**Math 10B with Professor Stankova**
**Worksheet, Discussion #14; Tuesday, 3/6/2018**
**GSI name: Roy Zhao**

# 1 Covariance

## 1.1 Concepts

1. The **Covariance** is defined as $Cov(X, Y) = E[XY] - E[X]E[Y]$. It measures how "independent" two random variables are. For **independent** random variables, we have $Cov(X, Y) = 0$. Note that we can recover the definition of regular variance because the covariance of a random variable with itself is $Cov(X, X) = E[X^2] - E[X]^2 = Var(X)$. We can update the formula for the variance of the sum of two random variables as $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$ which holds for **all** random variables. Properties that hold for the random variable are:

   - $Cov(X, Y) = Cov(Y, X)$
   - $Cov(X, Y + Z) = Cov(X, Y) + Cov(X, Z)$
   - $Cov(X, cY) = cCov(X, Y)$ for any constant $c$
   - $Cov(X, c) = 0$ for any constant $c$

## 1.2 Example

2. If $Var(X) = 1$ and $Cov(X, Y) = 0$, what is $Cov(X, X + Y)$?

## 1.3 Problems

3. True False The covariance of two random variables is always $\geq 0$.

4. True False For random variables $X, Y$ and constants $c, d$, we have $Cov(X + c, Y + d) = Cov(X, Y)$.

5. If $Var(X) = 4$ and $Var(Y) = 1$ and $Cov(X, Y) = -2$, what is $Var(X + 2Y)$?

## 1.4 Extra Problems

6. If $Var(X) = 1$ and $Var(Y) = 1$ and $Cov(X, Y) = 3$, what is $Var(2X + Y)$?

# 2 Z-Scores, CLT, LoLN

## 2.1 Concepts

7. 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 $\sigma = 1$.

   One key area these pop up in is when taking the average of a bunch of trials. The **Central Limit Theorem (CLT)** tells us that for $X_i$ independent and identically distributed (i.i.d.) (e.g. rolling a die multiple times) with $E[X_i] = \mu$ and $Var(X_i) = \sigma^2$, then the average that we get (e.g. the average number that we roll) is **approximately** normal distributed with mean $\mu$ and standard deviation $\sigma/\sqrt{n}$. 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 order to compute probabilities, we compute the $z$ score. Given a normal distribution with mean $\mu$ and standard deviation $\sigma$, 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 $\mu$. Namely, $E[\bar{X}] \to \mu$ and the probability that we are far away from the mean goes to 0.

## 2.2 Examples

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

9. 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?

## 2.3 Problems

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

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

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

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

14. Let $f$ be normally distributed with mean 5 and standard deviation 2. Calculate the probability $P(X \leq 3)$.

15. Let $f$ be normally distributed with mean 0 and standard deviation 5. Calculate the probability $P(-2 \leq X \leq -1)$.

16. 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?

17. 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?

## 2.4 Extra Problems

18. Let $f$ be normally distributed with mean 3 and standard deviation 5. Calculate the probability $P(X \geq 0)$.

19. Let $f$ be normally distributed with mean 2 and standard deviation 1. Calculate the probability $P(X \leq 0)$.

20. 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)?

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

22. 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?