Math 10A with Professor Stankova
Worksheet, Discussion #37; Monday, 11/27/2017
GSI name: Roy Zhao

Independent Events

Example

1. Calculate the standard deviation of the binomial distribution with n trials and $p = \frac{1}{2}$.

> **Solution:** Let X_i be the result of one trial which is 1 for a success and 0 for no success. First, we calculate that variance of a single trial. The mean is $\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1 = \frac{1}{2}$ and so the variance is
>
> $$\frac{1}{2}(0 - 1/2)^2 + \frac{1}{2}(1 - 1/2)^2 = \frac{1}{4}.$$
>
> Thus, the variance of n trials, since the n trials are independent and hence variance adds, is $n/4$. So, the standard deviation is $\sqrt{n}/2$.

2. I flip a coin and roll two die. What is the probability that I flip heads and roll snake eyes (two 1s)?

> **Solution:** The probability that I flip heads is $\frac{1}{2}$. The probability that I roll a 1 is $\frac{1}{6}$. Since the first and second die rolls are independent and independent to the coin flip, the probability of all three happening is
>
> $$\frac{1}{2} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{72}.$$

Problems

3. True **FALSE** Let X be 1 the first card in a deck is an ace of spades and 0 otherwise. Let Y be 1 if the last card in the deck is a king of hearts and 0 otherwise. Then X, Y are independent random variables.

> **Solution:** If we know that $X = 1$, then Y is 1 with probability $1/51$ instead of $1/52$ so we know a little bit more about Y if we know something about X. Without knowing X, the probability Y is 1 is $1/52$.

4. True **FALSE** The expected values of the sum of any two random variables add only if the random variables are independent.

Solution: This is true for all random variables.

5. **TRUE** False The variance of the sum of any two random variables add only if the random variables are independent.

Solution: This is true for only independent random variables since then

$$E[(X+Y)^2]-(E[X+Y])^2 = E[X^2]+2E[XY]+E[Y^2]-(E[X]^2+2E[X]E[Y]+E[Y]^2)$$

$$= Var(X) + Var(Y) + 2E[XY] - 2E[X]E[Y] = Var(X) + Var(Y)$$

since for independent random variables, we know that $E[X]E[Y] = E[XY]$.

6. True **FALSE** I pick two random people from a crowd. Let X be the height of the first person and Y the height of the second. Then X, Y are independent.

Solution: Suppose that there was one 10 foot person and everyone else is 5 feet. Well if I know that $X = 10$, then I know that $Y = 5$, and hence they aren't independent.

Central Limit Theorem

Example

7. Show that the distribution of $\overline{X}$, the average of n i.i.d. random variables with mean μ and standard deviation σ has mean μ and standard deviation $\sigma/\sqrt{n}$.

Solution: First, we note that the mean is

$$E[\bar{X}] = E[\frac{X_1 + \cdots + X_n}{n}] = \frac{E[X_1] + \cdots + E[X_n]}{n} = \frac{\mu n}{n} = \mu.$$

Then, the variance adds and note that $Var(cX) = c^2 Var(X)$ to get that

$$Var(\bar{X}) = \frac{1}{n^2} Var(X_1 + \cdots + X_n) = \frac{Var(X_1) + \cdots + Var(X_n)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Therefore the standard error or standard deviation is $\sqrt{(\sigma^2)/n} = \sigma/\sqrt{n}$.

8. Suppose that in the 2012 election, 55% of people preferred Obama over Romney. If I sample 100 random people (assume that they are independently chosen), what is the probability that a majority of them support Obama?

Solution: Let X be a random variable that is 1 if the person prefers Obama and 0 otherwise. Then, we know that $E[X] = 0.55$ and $Var(X) = 0.55 \cdot (1-0.55)^2 + 0.45(0-0.55)^2 = 0.2475$ so $SE(X) \approx 0.5$. Let $\bar{X}$ be the average of asking 100 people, and hence $\bar{X}$ is normally distributed with mean 0.55 and standard deviation $0.5/\sqrt{100} = 0.05$. I want to calculate $P(\bar{X} \geq 0.5) = 0.5 + z(\frac{|0.5-0.55|}{0.05}) \approx 0.5 + z(1) \approx$ 84.14%.

Problems

9. True **FALSE** The example above proves the central limit theorem.

Solution: The above part just shows that the distribution has mean μ and standard deviation $\sigma/\sqrt{n}$. You still need to show that the distribution is a normal distribution, and this is a lot harder to do and beyond the scope of this class.

10. **TRUE** False You can use the Central Limit Theorem to prove the Law of Large Numbers.

Solution: We know that $\overline{X}$ has distribution with mean μ and standard deviation $\sigma/\sqrt{n}$ and hence as $n \to \infty$, the standard deviation $\to 0$ which means that all of the probability gets concentrated at μ, which is exactly what the law of large numbers says.

11. **TRUE** False For a constant $c \geq 0$, we have that $SE(cX) = cSE(X)$.

Solution: This comes from the fact that $Var(cX) = c^2Var(X)$ and so $SE(cX) = \sqrt{Var(cX)} = \sqrt{c^2Var(X)} = c\sqrt{Var(X)} = cSE(X)$.

12. True **FALSE** Suppose I calculate that probability that in a sample of $10,000$ men, their average height is less than 66 inches is 99.9%. Then all but one or two men in a sample of $10,000$ men will have a height of less than 66 inches.

> **Solution:** The probability that the average is less than 66 inches is very high but that does not mean that same percentage will actually have height less than 66 inches. (Average vs individual values)

13. Suppose that the height of women is distributed with an average height of 63 inches and a standard deviation of 10 inches. Taking a sample of 100 women, what is the probability that the average of the heights of these 100 women is between 62 and 64 inches?

> **Solution:** The average height of 100 women will be approximately normally distributed with average 63 and standard deviation $10/\sqrt{100} = 1$. Therefore, $P(62 \leq \bar{X} \leq 64) = P(62 \leq \bar{X} \leq 63) + P(63 \leq \bar{X} \leq 64) = z(1) + z(1) = 2z(1)$.

14. Suppose the weight of newborns is distributed with an average weight of 8 ounces and a standard deviation of 1 ounce. Today, there were 25 babies born at the Berkeley hospital. What is the probability that the average weight of these newborns is less than 7.5 ounces?

> **Solution:** The average weight of these babies will be approximately normally distributed with mean 8 and standard deviation $1/\sqrt{25} = 0.2$. The probability is $P(\bar{X} \leq 7.5) = 0.5 - P(7.5 \leq \bar{X} \leq 8) = 0.5 - z(2.5)$.

15. Suppose that the average lifespan of a human is 75 years with a standard deviation of 10 years. What is the probability that in a class of 25 students, they will on average live longer than 80 years?

> **Solution:** The average lifespan of 25 students is approximately normally distributed with mean 75 and standard deviation $10/\sqrt{25} = 2$. Thus $P(\bar{X} \geq 80) = 0.5 - z(|80 - 75|/2) = 0.5 - z(2.5)$.

16. 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)?

> **Solution:** In 16 games, he will average 0.9 TDs/game with a standard deviation of $1/\sqrt{16} = 0.25$. So the probability that he averages at least 1 TD/game is $P(\bar{X} \geq 1) = 0.5 - P(0.9 \leq \bar{X} \leq 1) = 0.5 - z(\frac{|1-0.9|}{0.25}) = 0.5 - z(0.4)$.

17. Suppose that the average shopper spends 100 dollars during Black Friday, with a standard deviation of 50 dollars. What is the probability that a random sample of 25 shoppers will have spent more than \$3000?

> **Solution:** In a sample of 25 shoppers, the average shopper will spend 100 dollars with a standard deviation of $50/\sqrt{25} = 10$. Thus, the probability that a random sample will spend more than 3000 dollars is the probability that a random sample will average more than $3000/25 = 120$ dollars per person. This probability is $0.5 - z(|120 - 100|/10) = 0.5 - z(2)$.

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

> **Solution:** In a class of 25, the average score will be distributed with mean 60 and standard deviation $20/\sqrt{25} = 4$. The probability that they had an average score of at least 66 is $0.5 - z(|66 - 60|/4) = 0.5 - z(1.5)$.