Expected Values May 4, 2001

Juana: There's some pdf "cleanup" I need you to do before talking about expected values:

PDFs

All of the pdfs we talked about on Wednesday were examples of pdfs for discrete valued random variables. These pdfs explicitly assign a probability for every value of the random variable. That might do so with a table:

X = payoff on lottery

x	Prob
0	.25
$1	.25
$1.50	.20
$2	.10
$5	.15
$10	.05

Or a function:
X = number of children under the "one-boy" rule:
$P(X = x) = (1/2)^x$

There are some rules these pdfs must follow:
• probabilities are always between 0 and 1
• the sum over all possible values must equal 1

Sum the probabilities in the table, and you'll see this. It's harder to see for the function given because the sum goes from 1 to infinity. But if you took a calculus course, you'd learn how to do this, and would see that Sum(x = 1 to x = infty) = 1.

For continuous random variables, the situation is more difficult. The reason is that it is impossible to list ALL of the values of a continuous random variable. (Of course, it's impossible to list all of the value for the one-boy rule, too, but for a continuous variable, you can't even begin.)

for example,: X = time spent waiting in line at the coffee shop.

We might know that X will have values inbetween 0 minutes and, say, 15 minutes. But we can't just write down the values: 0, 1 min, 2, min. etc. because we've left out all the values inbetween 0 and 1, say. For example, .5. No matter which values you write down, you'll always be leaving infinitely many out.

A side effect of this is that, if X is continuous, then P(X =x) = 0 for any value of x. We'll get back to this later.

Recall: One-Boy Policy example.
China had a one-boy policy in some areas that meant that couples could have children until they had their first boy. Let X = number of children in such a family. We found last time that P(X = 2) = 1/4. Also, P(X = x) = $1/2^{x}$.

The interpretation of this probability is that if there were infinitely many couples under this policy, or it was carried out for infinitely many years, than 1/4 of all families would have two children.

Of course, there are not infinitely many. So in practice, the true proportions might differ slightly.

If we did a random sample of, say, 50 such families, we would expect the proportion of two-children families to be close to 1/4.

In our simulation of a random sample of 50 families, we found that about 18% had two children. This is not very close to 25%, but then this was a small sample.

Now let's change the question:

> In theory, how many children should the typical family have under this policy? Or, put differently, what's the typical theoretical number of children in these families?

This question is very similar to asking: What's the typical value on a list of data? Only instead of asking that question about a list of numbers, we're now asking it about a theoretical population.

The answer, of course, depends on what we mean by "typical." There are different ways of answering that question here, just as there were with data sets. But here's what we'll mean:

The Expected Value is the mean of the probability distribution function.
The mean is the "balancing point" of the pdf, just as the sample mean was the "balancing point" of a histogram.

The sample mean was computed by taking the average of the values in a list of numbers.

But the expected value or theoretical mean (also called population mean) is found by :

sum p(x)*x

that is, the value of the random variable times the probability of that value, all summed up.

Notation: E(X) is the "mean of X", and so is mu_x.

Examples:

I. Flip a coin three times. Let X be the number of heads. The pdf, we found last time, is

x	p(x)
0	1/8
1	3/8
2	3/8
3	1/8

What is the expected number of heads?

E(X) = 0*1/8 + 1 * (3/8) + 2*(3/8) + 3*(1/8) = 12/8 = 1.5

NOTE: Make a graph of PDF and show where mu_x occurs.

Other ways of asking the same question:

How many heads to you expect? Find E(X), find the mean of X, find the mean number of heads.

Note that we are asking about the expected number of heads IN REFERENCE TO A THEORETICAL MODEL. The number *you* get if you flip a coin 3 times will NOT be 1.50, I can guarantee you.

Interpretation of the expected value or mean:

Think of it as a long-run average. If you did this experiment (flip fair coin 3 times) MANY many times, then in the long run you'd have about 1.50 heads per experiment.

The term "expected value" comes from the gambling tradition. Here's an example:

Example II. American Roullette has a wheel with 38 numbers: 00, 0, 1,2,...,38. The 0 and 00 slot are green, half of the remaining numbers are red, half are black. One bet consists of placing $1 on either red or black. The wheel is spun and if the ball lands on the color you predicted, you are paid $2. If not, you are paid 0. Thus, if you "win", you're net is $1 and if you lose it is -$1.

Find your expected winnings. (Or "your mean winnings".)

First you need a random variable: X = amount you win.

Next, you need the pdf:

The values of X are -1,1

x	p(x)
-1	20/38
1	18/38

E(X) = -1(20/38) + 1(18/38) = -2/38 = -0.0526

In other words, you can "expect" to lose about 5 cents.

How is this possible, if you've only played once? After all, on a single spin you either win $1 or lose $1. So how can you lose 5 cents?

The answer is in the interpretation of expected value. It's a long-run value: after infinitely many plays, you'll have lost, on average 5 cents per play.

Note: It might seem strange that your "winnings" are a negative number (-0.05). This just means that you lost that amount.

This is what keeps casinos alive. You'll lose "only" 5 cents on average. But they'll gain 5 cents on average. At any given moment, probably thousands of people have placed a bet on red. The casino just made an average of 5 cents per person. Multiplied by many people over many nights, this is not small change.

Ex III. A class room has students of the following ages:

age	percent of class
18	50
19	25
20	15
21	10

Choose a student at random. What's the expected age?

Note that the percent of the class is the same as the probability of getting a student of that age. So that if X = age of randomly select student, then the table gives the pdf of X.
So E(X) = 18(.50) + 19*.25 + 20 *(.15) + 21 * .10 = 18.85.

Ex IV: Box Models

Often, when thinking about surveys or random samples, it is useful to think of the population as consisting of tickets in a box, and our random sample consists of drawing tickets out of the box.

For example, the last problem could be modelled as follows:
Make a box with 50 tickets with "18", 25 with "19", 15 with "20", and 5 with "21". Then draw a ticket at random. What's the expected value of the ticket? The probability of selecting an "18" is now 50/100, a "19" is 25/100, etc.

The answer is: (50/100)*18 + (25/100)*19 + etc = 17.8.

The answer is the same, of course, but note the following points:

- The box represents our population. Hence you can see that the expected value is the mean of the population.
- the expected value can also be computed just by taking the average of the tickets.

(Because: .50*18 = 50/100 * 18. We used the first version, .50*18 in Ex III, and we use the second here.)

How is the Population Mean related to the Sample Mean (the average)?

Suppose the classroom mentioned above is very big. (Thousands of students!) Here's how we might collect some data about the age of the class:
Design a survey to select 10 people at random from the class. This is identical to drawing 10 tickets with replacement from our box. Then we could make a histogram of their age, and compute the average age. The population mean is 17.8 (the average of the box.) The sample mean is the average of the 10 tickets.

A result called the Law of Large Numbers says that the sample average will be "close to" the population mean, as long as the sample was made from random, independent selections from the population. "Close to" is a subjective term, but (as we'll see later), we can get a pretty good idea of how close. For now, it's important to know that the larger the sample size (the more tickets you draw), the closer you'll get.

It's also important to notice the sample average is not always the same distance from the population mean. Sometimes its close, sometimes its far. But it tends to be -- it has a high probability of being -- closer rather than farther.

SPECIAL CASE

There is a certain type of situation that occurs so frequently that we have a short-cut way of dealing with it. Remember the experiment in which we flipped a coin 3 times and counted the number of heads? We could make a box model of that situation that looks like this:
Box: 1 "0" ticket; 3 "1" tickets, 3 "2" tickets, 1 "3" ticket.
Note that there are 8 tickets, so the probability of drawing a "1" ticket is the same as the probability of getting 1 heads in 3 tosses. The average of the box is 12/8 = 1.5, and this is the expected value if we reach in and randomly select a ticket.

Another way of doing the simulation, though, is to make a box with just two tickets:
0, 1.
We randomly select 3 tickets with replacement, and add the sum of the tickets. What's the expected value of this sum?

There's a simple formula:

If the only values on the tickets are "1" and "0" AND
If draws are done with replacement AND
If a predetermined number of draws are done AND
If X is the number of "1" tickets selected, **then**

E(X) = expected sum of tickets = expected number of "1" tickets = n*p where
n is the number of draws
p is the average of the box (and also the probability of selecting a "1").

So in our coin example: n = 3, p = .5, np = 1.5

If you draw 10 tickets with replacmenet, n = 10, p = .5, np = 5. So if you flip a coin 10 times, you expect 5 heads.

Random Variables with this structure are called Binomial Random Variables. They occur whenever:

1)The outcome of a trial is either success or failure (1 or 0) AND
2)There is a fixed number of trials AND
3)The trials are independent AND
4)You are interested in the number of successes THEN
a binomial random variable is a good model.

Example: A basketball player shoots 10 free throws. The probability of making a single throw is .25. How many should we expect? np = 10*.25 = 2.5.

Example: 30% of the people in LA will vote for Hahn for mayor. If we take a random sample with replacement of 50 people, how many should we expect to be Hahn supporters? 50*.30 = 15.

Example: Flip a coin until the first head appears. How many flips can we expect it to take? Can't apply this rule, because this is not binomial: there is no fixed n, for one. Also, we are not counting the number of successes, but instead the number of trials until the first success.

Example: Let X be the number of days of rain in a week. The probability of rain on any given day in LA is .10. How many days of rain can we expect in a week? Can't apply the binomial because successive days of rain are not independent. If it rains one day, it is more likely to rain the next.

SD:
sigma = square root of the sum of $(x - mu)^2$ p(x)

Classroom Example:
sigma = square root of $(18 - 18.85)^2 * .5 + (19 - 18.85)^2 * .25 + ..$
= 1.01366

Roulette
sigma = squareroot of $(-1 - -0.0526)^2 * (20/38) + (1 - -0.0526)^2 * (18/38) = 0.756$

If X is a binomial random variable:
the formula above simplifies to
sigma = sqrt(n * p * (1-p)) = sqroot of (number of trials * prob. of success * prob. of failure)

Example
Toss fair coin 3 times. X is number of heads. E(X) = np = 1.5
sigma = sqrt(3 * .5 * .5) = .5

Toss fair coin 100 times. E(X) = 50, sigma = 5

If pdf is symmetric, then about 68% of all values lie within 1 SD of mu
95% within 2, 99.7% within 3.

So if we flip a coin 100 times, 68% of the time we'll see between 45 and 55 heads, and almost always between 35 and 65 heads.