Class 8: Expected Value and Variance of Random Variables (Text: Section 4.4)

A **random variable** is a number whose value depends on random process. If *X* is a random variable (RV), it is a function whose input is an experiment and whose output is a number, *x*.

EXPECTED VALUE OF A RANDOM VARIABLE

If we draw one of 10 slips of paper from a bowl, with \$0 written on nine of them and \$100 on the tenth, we will win nothing $^9/_{10}$ of the time and a hundred dollars $^1/_{10}$ of the time. On average, over many games, how much do you win per game?

$$\text{Average win per game} =$$

The \$10 is called the **expected value** of the game. Note that \$10 is not a value you can actually win—that is either \$0 or \$100—but it is the average winning per game over a very large number of plays.

Expected value is a weighted average, with more probable values being weighted more heavily.
For a random variable, X, it is written $E(X)$ or μ.
If x_1, x_2 .... are the values that X takes with probabilities p_1, p_2, ... then

$$E(X) = \mu_X = p_1x_1 + p_2x_2 + \cdots + p_nx_n$$

Ex: For these values of the prizes won in the Arizona Lottery,

(a) What is probability of winning nothing?
(b) Find the average win per ticket over many plays. What does it tell you about the price of a ticket?

Lottery	Arizona Pick Five[1]				
Win	\$50,000	\$500	\$5	\$1	\$0
Probability	$^1/_{575,757}$	$^1/_{3387}$	$^1/_{103}$	$^1/_{10}$	0.89

Note:

- Expected value gives the amount the state pays out, on average, per ticket.
- In this case, $E(W)$, is not an amount you can actually win. It is the average value of your winnings, per ticket, over a large number of plays.
- Note that $E(X)$ does not include administrative costs.

[1] http://www.arizonalottery.com/Pick5.html?GameID=6&TokenID=150.135.120.6

<u>Ex: How much should you pay per year for collision insurance for your car?</u> Assume:

- Car is worth $30,000.
- Chance of accident in a year is $1/20$
- If there is an accident, the car is totaled.
- Insurance pays the whole cost of replacing the car.

<u>Ex: How much should you pay per year for collision car insurance for the same car</u> assuming

- Same probability of accident
- If there is an accident, cost of repairs is uniformly distributed between $0 and $30,000.
- Insurance covers all repair costs

<u>Ex: How much should you pay per year for collision car insurance with a $1000 deductible?</u> If there is an accident , the cost of repairs is uniformly distributed between $0 and $30,000. (Same car and probability of accident.)

VARIANCE AND STANDARD DEVIATION OF RANDOM VARIABLES

We define the **standard deviation** of a random variable in much the same way as the standard deviation of the sample: as the average distance from the expected value, with more probable values weighted more heavily.

The **variance** is the square of the standard deviation; it is useful when you need to find the standard deviation of sums and difference of random variables.

Interpretation and Notation:

Standard deviation is a weighted average of the distances from the mean (expected value), with more probable values being weighted more heavily.
For a random variable, X, the standard deviation is written $SD(X)$ or σ;
the variance is $Var(X)$ or $V(X)$ or σ^2.
If X has expected value μ, then the variance is

$$V(X) = \sigma^2 = p_1(x_1 - \mu)^2 + p_2(x_2 - \mu)^2 + \cdots + p_n(x_n - \mu)^2$$

$$SD(X) = \sqrt{V(X)}$$

Ex: Find the variance and standard deviation for Arizona Lottery

GENERAL FORMULAS

For a Discrete Random Variable:

If the p.d.f of the random variable X is given by

Values of X	x_1	x_2	x_3	$\cdots$	x_k
pdf	p_1	p_2	p_3	$\cdots$	p_k

then

Expected value	***Variance***	***Standard deviation***
$E(X) = \mu_x = \sum_{i=1}^{i=k} p_i x_i$	$Var(X) = \sigma_X^2 = \sum_{i=1}^{i=k} p_i (x_i - \mu_x)^2$	$SD(X) = \sigma_X = \sqrt{V(X)}$

For a Continuous Random Variable:

In this course, for a continuous random variable, X ,you will be given $E(X)$ and $V(X)$ or $SD(X)$. If $p(x)$ is the pdf of a continuous random variable X, although these formulas are ***not needed*** in Math 263, they are:

Expected value	***Variance***	***Standard deviation***
$E(X) = \mu_x = \int_{-\infty}^{\infty} xp(x)\,dx$	$Var(X) = \sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_x)^2 p(x)\,dx$	$SD(X) = \sigma_X = \sqrt{V(X)}$

SUMS AND DIFFERENCES of RANDOM VARIABLES

Expected Value:

1. If a and b are numbers $E(a + bX) = a + bE(X)$ that is, $\mu_{a+bX} = a + b\mu_X$
2. If X and Y are random variables $E(X + Y) = E(X) + E(Y)$ that is, $\mu_{X+Y} = \mu_X + \mu_Y$

Variance and Standard Deviation:

1. If a and b are numbers: $\sigma^2_{a+by} = V(a + bY) = b^2V(Y)$
2. If X and Y are independent random variable (so correlation $r = 0$)
$$\sigma^2_{x+y} = V(X + Y) = \sigma^2_x + \sigma^2_y$$
$$\sigma^2_{x-y} = V(X - Y) = \sigma^2_x + \sigma^2_y$$
3. (***Optional***) If X are Y are not independent (so correlation $r \neq 0$)
$$\sigma^2_{x+y} = V(X + Y) = \sigma^2_x + \sigma^2_y + 2r\sigma_x\sigma_y$$
$$\sigma^2_{x-y} = V(X - Y) = \sigma^2_x + \sigma^2_y - 2r\sigma_x\sigma_y$$

Ex: Ear lengths: For women, ear lengths are normally distributed with mean 2.06 inches and standard deviation 0.17 inches; for men they are normally distributed with mean 2.45 inches and standard deviation 0.17 inches

(a) A woman and a man are randomly selected. What are the mean and standard deviation of the difference in their ear lengths?

(b) Assuming that the difference in ear lengths is normally distributed (which it is), what is the probability that the woman's ear length is longer than the man?

(c) What difference does it make if the man and the woman are a couple?

Ex: What is variance and standard deviation of winnings from two lottery tickets, if the winnings are independent, given $\sigma_x = \$66$ and $\sigma_y = \$50$.

Ex: How do people determine how to invest in stock market? (Optional)

Stocks have higher earnings, but carry more risk.
Let X be yearly earnings of an "index fund" which is a combination of many stocks, with $E(X) = 15\%$ and $\sigma(X) = 25\%$. (Earnings are given as a percentage of what you have invested.)

T-bills: Treasury bills earn less, but vary less.
Let Y be yearly earning of a treasury bill, with $E(Y) = 5\%$ and $\sigma(Y) = 3\%$

Suppose the correlation between X and Y is $r = -0.1$

(a) What fraction of investment should go to which to maximize earnings? What is the standard deviation in this case?
(b) What is the expected value if the money is split 50-50 between the stocks and the T-bill? What is the standard deviation in this case?