Discrete Structures for Computer Science

William Garrison
bill@cs.pitt.edu
6311 Sennott Square

Lecture #26: Expected Value

Based on materials developed by Dr. Adam Lee

What is a random variable?

Definition: A random variable is a function X from the sample space of an experiment to the set of real numbers **R**. That is, a random variable assigns a real number to each possible outcome.

Note: Despite the name, X is <u>not</u> a variable, and is <u>not</u> random. X is a function!

Example: Suppose that a coin is flipped three times. Let X(s) be the random variable that equals the numbers of heads that appear when s is the outcome. Then X(s) takes the following values:

- X(HHH) = 3
- X(HHT) = X(HTH) = X(THH) = 2
- X(TTH) = X(THT) = X(HTT) = 1
- X(TTT) = 0

Random variables and distributions

Definition: The distribution of a random variable X on a sample space S is the set of pairs (r, p(X=r)) for all $r \in X(S)$, where p(X=r) is the probability that X takes the value r.

Note: A distribution is usually described by specifying p(X=r) for each $r \in X(S)$

Example: Assume that our coin flips from the previous slide were all equally likely to occur. We then get the following distribution for the random variable X:

- p(X=0) = 1/8
- p(X=1) = 3/8
- p(X=2) = 3/8
- p(X=3) = 1/8

Many times, we want to study the expected value of a random variable

Definition: The expected value (or expectation) of a random variable X(s) on the sample space S is equal to:

$$E(X) = \sum_{s \in S} p(s)X(s)$$

For every outcome…

… use the probability of that outcome occuring…

… to weight the value of the random variable for that outcome.

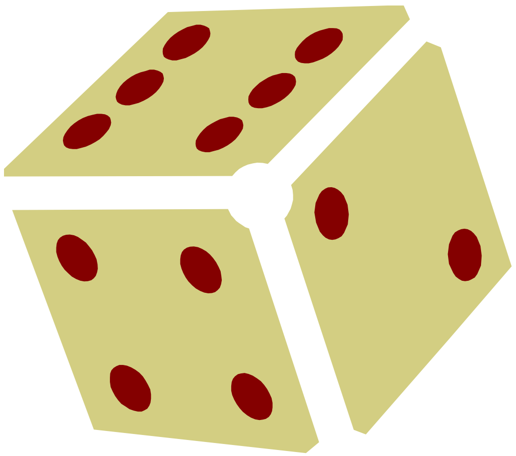
Note: The expected value of a random variable defined on an infinite sample space is defined iff the infinite series in the definition is absolutely convergent.

A roll of the dice…

Example: Let X be the number that comes up when a die is rolled. What is the expected value of X?

Solution:

- 6 possible outcomes: 1, 2, 3, 4, 5, 6
- Each outcomes occurs with the probability 1/6
- $E(X) = 1/6 + 2/6 + 3/6 + 4/6 + 5/6 + 6/6$
- $= 21/6$
- $= 7/2$



A flip of the coin…

Example: A fair coin is flipped three times. Let S be the sample space of the eight possible outcomes, and X be the random variable that assigns to an outcome the number of heads in that outcome. What is the expected value of X?

Solution:

- Since coin flips are independent, each outcome is equally likely
- E(X) = 1/8[X(HHH) + X(HHT) + X(HTH) + X(THH) + X(TTH) + X(THT) + X(HTT) + X(TTT)]
- = 1/8[3 + 2 + 2 + 2 + 1 + 1 + 1 + 0]
- = 12/8
- = 3/2

If S is large, the definition of expected value can be difficult to use directly

Definition: If X is a random variable and p(X=r) is the probability that X = r (i.e., p(X=r) = $\sum_{s\in S, X(s)=r} p(s)$), then

$$E(X) = \sum_{r \in X(S)} p(X = r)r$$

Each value of X…

… is weighted by its probability of occurrence.

Proof:

- Suppose that X is a random variable ranging over S
- Note that p(X=r) is the probability that X takes the value r
- This means that p(X=r) is the sum of the probabilities of the outcomes s∈S such that X(s) = r
- It thus follows that $E(X) = \sum_{r \in X(S)} p(X = r)r$ ❑

Rolling two dice

Example: Let X be the sum of the numbers that appear when a pair of fair dice is rolled. What is the expected value of X?

Recall from last week:

- $X(1,1) = 2$ — $p(X=2) = 1/36$
- $X(1,2) = X(2, 1) = 3$ — $p(X=3) = 2/36 = 1/18$
- $X(1,3) = X(2,2) = X(3,1) = 4$ — $p(X=4) = 3/36 = 1/12$
- $X(1,4) = X(2,3) = X(3,2) = X(4,1) = 5$ — $p(X=5) = 4/36 = 1/9$
- $X(1,5) = X(2,4) = X(3,3) = X(4,2) = X(5,1) = 6$ — $p(X=6) = 5/36$
- $X(1,6) = X(2,5) = X(3,4) = X(4,3) = X(5,2) = X(6,1) = 7$ — $p(X=7) = 6/36 = 1/6$
- $X(2,6) = X(3,5) = X(4,4) = X(5,3) = X(6,2) = 8$ — $p(X=8) = 5/36$
- $X(3,6) = X(4,5) = X(5,4) = X(6,3) = 9$ — $p(X=9) = 4/36 = 1/9$
- $X(4,6) = X(5,5) = X(6,4) = 10$ — $p(X=10) = 3/36 = 1/12$
- $X(5,6) = X(6,5) = 11$ — $p(X=11) = 2/36 = 1/18$
- $X(6,6) = 12$ — $p(X=12) = 1/36$

So we have that:

- $E(X) = 2(1/36) + 3(1/18) + 4(1/12) + 5(1/9) + 6(5/36) + 7(1/6) + 8(5/36) + 9(1/9) + 10(1/12) + 11(1/18) + 12(1/36)$
- $= 7$

We can apply this formula to reason about Bernoulli trials!

Theorem: The expected number of successes when n independent Bernoulli trials are performed, in which p is the probability of success, is np.

The proof of this theorem is straightforward (cf. Sec 7.4 of the text)

But let's think about it intuitively…

- 6 coin flips, how many will be heads?
- Bernoulli trials: n = 6, p = 0.5, q = 0.5
- Intuitively, you'd expect half of your flips to be heads
- Mathematically, 6 * 0.5 = 3

Expected values are linear!

Theorem: If $X_1, X_2, \ldots, X_n$ are random variables on S and if a and b are real numbers, then

1. $E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$
2. $E(aX + b) = aE(X) + b$

Proof:

- To prove the first result for n=2, note that
- $E(X_1 + X_2) = \Sigma_{s\in S}\, p(s)(X_1(s) + X_2(s))$ — Def'n of E(X)
- $= \Sigma_{s\in S}\, p(s)X_1(s) + \Sigma_{s\in S}\, p(s)X_2(s)$ — Property of summations
- $= E(X_1) + E(X_2)$ — Def'n of E(X)
- The case with n variables is an easy proof by induction

- To prove the second property, note that
- $E(aX + b) = \Sigma_{s\in S}\, p(s)(aX(s) + b)$ — Def'n of E(X)
- $= \Sigma_{s\in S}\, p(s)aX(s) + \Sigma_{s\in S}\, p(s)b$ — Property of summations
- $= a\Sigma_{s\in S}\, p(s)X(s) + b\Sigma_{s\in S}\, p(s)$ — Property of summations
- $= aE(X) + b$ ❑ — Def'n of E(X), $\Sigma_{s\in S}\, p(s) = 1$

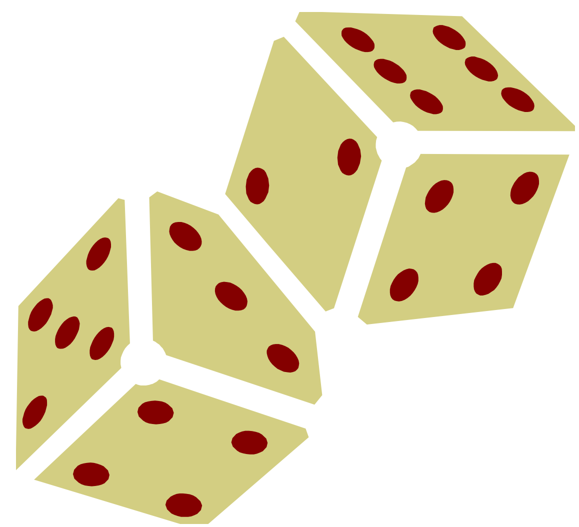
Dice, revisited

Example: What is the expected value of the sum of the numbers that appear when two fair dice are rolled?

Solution:

- Let X_1 and X_2 be random variables indicating the value on the first and second die, respectively
- Want to calculate $E(X_1+X_2)$
- By the previous theorem, we have that $E(X_1+X_2) = E(X_1)+E(X_2)$
- From earlier in lecture, we know that $E(X_1) = E(X_2) = 7/2$
- So, $E(X_1+X_2) = 7/2 + 7/2 = 7$

Note: This agrees with the (more complicated) calculation that we made earlier in lecture.



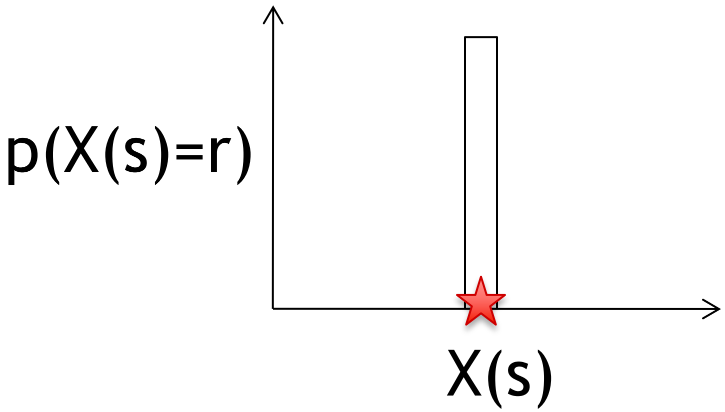
In-class exercises

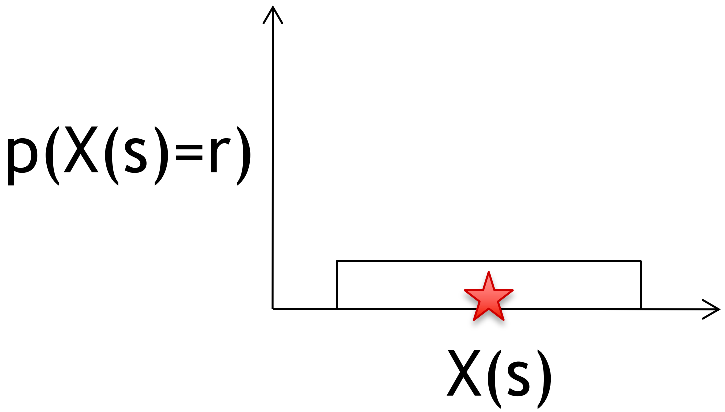
Problem 1: Consider a die in which the number 5 is two times as likely to be rolled as any other number. What is the expected value of this die?

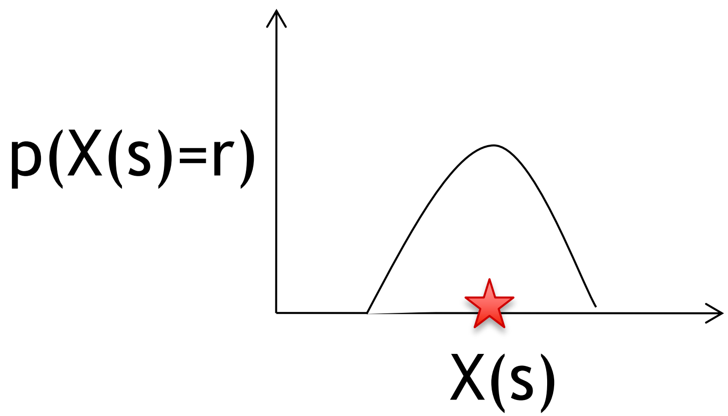
Problem 2: Alice and Bob regularly play chess together. Historically, Alice wins 70% of the time. If Alice and Bob play 7 games of chess, how many games can Alice be expected to win?

Sometimes we need more information than the expected value can give us

The expected value of a random variable doesn't tell us the whole story…







The variance of a random variable gives us information about how wide it is spread

Definition: The variance of a random variable X on a sample space S is defined as:

$$V(X) = \sum_{s \in S} \underbrace{\left(X(s) - E(X)\right)^2}_{\text{Squared difference from expected value}} p(s)$$

Squared difference from expected value

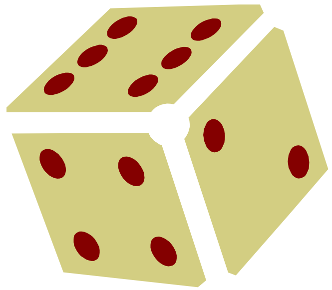
Weighted by probability of occurrence

Definition: The standard deviation of a random variable X on a sample space S is defined as $\sqrt{V(X)}$.

Variance of a die

Example: A fair die is rolled. What is the variance of the random variable X representing the face that appears?

Solution:



- Recall that $E(X) = 3.5$
- $X(1) = 1$, $p(1)=1/6$
- $X(2) = 2$, $p(2)=1/6$
- $X(3) = 3$, $p(3)=1/6$
- $X(4) = 4$, $p(4)=1/6$
- $X(5) = 5$, $p(5)=1/6$
- $X(6) = 6$, $p(6)=1/6$
- Thus, $V(X) = (1/6)(1-3.5)^2 + (1/6)(2-3.5)^2 + (1/6)(3-3.5)^2 + (1/6)(4-3.5)^2 + (1/6)(5-3.5)^2 + (1/6)(6-3.5)^2$
- $V(X) = 6.25/6 + 2.25/6 + 0.25/6 + 0.25/6 + 2.25/6 + 6.25/6$
- $V(X) = 17.5/6 \approx 2.92$

Variance: The short form

Theorem: If X is a random variable on a sample space S, then $V(X) = E(X^2) - E(X)^2$.

Proof:

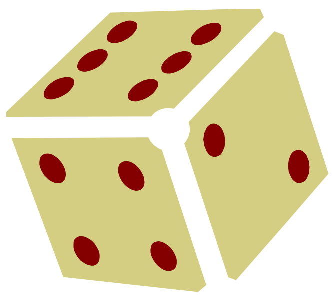
- $V(X) = \sum_{s \in S} (X(s) - E(X))^2 p(s)$
- $= \sum_{s \in S} X(s)^2 p(s) - 2E(X)\sum_{s \in S} X(s)p(s) + E(X)^2 \sum_{s \in S} p(s)$
- $= E(X^2) - 2E(X)E(X) + E(X)^2$
- $= E(X^2) - E(X)^2$ ❑

Variance of a die, revisited

Example: A fair die is rolled. What is the variance of the random variable X representing the face that appears?

Solution:

- Recall that $E(X) = 3.5$
- $X^2(1) = 1$, $p(1)=1/6$
- $X^2(2) = 4$, $p(2)=1/6$
- $X^2(3) = 9$, $p(3)=1/6$
- $X^2(4) = 16$, $p(4)=1/6$
- $X^2(5) = 25$, $p(5)=1/6$
- $X^2(6) = 36$, $p(6)=1/6$
- Thus, $E(X^2) = (1/6)(1) + (1/6)(4) + (1/6)(9) + (1/6)(16) + (1/6)(25) + (1/6)(36)$
- $E(X^2) = 1/6 + 4/6 + 9/6 + 16/6 + 25/6 + 36/6 = 91/6$
- $V(X) = E(X^2) - E(X)^2 = 91/6 - 3.5^2 \approx 2.92$

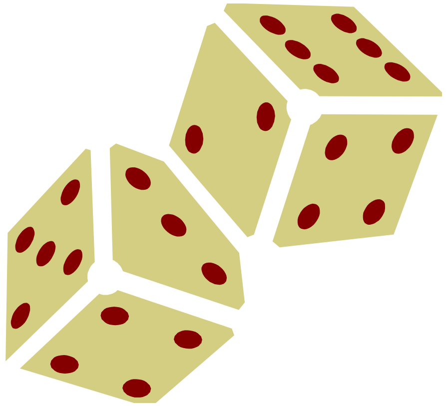


Multiple Dice

Example: Two dice are rolled. What is the variance of the random variable $X((j,k)) = 2j$, where j is the number of appearing on the first die and k is the number appearing on the second die.

Solution:

- $V(X) = E(X^2) - E(X)^2$
- Note that $p(X=k) = 1/6$ for $k = 2,4,6,8,10,12$ and is 0 otherwise
- $E(X) = (2+4+6+8+10+12)/6 = 7$
- $E(X^2) = (2^2+4^2+6^2+8^2+10^2+12^2)/6 = 182/3$
- So $V(X) = 182/3 - 49 = 35/3$



Variance of a Bernoulli Distribution

Example: What is the variance of random variable X with X(t)=1 if a Bernoulli trial is a success and X(t)=0 otherwise? Assume that the probability of success is p.

Solution:

§7.4 also proves that the variance of n Bernoulli trials is npq

- Note that X takes only the values 0 and 1
- Hence, $X(t) = X^2(t)$
- $V(X) = E(X^2) - E(X)^2$
- $\quad = p - p^2$
- $\quad = p(1-p)$
- $\quad = pq$

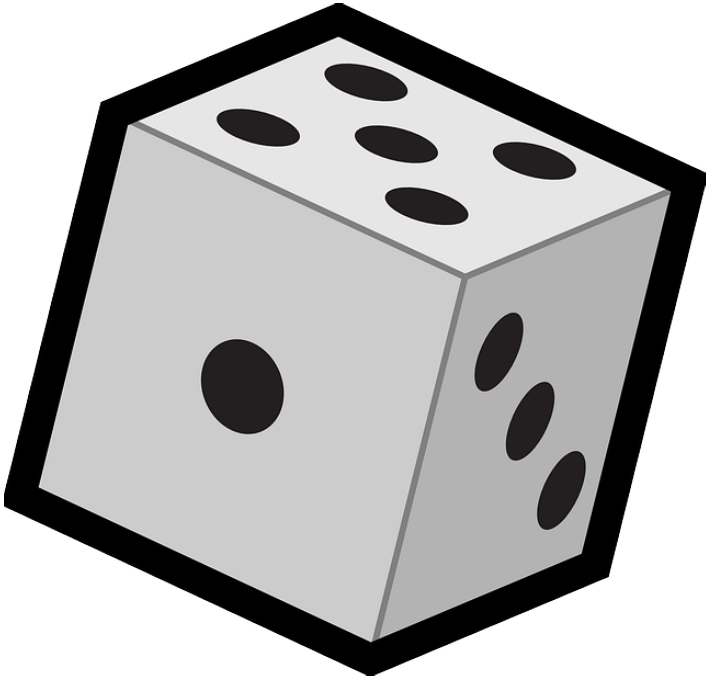
This tells us that the variance of ANY Bernoulli distribution is pq!

Variance of n Bernoulli trials

Example: A fair die is rolled 5 times. Let X be the random variable that assigns to an outcome the number of throws less than 3. What is the variance of X?

Solution:

- $n = 5$, $p = 1/3$, $q = 2/3$
- $V(X) = npq = 5 * 1/3 * 2/3 \approx 1.11$



In-class exercises

Problem 3: What is the variance of the number of heads that come up when a fair coin is flipped 10 times?

Problem 4: Let X be a random variable that equals the number of tails minus the number of heads when 3 fair coins are flipped. What is the expected value of X? What is the variance of X?

Final Thoughts

- Analyzing the expected value of a random variable allows us to answer a range of interesting questions
- The variance of a random variable tells us about the spread of values that the random variable can take