Chapter 16: Random Variables

So we've talked about variables. And we've talked about things that are random. Now it's time to put the two together.

The Idea

A **Random Variable** measures the (quantitative) result of some random experiment. Pick a person at random, and measure his/her age: you've got a random variable (age). Pick a group of 10 cars and count how many are red: you've got a random variable (number that are red).

Random variables force every experiment to produce a quantitative result…and we can do mathy things with quantities!

We'll focus the discussion (for now) on discrete random variables. Thus, our random variables can only take on certain (discrete) values…which means that we can usually look at the distribution of the variable in the form of a table (remember that a distribution shows what values can occur, and how often each one does occur).

The values of the variable must be disjoint—there can't be any way for the underlying experiment to produce two values at the same time. Also, the values of the variable must be exhaustive—they have to cover every possible outcome. As a result of this, the probabilities of the values must add to one.

A quick note on notation…capital letters refer to the variable itself, and lowercase letters refer to a value of the variable. As usual, don't get too worked up about the notation…we'll keep it down to an absolute minimum.

Examples

[1.] Let *X* represent the number of siblings that a student (of a particular statistics class) has. The probability distribution of *X* is (partially) given below.

Table 1 - Distribution of Siblings

x	**0**	**1**	**2**	**3**	**4**
$P(X = x)$	0.200	0.425	0.275	0.075	???

What is the probability that *X* takes the value 4?

Since this represents all possible outcomes, the probabilities must sum to one—thus, $P(X = 4) = 0.025$.

[2.] Flip three coins—let *H* represent the number of heads. Write down the distribution of *H*.

We've written down a universe for this experiment before—just take that and convert the letters to numbers of heads.

h	0	1	2	3
$P(H = h)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Expected Value

The term expected value is code for the mean. Since these distributions represent all possible values, we use μ to represent the mean of the variable.

Believe it or not, the formula for finding the mean of a discrete random variable is the same as that for the sample mean of a data set…though they look a little different, they really are doing the same thing in the same way.

Equation 1 - The Mean of a Discrete Random Variable

$$E(x) = \mu_X = \sum_{i=1}^{n} x_i \cdot P(X = x_i)$$

The way to think of this formula is that you are multiplying each value of the variable by its probability, and then adding up the results.

Examples

[3.] Find the mean of *X* from example 1.

Even though your calculator will do this for you, you should show the work on the AP Exam.

$\mu_X = (0)(0.2)+(1)(0.425)+(2)(0.275)+(3)(0.075)+(4)(0.025)=1.3$. The mean is 1.3 siblings.

[4.] Find the mean of *Y* from example 2.

$\mu_Y = (0)\left(\frac{1}{8}\right)+(1)\left(\frac{3}{8}\right)+(2)\left(\frac{3}{8}\right)+(3)\left(\frac{1}{8}\right)=1.5$. We can expect an average of 1.5 heads.

[5.] The distribution of occupancy of rental properties in Irmo is given in the table below.

o	1	2	3	4	5	6	7
$P(O=o)$	0.271	0.231	0.218	0.154	0.07	0.045	0.011

Find the mean occupancy of a rental property in Irmo.

$\mu_O = (1)(0.271)+(2)(0.231)+(3)(0.218)+(4)(0.154)+(5)(0.07)$
$+(6)(0.045)+(7)(0.011)=2.7$. There are an average of 2.7 occupants per household.

Variance

We've just measured the center of the variable's distribution—naturally, spread comes next! Since we are deep in the theory, we'll focus on calculating variance. Remember that it only takes one more step to change variance into standard deviation!

Equation 2 - The Variance of a Discrete Random Variable

$$Var(X) = \sigma_X^2 = \sum_{i=1}^{n}\left((x_i - \mu_X)^2 \cdot P(X = x_i)\right)$$

Don't worry too much about that formula…we'll let the calculator do the work for us.

Examples

[6.] Find the variance of *X* from example 1.

Again…even though your calculator will do this for you, you should show the work.

$\sigma_X^2 = (0-1.3)^2(0.2)+(1-1.3)^2(0.425)+(2-1.3)^2(0.275)$

$+(3-1.3)^2(0.075)+(4-1.3)^2(0.025)=0.91$. The variance is 0.91 siblings squared.

[7.] Find the variance of *Y* from example 2.

OK—you get the idea. I'm just going to bang out the rest of them. No, that doesn't mean you get to skip the work.

The variance is 0.75…I suppose the units on that would be heads squared.

[8.] Find the standard deviation of *O* from example 5.

Careful now! This isn't variance this time.

The standard deviation is 1.4967 occupants.

[9.] Roll two dice and let *S* represent the sum of the pips. Find the standard deviation of *S*.

A new one! We need the distribution first. We listed it in a previous chapter…I'm going to take that and convert it to a list of sums.

s	2	3	4	5	6	7	8	9	10	11	12
$P(S=s)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Now, do the math…the standard deviation is 2.4152 pips.

Linear Transformations

There are times when we need to make adjustments to our random variables—perhaps they were measured with the wrong units; maybe we forgot to measure from the zero mark on the ruler, and accidentally added 1cm to each measurement. Fixing these errors (multiplying by 2.54 to change inches to centimeters; subtracting one) are examples of **Linear Transformations of Random Variables**.

Adding a Constant

If *Y* is a random variable, then $Y+1$ is also a random variable, where each datum of the set has 1 added to it. Yes, there are situations in real life where this is needed!

What will this adjustment do to the mean? Think about it for a second…isn't it obvious? The new mean will be one greater than the old mean! It's not hard to prove if you know how to deal with summation notation—a topic typically introduced in PreCalculus.

At any rate, this means that in general, adding a constant to a random variable will add that same constant to the mean of the variable.

Equation 3 - The Effect of Adding a Constant on the Mean

$$\mu_{X+a} = \mu_X + a$$

What about spread? That's a little harder, but not much. Because of the subtraction inside of the formula for variance, and because the one gets added to both the variable and the mean of the variable, it turns out that adding one does not change the variance.

In general, adding a constant to a random variable will not change the variance.

Equation 4 - The Effect of Adding a Constant on the Variance

$$\sigma_{X+a}^2 = \sigma_X^2$$

Multiplying by a Constant

If *X* is a random variable, then $3X$ is also a random variable, where each datum of the set is multiplied by 3.

Be careful! In Algebra, $3X = X + X + X$ …but in statistics, this isn't necessarily true. In statistics, $3X$ means to multiply every value of the variable by 3, but $X + X + X$ means to repeat the experiment three times and add the results. Those are not the same thing!

Back to the main thought…what will that multiplication do to the mean? Think about it for a second…

That's right! The mean will get multiplied by three also.

In general, multiplying a variable by a constant will also multiply the mean.

Equation 5 - The Effect of Multiplying by a Constant on the Mean

$$\mu_{bX} = b \cdot \mu_X$$

Now, what about variance? Again, this isn't quite the same as the mean…it turns out that the additional constant gets factored out. However, remember that there is a squaring operation in that formula. Thus, multiplying a variable by a constant will multiply the variance by the square of the constant.

Equation 6 - The Effect of Multiplying by a Constant on the Variance

$$\sigma_{bX}^2 = b^2 \cdot \sigma_X^2$$

Now, a word of caution…when you take the square root to find standard deviation, what happens to that b^2?

If you're being really careful, then you know that $\sqrt{b^2} = |b|$. This should make sense…multiplying every datum in the data by a negative *could not* make the standard deviation zero! Thus:

Equation 7 - The Effect of Multiplying by a Constant on the Standard Deviation

$$\sigma_{bX} = |b| \cdot \sigma_X$$

Examples

[10.] Back in example 1, *X* represented the number of siblings a student had…so if you add one, you should get the number of children in the household. Let *T* be the total number of children in the household.

Find the mean and variance of *T*.

The mean will be one larger—specifically, 2.3 people. The variance will be unchanged at 0.91 people squared.

[11.] Dave works at a tax firm where the charge for the service depends on the amount of your refund. In particular, the firm charges 1% of the refund amount plus a fixed fee of $250.00. The amount of client refunds varies with mean $2545 and standard deviation $550. What are the mean and standard deviation of the firm's service charges?

Let's write that out with actual variables…let *R* be the amount of the refund, and let *F* be the amount of the fees. We're given $\mu_R = 2545$ and $\sigma_R = 550$, and the relationship between the variables is $F = 0.01 \cdot R + 250$. Thus, $\mu_F = 0.01 \cdot \mu_R + 250 = 275.45$ dollars. For the standard deviation, we first need the variance: $\sigma_F^2 = \left(0.01\right)^2 \cdot \sigma_R^2 = 30.25$, so the standard deviation is $\sigma_F = \sqrt{30.25} = 5.5$ dollars.

[12.] Zeke did some measuring of the distances of North American rivers for his Geography class. The lengths vary with mean 591.18 miles and standard deviation 493.87 miles. His teacher was upset that Zeke measured in miles and wants everything converted to kilometers (the conversion is 1 mile is approximately 1.61 kilometers). What are the mean and standard deviation of Zeke's measurements once they are converted to kilometers?

The mean in kilometers will be $\left(1.61\right)\left(591.18\right) = 951.8$, and the standard deviation in kilometers will be $\left(1.61\right)\left(493.87\right) = 795.1307$.

Linear Combinations

In addition to transforming single variables, it is often desired to combine different random variables…for example, we may want to add one variable that measures the distance that people travel for Spring Break with another variable that measures the distance that people travel over the Thanksgiving Break.

The question is: what happens to the means and variances of the original variables? How do they combine to find the mean and variance of the new combined variable?

So far, transforming the mean has been easy…if you add two variables, could the new mean just be the sum of the original means? Could it be that simple?

It is.

Equation 8 - The Mean of a Linear Combination of Random Variables

$$\mu_{X+Y} = \mu_X + \mu_Y$$

Variance has been the problem…it's never worked out as nicely as the mean. Thus, it's not hard to guess that the variance of the combined variable won't just be the sum of the original variances.

…except that this is exactly what happens! Isn't that great?

Of course, there is one caveat: this only works if the variables are independent. The reason, it turns out, is Algebra…specifically, binomial expansion.

Humor me for a moment…work out $\left(X+Y\right)^2$.

If you wrote $\left(X+Y\right)^2 = X^2 + Y^2$, then you've got problems…you've forgotten how to expand a binomial raised to a power.

For the record, the correct answer is $\left(X+Y\right)^2 = X^2 + 2XY + Y^2$. Each of these pieces represents something in the problem: $\left(X+Y\right)^2$ is the variance of the combined variable, X^2 and Y^2 are the variances of the original variables, and $2XY$ represents something called covariance—a measure of how strongly two variables vary together. If two variables are independent, then their covariance is zero…which gives us the nice result.

So—as long as two variables are independent, their variances add.

Equation 9 - The Variance of a Linear Combination of Independent Random Variables

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$

I'll leave it to you to think about what happens when you subtract two variables…I will remind you that $X - Y = X + \left(-1\right) \cdot Y$.

Examples

[13.] Three dice are rolled and the sum of the pips found. What is are the mean and standard deviation of the sum?

Let's figure out the mean and variance for one die and go from there. The values of the variable are one through six, and the probabilities are each $\frac{1}{6}$, which makes $\mu_D = 3.5$ and $\sigma_D^2 = 2.9167$. The sum of three dice means $S = D + D + D$, so $\mu_S = \mu_D + \mu_D + \mu_D = 10.5$. Furthermore, the variances will add, so $\sigma_S^2 = \sigma_D^2 + \sigma_D^2 + \sigma_D^2 = 8.75$. Thus, the standard deviation will be $\sqrt{8.75} = 2.958$.

All of this assumes that the dice are independent…but that's a fairly safe bet.

[14.] Student scores on the Math section of the SAT vary with mean 500 and standard deviation 100. Scores on the Critical Reading section also vary with mean 500 and standard deviation 100. Studies suggest that these score are well correlated—high scores in one area tend to be paired with high scores in the other (and vice versa, with respect to low scores).

Find the mean and standard deviation of the composite score (m + cr).

A-ha! The answer is that the mean will be 1000, and that the standard deviation cannot be found. The given information suggests that the variables are not independent, so we can only find the mean—not the standard deviation.

[15.] A study of patients with Parkinson's disease measured the effects of two treatments (lines on the ground, and a flashing light) on their cadence (steps per minute). The study found that patients using the lines had a cadence that varied with mean 105 and standard deviation 14.1, and patients using the flashing light had cadences that varied with mean 112 and standard deviation 11.7. No patient was tested using both treatments.

Find the mean and standard deviation of the difference in cadences for the two treatments.

Since no patient experienced both treatments, it's probably a fair bet that the variables are independent. Let *L* represent the line treatment, and let *F* represent the flashing light treatment. Furthermore, let's subtract $F-L$. The mean of the new variable is $\mu_{F-L}=112-105=7$ steps per minute. The variance will be $\sigma_{F-L}^2=\sigma_F^2+\sigma_L^2=(14.1)^2+(11.7)^2=335.7$, so the standard deviation will be $\sigma_{F-L}=\sqrt{335.7}=18.322$ steps per minute.

A few of you are wondering why I didn't subtract the variances…here's a little hint: $\sigma_{F-L}^2=\sigma_{F+(-1)(L)}^2=\sigma_F^2+(-1)^2\sigma_L^2$.

Continuous Random Variables

We've focused on discrete random variables…the textbook goes ahead and throws in the continuous random variables also, but I'm going to save that for later. Continuous random variables are wildly important…but let's digest what we've got here first.