σ for the averaged random variable $\bar{X}$. Now the reason that $\mu = \bar{\mu}$ is because if you think about it, we are still taking the average over the same population (everyone in $10B$). The reason that $\bar{\sigma} \leq \sigma$ is because if we take the average test score of a particular section, if there are students with high scores, it is likely their score will cancel with students with lower scores. So, it is less likely that a section average will be far away from the class average.

In order to compute probabilities, we compute the z score. Given a normal distribution with mean μ and standard deviation σ, the z score of a value a is $\frac{|a-\mu|}{\sigma}$. Then we look up this value in a table.

The **Law of Large Numbers** is a weaker statement that just says that as we take averages and let $n \to \infty$, then the sample mean becomes closer and closer to the actual mean μ. Namely, $E[\bar{X}] \to \mu$ and the probability that we are far away from the mean goes to 0. Mathematically, it says that for any fixed $\epsilon > 0$, we have

$$\lim_{n\to\infty} P(|\bar{X} - \mu| > \epsilon) = 0.$$

1.2 Examples

2. Let f be normally distributed with mean -2 and standard deviation 4. Calculate the probability $P(-1 \leq X \leq 1)$.

3. Suppose you ask a group of 9 voters whether they voted for Romney or Obama and suppose that they are equally likely to have voted for either. Calculate the exact probability that less than a third of them voted for Romney. Use the CLT to approximate the probability that less than a third of them voted for Romney.

1.3 Problems

4. True False We can only use the z score to calculate probabilities of normal distributions (bell curves).

5. True False The normal distribution with positive mean can only take on positive values. ($P(X \leq 0) = 0$)

6. True False You can use the Central Limit Theorem to prove the Law of Large Numbers.

7. True False Suppose I calculate that probability that in a sample of 10,000 men, their average height is less than 66 inches is 99.999%. Then all but one or two men in a sample of 10,000 men will have a height of less than 66 inches.

8. Let f be normally distributed with mean 1 and standard deviation 4. Calculate the probability $P(X \geq 3)$.

9. Let f be normally distributed with mean -2 and standard deviation 4. Calculate the probability $P(-3 \leq X \leq 1)$.

10. Suppose that on the most recent midterm, the average was 60 and the standard deviation 20. What is the approximate probability that a class of 25 had an average score of at least 66?

11. The newest Berkeley quarterback throws an average of 0.9 TDs/game with a standard deviation of 1. What is the probability that he averages at least 1 TD/game next season (16 total games)?

12. Suppose that the height of NBA players is distributed with an average height of 83 inches and a standard deviation of 10 inches. Taking a sample of 100 players, approximate the probability that the average of the heights of these 100 players is between 82 and 84 inches.