Math 10B with Professor Stankova
Worksheet, Discussion #15; Thursday, 3/14/2019
GSI name: Roy Zhao

1 Variance and Covariance

1.1 Concepts

1.

Distribution	PMF	$E(X)$	Variance
Uniform	If $\#R(X) = n$, then $f(x) = \frac{1}{n}$ for all $x \in R(X)$.	$\sum_{i=1}^{n} \frac{x_i}{n}$	$\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{n}$
Bernoulli Trial	$f(0) = 1 - p$, $f(1) = p$	p	$Var(X) = p(1-p)$
Binomial	$f(k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Geometric	$f(k) = (1-p)^k p$	$\frac{1-p}{p}$	$Var(X) = \frac{1-p}{p^2}$
Hyper-Geometric	$f(k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$	$\frac{nm}{N}$	$\frac{nm(N-m)(N-n)}{N^2(N-1)}$
Poisson	$f(k) = \frac{\lambda^k e^{-\lambda}}{k!}$	λ	λ

The **Variance** is defined as $Var(X) = E((X-\mu)^2)$. An easier form is $E(X^2) - E(X)^2$. It satisfies some properties:

- $Var(c) = 0$
- $Var(cX) = c^2 Var(X)$
- $Var(X+Y) = Var(X) + V(Y)$ for **independent** random variables.

The **Standard Error** is defined as $SE(X) = \sqrt{Var(X)}$. We use it to get the same units as X.

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 Examples

2. Let X be a uniform random variable with range $\{-1, 0, 1\}$. Let $Y = X^2$. Calculate $Cov(X, Y)$.

1.3 Problems

3. True False For independent random variables X, Y we have $Var(X - Y) = Var(X) - Var(Y)$.
4. True False The product of two Bernoulli trials is another Bernoulli trial.
5. True False If c is a constant, then $Var(X + c) = Var(X)$.
6. True False The covariance of two random variables is always ≥ 0.
7. True False For random variables X, Y and constants c, d, we have $Cov(X + c, Y + d) = Cov(X, Y)$.
8. While pulling out of a box of cookies, what is the expected number of cookies I have to pull out before I pull out an oatmeal raisin if 20% of cookies are oatmeal raisin? What is the variance?
9. I flip a coin some number of times and I expected to see 90 heads with a standard deviation of 3 heads. What is the probability that I actually see 95 heads?
10. I am at a casino and play a game and am expected to gain 10 cents per play with a variance of 1$^2 if I bet $10. What is the expected value and variance when I bet $100 instead?
11. Prove the short cut formula for variance from the definition of variance.
12. Suppose that I flip a fair coin 10 times. Let T be the number of tails I get and H the number of heads. Calculate $E[T], E[H], Var[T], Var[H], Var[T + H]$. Now calculate $E[T - H]$ and $Var[T - H]$.

2 Average of Random Variables

2.1 Concepts

13. 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 μ 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 $Var(\bar{X}) = \sigma^2/n$.

2.2 Examples

14. 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 $\sigma/\sqrt{n}$.

15. 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 expected value and standard deviation for the average height?

2.3 Problems

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

17. True False Suppose that I roll a fair die 100 times. Then since the expected value of the average die roll is 3.5, I will roll a $1, 2$, or 3 50 times and a $4, 5$, or 6 50 times.

18. 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 expected value and variance of the average weight of these babies?

19. Suppose that the average lifespan of a human is 75 years with a standard deviation of 10 years. What is the mean and standard error of the average lifespan of a class of 25 students?

20. 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 expected value and variance for the percentage of the people sampled who support Obama?

21. The newest Berkeley quarterback throws an average of 0.75 TDs/game with a standard deviation of 1. What is his expected value and standard deviation for the number of TDs he throws next season (16 total games)?

22. Let $X_1, \ldots, X_4$ be i.i.d Bernoulli trials with $p = \frac{3}{4}$. Let $\bar{X}$ be the average of them. What is $Var[\bar{X}]$? Find $Cov(X_1, \bar{X})$ (Hint: Write $\bar{X} = \frac{1}{4}(X_1 + X_2 + X_3 + X_4)$).