INFORMATION & COMPUTER SCIENCES
UNIVERSITY of HAWAIʻI at MĀNOA

Ch 7.4: Expected Value and Variance

ICS 141: Discrete Mathematics for Computer Science I

KYLE BERNEY
DEPARTMENT OF ICS, UNIVERSITY OF HAWAII AT MANOA

Random Variables

- *Recall:* A random variable is a mapping from the sample space of an experiment to the set of real numbers

- Let X be a random variable
- The probability that X takes the value $r \in \mathbb{R}$ is denoted as $\Pr(X = r)$

Expected Value

- Definition 1: The expected value of the random variable X on the sample space S, denoted $E[X]$, is equal to

$$E[X] = \sum_{s \in S} \Pr(s) X(s)$$

Expected Value

- Ex: Let X be a random variable defined as the number that comes up when a fair die is rolled. What is the expected value of X?

Expected Value

- <u>Ex:</u> Let X be a random variable defined as the number that comes up when a fair die is rolled. What is the expected value of X?
- <u>Solution:</u> The random variable X can take the values 1, 2, 3, 4, 5, or 6, each with probability $1/6$. It follows that

$$\begin{aligned} E[X] &= \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 \\ &= \frac{21}{6} \\ &= \frac{7}{2} \end{aligned}$$

Expected Value

- Ex: Let X be a random variable defined as the total number of heads when a fair coin is flipped three times. What is the expected value of X?

Expected Value

- Ex: Let X be a random variable defined as the total number of heads when a fair coin is flipped three times. What is the expected value of X?
- Solution: The random variable X can take the values 0, 1, 2, or 3. Since there are 8 total possible outcomes, each outcome has a probability of $1/8$. It follows that

$$\begin{aligned} E[X] &= \frac{1}{8}(X(HHH) + X(HHT) + X(HTH) + X(HTT) \\ &\quad + X(THH) + X(THT) + X(TTH) + X(TTT)) \\ &= \frac{1}{8}(3 + 2 + 2 + 1 + 2 + 1 + 1 + 0) \\ &= \frac{12}{8} = \frac{3}{2} \end{aligned}$$

Expected Value

- When an experiment has relatively few outcomes, we can compute the expected value of a random variable directly from its definition by enumerating all of its outcomes
- However, when there a large number of outcomes, we rely on the following result

- Theorem 1: Let X be a random variable where $\Pr(X = r)$ is the probability that X takes the value of r.

$$E[X] = \sum_{r \in X(S)} r \cdot \Pr(X = r)$$

Expected Value

- Ex: What is the expected value of the sum of the numbers that appear when a pair of fair dice are rolled?

Expected Value

- Solution: Let X be a random variable defined as the sum of the numbers that appear when a pair of dice are rolled.
  - $X = 2 \Rightarrow \{(1,1)\}$
  - $X = 3 \Rightarrow \{(1,2),(2,1)\}$
  - $X = 4 \Rightarrow \{(1,3),(2,2),(3,1)\}$
  - $X = 5 \Rightarrow \{(1,4),(2,3),(3,2),(4,1)\}$
  - $X = 6 \Rightarrow \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$
  - $X = 7 \Rightarrow \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$
  - $X = 8 \Rightarrow \{(2,6),(3,5),(4,4),(5,3),(6,2)\}$
  - $X = 9 \Rightarrow \{(3,6),(4,5),(5,4),(6,3)\}$
  - $X = 10 \Rightarrow \{(4,6),(5,5),(6,4)\}$
  - $X = 11 \Rightarrow \{(5,6),(6,5)\}$
  - $X = 12 \Rightarrow \{(6,6)\}$

Expected Value

- Solution: Hence,
  - $\Pr(X = 2) = 1/36$
  - $\Pr(X = 3) = 2/36$
  - $\Pr(X = 4) = 3/36$
  - $\Pr(X = 5) = 4/36$
  - $\Pr(X = 6) = 5/36$
  - $\Pr(X = 7) = 6/36$
  - $\Pr(X = 8) = 5/36$
  - $\Pr(X = 9) = 4/36$
  - $\Pr(X = 10) = 3/36$
  - $\Pr(X = 11) = 2/36$
  - $\Pr(X = 12) = 1/36$

Expected Value

- Solution: Therefore,

$$
\begin{aligned}
E[X] &= \sum_{r=2}^{12} r \cdot \Pr(X = r) \\
&= 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + 5 \cdot \frac{4}{36} + 6 \cdot \frac{5}{36} + 7 \cdot \frac{6}{36} \\
&\quad + 8 \cdot \frac{5}{36} + 9 \cdot \frac{4}{36} + 10 \cdot \frac{3}{36} + 11 \cdot \frac{2}{36} + 12 \cdot \frac{1}{36} \\
&= \frac{252}{36} \\
&= 7
\end{aligned}
$$

Expected Value

- Wait... we said that we could rely on Theorem 1 so that we do not have to enumerate all possible outcomes
- However in the previous example, in order to calculate $\Pr(X = r)$ for $r = 2, 3, \ldots, 12$, we had to enumerate all possible outcomes!

- Indicator random variables and linearity of expectations allow us to simplify calculations

Indicator Random Variables

- Definition 2: Given a sample space S and an event A, the indicator random variable I_A associated with the event A is defined as

$$I_A = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A \text{ does not occur} \end{cases}$$

Indicator Random Variables

- Definition 2: Given a sample space S and an event A, the indicator random variable I_A associated with the event A is defined as

$$I_A = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A \text{ does not occur} \end{cases}$$

- Ex: Suppose that we flip a fair coin. Let X_H be an indicator random variable associated with the coin coming up heads.

$$X_H = \begin{cases} 1 & \text{if } H \text{ occurs} \\ 0 & \text{if } T \text{ occurs} \end{cases}$$

Indicator Random Variables

- Definition 2: Given a sample space S and an event A, the indicator random variable I_A associated with the event A is defined as

$$I_A = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A \text{ does not occur} \end{cases}$$

- Ex: Suppose that we roll a fair die. Let X_i be an indicator random variable associated with the roll resulting in i, for $i = 1, 2, \ldots, 6$.

$$X_i = \begin{cases} 1 & \text{if } i \text{ is rolled} \\ 0 & \text{otherwise} \end{cases}$$

Expected Value

- Lemma 1: Let I_A be an indicator random variable associated with the event A.

$$E[I_A] = \Pr(A)$$

Expected Value

- Lemma 1: Let I_A be an indicator random variable associated with the event A.

$$E[I_A] = \Pr(A)$$

- Proof: By definittion of an indicator random variable and the definition of the expected value, we have

$$\begin{aligned} E[I_A] &= 1 \cdot \Pr(A) + 0 \cdot \Pr(\overline{A}) \\ &= \Pr(A) \end{aligned}$$

$\blacksquare$

Linearity of Expectations

- <u>Theorem 2:</u> (Linearity of Expectations) For $i = 1, 2, \ldots, n$, let X_i be a random variable and let a and b be real numbers.
  1. $E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$
  2. $E[aX + b] = aE[X] + b$

Linearity of Expectations

- Proof: (1.)

  Let n be an arbitrary positive integer.

  Inductive Hypothesis: Assume inductively that for all integers k, such that $0 < k < n$, $P(k)$ is true. In other words,

  $$E\left[\sum_{i=1}^{k} X_i\right] = \sum_{i=1}^{k} E[X_i]$$

  Base Case: Assume $n = 1$.

  Trivially, we have

  $$E\left[\sum_{i=1}^{1} X_i\right] = E[X_1] = \sum_{i=1}^{1} E[X_i]$$

Linearity of Expectations

- <u>Proof:</u> (1.)
  <u>Inductive Case:</u> Assume $n > 1$.
  By definition of the expected value, it follows that

$$
\begin{aligned}
E\left[\sum_{i=1}^{n} X_i\right] &= \sum_{s\in S}\left(\Pr(s)\sum_{i=1}^{n} X_i(s)\right) \\
&= \sum_{s\in S}\Pr(s)\left(\left(\sum_{i=1}^{n-1} X_i(s)\right) + X_n(s)\right) \\
&= \sum_{s\in S}\left(\Pr(s)\sum_{i=1}^{n-1} X_i(s)\right) + \sum_{s\in S}\Pr(s)X_n(s) \\
&= E\left[\sum_{i=1}^{n-1} X_i\right] + E[X_n]
\end{aligned}
$$

Linearity of Expectations

- Proof: (1.)
  Inductive Case: Assume $n > 1$.
  Since $0 < n-1 < n$, from our inductive hypothesis we know that

$$E\left[\sum_{i=1}^{n-1} X_i\right] = \sum_{i=1}^{n-1} E[X_i]$$

Therefore,

$$E\left[\sum_{i=1}^{n-1} X_i\right] + E[X_n] = \left(\sum_{i=1}^{n-1} E[X_i]\right) + E[X_n]$$

$$= \sum_{i=1}^{n} E[X_i]$$

$\blacksquare$

Linearity of Expectations

- Proof: (2.)
  By definition of the expected value, it follows that

$$\begin{aligned}
E[aX+b] &= \sum_{s\in S} \Pr(s)(aX(s)+b) \\
&= a\left(\sum_{s\in S}\Pr(s)X(s)\right) + b\sum_{s\in S}\Pr(s) \\
&= aE[X]+b, \quad \text{since} \sum_{s\in S}\Pr(s) = 1
\end{aligned}$$

$\blacksquare$

Expected Value

- <u>Ex:</u> What is the expected value of the sum of the numbers that appear when a pair of fair dice are rolled?

Expected Value

- Solution: For $i = 1, 2, \ldots, 6$, let X_i be an indicator random variable such that

$$X_i = \begin{cases} 1 & \text{if } i \text{ is rolled} \\ 0 & \text{otherwise} \end{cases}$$

Let Y_1 be a random variable such that

$$Y_1 = \sum_{i=1}^{6} i \cdot X_i$$

In other words, Y_1 represents the outcome of the first die.

Expected Value

- Solution:
  Similarly, let Y_2 be a random variable that represents the outcome of the second die.

$$Y_2 = \sum_{i=1}^{6} i \cdot X_i$$

  Lastly, let $Z = Y_1 + Y_2$ be a random variable defined as the sum of the numbers that appears when a pair of fair dice are rolled.

Expected Value

- <u>Solution:</u>
  Thus, using linearity of expectations

$$
\begin{aligned}
E[Z] &= E[Y_1 + Y_2] \\
&= E[Y_1] + E[Y_2] \\
&= E\left[\sum_{i=1}^{6} i \cdot X_i\right] + E\left[\sum_{i=1}^{6} i \cdot X_i\right] \\
&= \left(\sum_{i=1}^{6} i \cdot E[X_i]\right) + \left(\sum_{i=1}^{6} i \cdot E[X_i]\right) \\
&= 2\sum_{i=1}^{6} i \cdot E[X_i]
\end{aligned}
$$

Expected Value

- Solution:
  Since all X_i's are indicator random variables,
  $E[X_i] = \Pr(X_i = 1) = 1/6$

$$2\sum_{i=1}^{6} i \cdot E[X_i] = 2\sum_{i=1}^{6} i \cdot \frac{1}{6}$$

$$= \frac{1}{3}\sum_{i=1}^{6} i$$

$$= \frac{21}{3}$$

$$= 7$$

Expected Value

- Ex: Let $A[1 \ldots n]$ be an array of n integers. If $i < j$ and $A[i] > A[j]$, then the pair (i, j) is called an inversion of A (i.e., they are out of order with respect to each other). Suppose that the elements of A form a uniform random permutation of the elements $\{1, 2, \ldots, n\}$. What is the expected number of inversions?

Expected Value

- Solution: For $i < j$, let $X_{i,j}$ be an indicator random variable associated with the event that $A[i] > A[j]$ (i.e., they are inverted). Since the n elements of A are a uniform random permutation, we have that

$$\Pr(X_{i,j} = 1) = \frac{1}{2}$$

In other words, it is equally likely that $A[i] > A[j]$ or $A[i] < A[j]$ (there are an equal number of permutations such that $A[i]$ preceeds $A[j]$ and $A[j]$ preceeds $A[i]$).
Hence,

$$E[X_{i,j}] = \frac{1}{2}$$

Expected Value

- <u>Solution:</u>
  Let X be a random variable denoting the total number of inverted pairs in the array A. In other words, X is the sum of all $X_{i,j}$ that meet the constraint that $1 \le i < j \le n$.

$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{i,j}$$

Expected Value

- Solution:
  Using linearity of expectations, we have

$$E[X] = E\left[\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} X_{i,j}\right]$$

$$= \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} E[X_{i,j}]$$

$$= \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \frac{1}{2}$$

Expected Value

- <u>Solution:</u>
  Using linearity of expectations, we have

$$
\begin{aligned}
\frac{1}{2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} 1 &= \frac{1}{2}\sum_{i=1}^{n-1}(n-i) \\
&= \frac{1}{2}\left(\sum_{i=1}^{n-1} n - \sum_{i=1}^{n-1} i\right) \\
&= \frac{1}{2}\left(n(n-1) - \frac{n(n-1)}{2}\right) \\
&= \frac{n(n-1)}{4}
\end{aligned}
$$

Independent Random Variables

- <u>Definition 3:</u> The random variables X and Y on a sample space S are <u>independent</u> if

$$\Pr(X = r_1 \text{ and } Y = r_2) = \Pr(X = r_1)\Pr(Y = r_2)$$

Independent Random Variables

- *Recall:* From Slide 12, we considered the problem of rolling two fair dice and asked what is the expected value.
  - Let Y_1 be a random variable that represents the outcome of the first die
  - Let Y_2 be a random variable that represents the outcome of the second die
- *Question:* Are Y_1 and Y_2 independent?

Independent Random Variables

- <u>Solution:</u> Let $S = \{1, 2, 3, 4, 5, 6\}$ and let $i, j \in S$. Since each of the 36 possible outcomes are equally likely, we have that

$$\Pr(Y_1 = i \text{ and } Y_2 = j) = \frac{1}{36}$$

Furthermore, each of the 6 outcomes on each die are equally likely

$$\Pr(Y_1 = i) = \Pr(Y_2 = j) = \frac{1}{6}$$

Hence,

$$\Pr(Y_1 = i \text{ and } Y_2 = j) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = \Pr(Y_1 = i)\Pr(Y_2 = j)$$

Therefore, Y_1 and Y_2 are independent random variables.

Independent Random Variables

- Theorem 3: If X and Y are independent random variables on a sample space S, then

$$E[XY] = E[X]E[Y]$$

Independent Random Variables

- Theorem 3: If X and Y are independent random variables on a sample space S, then

$$E[XY] = E[X]E[Y]$$

- Proof: See textbook for a proof.

Variance

- Expected value of a random variable tells us its "average" value
- Variance of a random variable measures how widely a random variable is distributed about its expected value

- Definition 4: Let X be a random variable on a sample space S. The variance of X, denoted $V[X]$, is

$$V[X] = \sum_{s \in S} (X(s) - E[X])^2 \Pr(s)$$

- In other words, $V[X]$ is the weighted average of the square of the deviation of X
- The standard deviation of X, denoted $\sigma(X)$, is defined as

$$\sqrt{V[X]}$$

Variance

- Theorem 4: If X is a random variable on a sample space S, then

$$V[X] = E[X^2] - E[X]^2$$

Variance

- <u>Proof:</u>

$$
\begin{aligned}
V[X] &= \sum_{s \in S} (X(s) - E[X])^2 \Pr(s) \\
&= \sum_{s \in S} (X^2(s) - 2X(s)E[X] + E[X]^2) \Pr(s) \\
&= \left(\sum_{s \in S} X^2(s) \Pr(s) \right) - \left(2E[X] \sum_{s \in S} X(s) \Pr(s) \right) \\
&\quad + \left(E[X]^2 \sum_{s \in S} \Pr(s) \right) \\
&= E[X^2] - 2E[X]E[X] + E[X]^2 \\
&= E[X^2] - E[X]^2
\end{aligned}
$$

■

Variance

- Corollary 1: If X is a random variable on a sample space S and $E[X] = \mu$, then

$$V[X] = E[(X - \mu)^2]$$

Variance

- Corollary 1: If X is a random variable on a sample space S and $E[X] = \mu$, then

$$V[X] = E[(X - \mu)^2]$$

- Proof:

$$\begin{aligned} E[(X - \mu)^2] &= E[X^2 - 2X\mu + \mu^2] \\ &= E[X^2] - 2\mu E[X] + \mu^2 \\ &= E[X^2] - 2\mu^2 + \mu^2 \\ &= E[X^2] - \mu^2 \\ &= E[X^2] - E[X]^2 \\ &= V[X] \end{aligned}$$

Variance

- <u>Ex:</u> Let X be a random variable that defined as the number that comes up when a fair die is rolled. What is the variance of X?

Variance

- <u>Ex:</u> Let X be a random variable that defined as the number that comes up when a fair die is rolled. What is the variance of X?

- <u>Solution:</u> From a previous example, we know that $E[X] = 7/2$. To find $E[X^2]$, note that X^2 takes the values i^2 for $i = 1, 2, \ldots 6$, each with probability $1/6$. Hence,

$$E[X^2] = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) = \frac{91}{6}$$

Therefore,

$$\begin{aligned} V[X] &= E[X^2] - E[X] \\ &= \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12} \end{aligned}$$

Bienyamé's Formula

- <u>Theorem 5:</u> (Bienyamé's Formula) If X and Y are independent random variables on a sample space S, then

$$V[X+Y] = V[X] + V[Y]$$

Furthermore, for $i = 1, 2, \ldots, n$, if X_i's are pairwise independent random variables on S, then

$$V\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} V[X_i]$$

Bienyamé's Formula

- Proof: It follows from Theorem 4 that

$$\begin{aligned}
V[X+Y] &= E[(X+Y)^2] - E[X+Y]^2 \\
&= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 \\
&= E[X^2] + 2E[XY] + E[Y^2] \\
&\quad - (E[X]^2 + 2E[X]E[Y] + E[Y]^2) \\
&= E[X^2] + 2E[XY] + E[Y^2] \\
&\quad - E[X]^2 - 2E[X]E[Y] - E[Y]^2
\end{aligned}$$

Bienyamé's Formula

- Proof:
  Since X and Y are independent, from Theorem 3 we have that $E[XY] = E[X]E[Y]$. Therefore,

$$E[X^2] + 2E[XY] + E[Y^2] - E[X]^2 - 2E[X]E[Y] - E[Y]^2$$
$$= E[X^2] + 2E[X]E[Y] + E[Y^2] - E[X]^2 - 2E[X]E[Y] - E[Y]^2$$
$$= E[X^2] - E[X]^2 + E[Y^2] - E[Y]^2$$
$$= V[X] + V[Y]$$

  The proof for the n pairwise independent random variables can be performed by generalizing this proof and using mathematical induction. (Try it yourself! See Exercise 33 in the textbook.)

Chebyshev's Inequality

- How likely is it that a random variable takes a value far from its expected value?
- Theorem 6: (Chebyshev's Inequality) Let X be a random variable on a sample space S. If r is a positive real number, then

$$\Pr(|X(s) - E[X]| \geq r) \leq \frac{V[X]}{r^2}$$