MAT1830

Lecture 24: Expectation and variance

A standard die is rolled some number of times and the average of the rolls is calculated. If the die is rolled only once this average is just the value rolled and is equally likely to be 1, 2, 3, 4, 5 or 6. If the die is rolled ten times, then the average might be between 1 and 2 but this is pretty unlikely – it's much more likely to be between 3 and 4. If the die is rolled ten thousand times, then we can be almost certain that the average will be very close to 3.5. We will see that 3.5 is the expected value of a random variable representing the die roll.

24.1 Expected value

When we said "average" above, we really meant "mean". Remember that the *mean* of a collection of numbers is the sum of the numbers divided by how many of them there are. So the mean of $x_1, \ldots, x_t$ is $\frac{x_1+\cdots+x_t}{t}$. The mean of 2,2,3 and 11 is $\frac{2+2+3+11}{4} = 4.5$, for example.

The expected value of a random variable is calculated as a weighted average of its possible values.

If X is a random variable with distribution

$$\begin{array}{r||c|c|c|c} x & x_1 & x_2 & \cdots & x_t \\ \hline \Pr(X = x) & p_1 & p_2 & \cdots & p_t \end{array},$$

then the *expected value* of X is

$$\mathrm{E}[X] = p_1 x_1 + p_2 x_2 + \cdots + p_t x_t.$$

Example. If X is a random variable representing a die roll, then

$$\mathrm{E}[X] = \frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \cdots + \frac{1}{6} \times 6 = 3.5.$$

Suppose you pay \$1 for a lottery ticket that has a $\frac{1}{5\,000\,000}$ chance of winning \$1 000 000. Let X be you net winnings (in dollars).

Then the probability distribution of X is

$$\begin{array}{c||c|c} x & -1 & 999\,999 \\ \hline \Pr(X=x) & \frac{4\,999\,999}{5\,000\,000} & \frac{1}{5\,000\,000} \end{array}.$$

Then mean of the possible outcomes is $\frac{-1+999\,999}{2} = 499\,999$.

But if we weight the average according the probabilities we get

$$\frac{4\,999\,999}{5\,000\,000} \times -1 + \frac{1}{5\,000\,000} \times 999\,999 = -\frac{4}{5}.$$

$-\frac{4}{5}$ is the **expected value** of X.

24.1 Expected value

When we said "average" above, we really meant "mean". Remember that the *mean* of a collection of numbers is the sum of the numbers divided by how many of them there are. So the mean of $x_1, \ldots, x_t$ is $\frac{x_1+\cdots+x_t}{t}$. The mean of 2,2,3 and 11 is $\frac{2+2+3+11}{4} = 4.5$, for example.

The expected value of a random variable is calculated as a weighted average of its possible values.

If X is a random variable with distribution

x	x_1	x_2	$\cdots$	x_t
$\Pr(X = x)$	p_1	p_2	$\cdots$	p_t

,

then the *expected value* of X is

$$\mathrm{E}[X] = p_1 x_1 + p_2 x_2 + \cdots + p_t x_t.$$

Example. If X is a random variable representing a die roll, then

$$\mathrm{E}[X] = \frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \cdots + \frac{1}{6} \times 6 = 3.5.$$

Question Let Y be a random variable with distribution

$$\begin{array}{c||c|c|c|c} y & 0 & 1 & 2 & 3 \\ \hline \Pr(Y = y) & \frac{1}{8} & \frac{1}{4} & \frac{3}{8} & \frac{1}{4} \end{array}.$$

What is $\mathrm{E}[Y]$?

Answer

$$\begin{aligned} \mathrm{E}[Y] &= \tfrac{1}{8} \times 0 + \tfrac{1}{4} \times 1 + \tfrac{3}{8} \times 2 + \tfrac{1}{4} \times 3 \\ &= 0 + \tfrac{1}{4} + \tfrac{3}{4} + \tfrac{3}{4} \\ &= \tfrac{7}{4} \end{aligned}$$

Example. Someone estimates that each year the share price of Acme Corporation has a 10% chance of increasing by \$10, a 50% chance of increasing by \$4, and a 40% chance of falling by \$10. Assuming that this estimate is good, are Acme shares likely to increase in value over the long term?

We can represent the change in the Acme share price by a random variable X with distribution

x	-10	4	10
$\Pr(X = x)$	$\frac{2}{5}$	$\frac{1}{2}$	$\frac{1}{10}$

Then

$$\mathrm{E}[X] = \frac{2}{5} \times -10 + \frac{1}{2} \times 4 + \frac{1}{10} \times 10 = -1$$

Because this value is negative, Acme shares will almost certainly decrease in value over the long term.

Notice that it was important that we weighted our average using the probabilities here. If we had just taken the average of -10, 4 and 10 we would have gotten the wrong answer by ignoring the fact that some values were more likely than others.

Flux Exercise

Let X be a random variable.

A. The expected value of X is the value that X is most likely to take.

B. The probability of X being above its expected value equals the probability that X is below its expected value.

C. The expected value of X is always positive.

D. All of the above.

E. None of the above.

Answer E

A is wrong: It may not even be **possible** for the value of a random variable to equal its expected value!

B and C are wrong: If you buy a lottery ticket your expected value is negative, and your probability of being below your expected value is much higher than your probability of being above it.

24.2 Law of large numbers

Our initial die-rolling example hinted that the average of a large number of independent trials will get very close to the expected value. This is mathematically guaranteed by a famous theorem called the *law of large numbers*.

Let $X_1, X_2, \ldots$ be independent random variables, all with the same distribution and expected value μ. Then

$$\lim_{n \to \infty} \frac{1}{n}(X_1 + \cdots + X_n) = \mu.$$

24.3 Linearity of expectation

We saw in the last lecture that adding random variables can be difficult. Finding the expected value of a sum of random variables is easy if we know the expected values of the variables.

> If X and Y are random variables, then
> $$\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y].$$

This works even if X and Y are not independent.

Similarly, finding the expected value of a scalar multiple of a random variable is easy if we know the expected value of the variable.

> If X is a random variable and $s \in \mathbb{R}$, then
> $$\mathrm{E}[sX] = s\mathrm{E}[X].$$

Example. Two standard dice are rolled. What is the expected total?

Let X_1 and X_2 be random variables representing the first and second die rolls. From the earlier example $\mathrm{E}[X_1] = \mathrm{E}[X_2] = 3.5$ and so

$$\mathrm{E}[X_1 + X_2] = \mathrm{E}[X_1] + \mathrm{E}[X_2] = 3.5 + 3.5 = 7.$$

Example. What is the expected number of '11' substrings in a binary string of length 5 chosen uniformly at random?

For $i = 1, \dots, 4$, let X_i be a random variable that is equal to 1 if the ith and $(i+1)$th bits of the string are both 1 and is equal to 0 otherwise. Then $X_1 + \cdots + X_4$ is the number of '11' substrings in the string. Because the bits are independent, $\Pr(X_i = 1) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$ and $\mathrm{E}[X_i] = \frac{1}{4}$ for $i = 1, \dots, 4$. So,

$$\mathrm{E}[X_1 + \cdots + X_4] = \mathrm{E}[X_1] + \cdots + \mathrm{E}[X_4] = \frac{4}{4} = 1.$$

Note that the variables $X_1, \dots, X_4$ in the above example were not independent, but we were still allowed to use linearity of expectation.

Question What is the expected number of '11' substrings in a binary string of length 5 chosen uniformly at random?

Answer Let X_1, X_2, X_3, X_4 be random variables such that
$X_1 = 1$ if the string is $1\ 1 * * *$ and $X_1 = 0$ otherwise,
$X_2 = 1$ if the string is $* \ 1\ 1 * *$ and $X_2 = 0$ otherwise,
$X_3 = 1$ if the string is $* * 1\ 1 *$ and $X_3 = 0$ otherwise,
$X_4 = 1$ if the string is $* * * 1\ 1$ and $X_4 = 0$ otherwise.

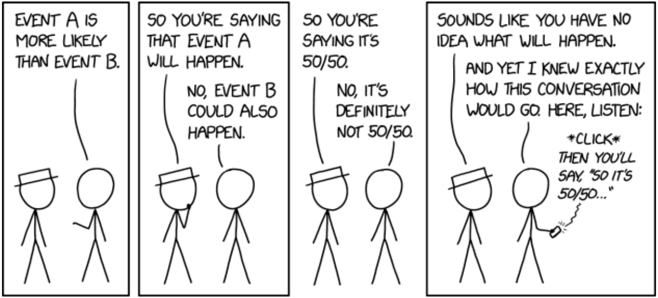
Then $X_1 + X_2 + X_3 + X_4$ is the number of '11' substrings in the string.

So we want to find $\mathrm{E}[X_1 + X_2 + X_3 + X_4]$.

Now $\Pr(X_1 = 1) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.
So $\mathrm{E}[X_1] = \frac{1}{4} \times 1 + \frac{3}{4} \times 0 = \frac{1}{4}$.
Similarly, $\mathrm{E}[X_2] = \frac{1}{4}$, $\mathrm{E}[X_3] = \frac{1}{4}$ and $\mathrm{E}[X_4] = \frac{1}{4}$.

By linearity of expectation,
$\mathrm{E}[X_1 + X_2 + X_3 + X_4] = \mathrm{E}[X_1] + \mathrm{E}[X_2] + \mathrm{E}[X_3] + \mathrm{E}[X_4] = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$.

This works even though X_1, X_2, X_3, X_4 are definitely not independent!


EVENT A IS MORE LIKELY THAN EVENT B.
SO YOU'RE SAYING THAT EVENT A WILL HAPPEN.
NO, EVENT B COULD ALSO HAPPEN.
SO YOU'RE SAYING IT'S 50/50.
NO, IT'S DEFINITELY NOT 50/50.
SOUNDS LIKE YOU HAVE NO IDEA WHAT WILL HAPPEN.
AND YET I KNEW EXACTLY HOW THIS CONVERSATION WOULD GO. HERE, LISTEN:
CLICK
THEN YOU'LL SAY, "SO IT'S 50/50..."

24.4 Variance

Think of the random variables X, Y and Z whose distributions are given below.

x	-1	99
$\Pr(X = x)$	$\frac{99}{100}$	$\frac{1}{100}$

y	-1	1
$\Pr(Y = y)$	$\frac{1}{2}$	$\frac{1}{2}$

z	-50	50
$\Pr(Z = z)$	$\frac{1}{2}$	$\frac{1}{2}$

These variables are very different. Perhaps X corresponds to buying a raffle ticket, Y to making a small bet on a coin flip, and Z to making a large bet on a coin flip. However, if you only consider expected value, all of these variables look the same – they each have expected value 0.

To give a bit more information about a random variable we can define its *variance*, which measures how "spread out" its distribution is.

If X is a random variables with $\mathrm{E}[X] = \mu$,

$$\mathrm{Var}[X] = \mathrm{E}[(X - \mu)^2].$$

So the variance is a measure of how much we expect the variable to differ from its expected value.

Unpacking that definition we get...

If X is a random variable with distribution

x	x_1	x_2	$\cdots$	x_t
$\Pr(X = x)$	p_1	p_2	$\cdots$	p_t

,

and $\mathrm{E}[X] = \mu$, then the variance of X is

$$\mathrm{Var}[X] = p_1(x_1 - \mu)^2 + p_2(x_2 - \mu)^2 + \cdots + p_t(x_t - \mu)^2.$$

Example Let Y be a random variable with distribution

y	0	2	3
$\Pr(Y = y)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

Then $\mathrm{E}[Y] = \frac{1}{6} \times 0 + \frac{1}{2} \times 2 + \frac{1}{3} \times 3 = 0 + 1 + 1 = 2$.
So $\mathrm{Var}[Y] = \frac{1}{6} \times (0-2)^2 + \frac{1}{2} \times (2-2)^2 + \frac{1}{3} \times (3-2)^2 = \frac{2}{3} + 0 + \frac{1}{3} = 1$.

x	-1	99
$\Pr(X = x)$	$\frac{99}{100}$	$\frac{1}{100}$

y	-1	1
$\Pr(Y = y)$	$\frac{1}{2}$	$\frac{1}{2}$

z	-50	50
$\Pr(Z = z)$	$\frac{1}{2}$	$\frac{1}{2}$

Example. The variable X above will be 1 smaller than its expected value with probability $\frac{99}{100}$ and will be 99 larger than its expected value with probability $\frac{1}{100}$. So

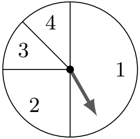
$$\mathrm{Var}[X] = \frac{99}{100} \times (-1)^2 + \frac{1}{100} \times 99^2 = 99.$$

Similarly,

$$\begin{aligned} \mathrm{Var}[Y] &= \tfrac{1}{2} \times (-1)^2 + \tfrac{1}{2} \times 1^2 = 1 \\ \mathrm{Var}[Z] &= \tfrac{1}{2} \times (-50)^2 + \tfrac{1}{2} \times 50^2 = 2500. \end{aligned}$$

Notice that the variance of X is much smaller than the variance of Z because X is very likely to be close to its expected value whereas Z will certainly be far from its expected value.

Question



24.2 Let X be the sum of 1000 spins of the spinner from Lecture 21, and let Y be 1000 times the result of a single spin. Find $\mathrm{E}[X]$ and $\mathrm{E}[Y]$.

Answer

Let S be the result of one spin. So $Y = 1000S$.

Then $\mathrm{E}[S] = \frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{8} \times 3 + \frac{1}{8} \times 4 = \frac{15}{8} = 1.875$.

So $\mathrm{E}[Y] = \mathrm{E}[1000S] = 1000\mathrm{E}[S] = 1875$.

We won't write down the distribution of X (why not?) but we don't need to. Since $X = S_1 + S_2 + \cdots + S_{1000}$ where each S_i has the same distribution as S, we can use linearity of expectation to find that
$\mathrm{E}[X] = \mathrm{E}[S_1] + \mathrm{E}[S_2] + \cdots + \mathrm{E}[S_{1000}] = 1000\mathrm{E}[S] = 1875$.

Flux Exercise

Let X be the sum of 1000 spins of the spinner and let Y be 1000 times the result of a single spin. Which of X and Y do you think would have greater variance?

A. X will have greater variance than Y

B. Y will have greater variance than X

C. X and Y will have equal variance

Answer B

Y will have much larger variance than X, since it has a high probability of being far from its expected value.

Example. Let X be a random variable with distribution given by

x	0	2	6
$\Pr(X = x)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

.

Then the expected value of X is

$$\mathrm{E}[X] = \frac{1}{6} \times 0 + \frac{1}{2} \times 2 + \frac{1}{3} \times 6 = 3.$$

So, the variance of X is

$$\mathrm{Var}[X] = \frac{1}{6} \times (0-3)^2 + \frac{1}{2} \times (2-3)^2 + \frac{1}{3} \times (6-3)^2 = 5$$


LIKELIHOOD YOU WILL GET CODE WORKING
BASED ON HOW YOU'RE SUPPOSED TO INSTALL IT:
VERY LIKELY
APP STORE OR PACKAGE MANAGER
GITHUB LINK
SOURCEFORGE LINK
GEOCITIES/TRIPOD LINK
COPY-AND-PASTE EXAMPLE CODE FROM PAPER'S APPENDIX
ANYTHING THAT "REQUIRES ONLY MINIMAL CONFIGURATION AND TWEAKING"
UNLIKELY

Question

24.3 Let X be the number of heads occurring when three fair coins are flipped. Find $\mathrm{E}[X]$ and $\mathrm{Var}[X]$.

Answer The distribution of X is

x	0	1	2	3
$\Pr(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

So $\mathrm{E}[X] = \frac{1}{8} \times 0 + \frac{3}{8} \times 1 + \frac{3}{8} \times 2 + \frac{1}{8} \times 3 = \frac{3}{2}$.

$$\begin{aligned}\mathrm{Var}[X] &= \tfrac{1}{8} \times (0 - \tfrac{3}{2})^2 + \tfrac{3}{8} \times (1 - \tfrac{3}{2})^2 + \tfrac{3}{8} \times (2 - \tfrac{3}{2})^2 + \tfrac{1}{8} \times (3 - \tfrac{3}{2})^2 \\ &= \tfrac{1}{8} \times \tfrac{9}{4} + \tfrac{3}{8} \times \tfrac{1}{4} + \tfrac{3}{8} \times \tfrac{1}{4} + \tfrac{1}{8} \times \tfrac{9}{4} = \tfrac{3}{4}.\end{aligned}$$