We will cover these parts of the book (8th edition):

6.3-6.4
7.1
7.2.1-7.2.7
7.4.1-7.4.4
7.4.6

Permutations and Combinations

- How many different sets of 3 people can we pick from a group of 6?
- There are 6 choices for the first person, 5 for the second one, and 4 for the third one, so there are 6·5·4 = 120 ways to do this.
- This is not the correct result!
- For example, picking person C, then person A, and then person E leads to the **same group** as first picking E, then C, and then A.
- However, these cases are counted **separately** in the above equation.

Permutations and Combinations

- So how can we compute how many different subsets of people can be picked (that is, we want to disregard the order of picking) ?

- To find out about this, we need to look at **permutations.**

- A **permutation** of a set of distinct objects is an ordered arrangement of these objects.

- An ordered arrangement of r elements of a set is called an **r-permutation**.

Permutations and Combinations

- **Example:** Let S = {1, 2, 3}.
- The arrangement 3, 1, 2 is a permutation of S.
- The arrangement 3, 2 is a 2-permutation of S.

- The number of r-permutations of a set with n distinct elements is denoted by **P(n, r).**

- We can calculate P(n, r) with the product rule:
- $P(n, r) = n \cdot (n - 1) \cdot (n - 2) \cdot \ldots \cdot (n - r + 1)$.

- (n choices for the first element, (n – 1) for the second one, (n – 2) for the third one…)

Permutations and Combinations

- **Example:**
- P(8, 3) = 8·7·6 = 336
- = (8·7·6·5·4·3·2·1)/(5·4·3·2·1)
- **General formula:**
- $P(n,r) = \frac{n!}{(n-r)!}$
- Knowing this, we can return to our initial question:
- How many different sets of 3 people can we pick from a group of 6?

UMASS BOSTON

Permutations and Combinations

- An **r-combination** of elements of a set is an unordered selection of r elements from the set.
- Thus, an r-combination is simply a subset of the set with r elements.

- **Example:** Let S = {1, 2, 3, 4}.
- Then {1, 3, 4} is a 3-combination from S.

- The number of r-combinations of a set with n distinct elements is denoted by C(n, r).

- **Example:** C(4, 2) = 6, since, for example, the 2-combinations of a set {1, 2, 3, 4} are {1, 2}, {1, 3}, {1, 4}, {2, 3}, {2, 4}, {3, 4}.

Permutations and Combinations

- How can we calculate C(n, r)?
- Consider that we can obtain the r-permutations of a set in the following way:
- **First,** we form all the r-combinations of the set (there are C(n, r) such r-combinations).
- **Then,** we generate all possible orderings within each of these r-combinations (there are P(r, r) such orderings in each case).
- Therefore, we have:
- $P(n, r) = C(n, r) \cdot P(r, r)$

Permutations and Combinations

- $C(n,r) = \frac{P(n,r)}{P(r,r)} = \frac{\frac{n!}{(n-r)!}}{\frac{r!}{(r-r)!}} = \frac{n!}{r!(n-r)!}$

- Now we can answer our initial question:

- How many ways are there to pick a set of 3 people from a group of 6 (disregarding the order of picking)?

- C(6, 3) = 6!/(3!·3!) = 720/(6·6) = 720/36 = 20

- There are 20 different ways, that is, 20 different groups to be picked.

Permutations and Combinations

- **Corollary:**
- Let n and r be nonnegative integers with $r \leq n$.
- Then $C(n, r) = C(n, n - r)$.

- Note that “picking a group of r people from a group of n people” **is the same as** “splitting a group of n people into a group of r people and another group of (n – r) people”.

Combinations

- **Proof:**

$$C(n,n-r)=\frac{n!}{(n-r)![n-(n-r)]!}=\frac{n!}{(n-r)!r!}=C(n,r)$$

This symmetry is intuitively plausible. For example, let us consider a set containing six elements (n = 6).

Picking two elements and leaving four is essentially the same as picking four elements and leaving two.

In either case, our number of choices is the number of possibilities to divide the set into one set containing two elements and another set containing four elements.

Permutations and Combinations

- **Example:**
- A soccer club has 8 female and 7 male members. For today's match, the coach wants to have 6 female and 5 male players on the grass. How many possible configurations are there?

- $C(8, 6) \cdot C(7, 5) = 8!/(6!\cdot 2!) \cdot 7!/(5!\cdot 2!)$
- $= 28\cdot 21$
- $= 588$

Binomial Coefficients

- Expressions of the form $C(n,k) = \binom{n}{k}$ are also called **binomial coefficients**.
- How come?
- A **binomial expression** is the sum of two terms, such as (a + b).
- Now consider $(a + b)^2 = (a + b)(a + b)$.
- When expanding such expressions, we have to form all possible products of a term in the first factor and a term in the second factor:
- $(a + b)^2 = a \cdot a + a \cdot b + b \cdot a + b \cdot b$
- Then we can sum identical terms:
- $(a + b)^2 = a^2 + 2ab + b^2$

Binomial Coefficients

- For $(a + b)^3 = (a + b)(a + b)(a + b)$ we have
- $(a + b)^3 = aaa + aab + aba + abb + baa + bab + bba + bbb$
- $(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$
- There is only one term a^3, because there is only one possibility to form it: Choose **a** from all three factors: C(3, 3) = 1.
- There is three times the term a^2b, because there are three possibilities to choose **a** from a **subset** of two out of the three factors: C(3, 2) = 3.
- Similarly, there is three times the term ab^2 (C(3, 1) = 3) and once the term b^3 (C(3, 0) = 1).

Binomial Coefficients

- This leads us to the following formula:
- $(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^{n-j} y^j$ (Binomial Theorem)

$$= \binom{n}{0} x^n + \binom{n}{1} x^{n-1} y + \cdots + \binom{n}{n-1} x y^{n-1} + \binom{n}{n} y^n$$

- **Proof:**
- The terms in the product when it is expanded are of the form $x^{n-j}y^j$ for $j = 0, 1, 2, \ldots, n$. To count the number of terms of the form $x^{n-j}y^j$, note that to obtain such a term it is necessary to choose $n-j$ *xs* from the n binomial factors (so that the other j terms in the product are *ys*). Therefore, the coefficient of $x^{n-j}y^j$ is $\binom{n}{n-j}$, which is equal to $\binom{n}{j}$. This proves the theorem.

Binomial Coefficients

- **Example1:**
- What is the coefficient of $x^{12}y^{13}$ in the expansion of $(x+y)^{25}$?

- **Solution:**
- From the binomial theorem it follows that this coefficient is
- $\binom{25}{13} = \frac{25!}{13! * 12!} = 5200300$

Binomial Coefficients

- **Example2:** Prove $\sum_{k=0}^{n} \binom{n}{k} = 2^n$
- Use Binomial theorem. x=1, y=1

- **Example3:** Prove $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$
- x=-1, y=1

- **Example4:** Prove $\sum_{k=0}^{n} 2^k \binom{n}{k} = 3^n$
- x=1, y=2

Combinations

- **Pascal's Identity:**

- Let n and k be positive integers with n ≥ k.
Then C(n + 1, k) = C(n, k – 1) + C(n, k).

- $\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$

- How can this be explained?

- What is it good for?

Combinations

- Imagine a set S containing n elements and a set T containing (n + 1) elements, namely all elements in S plus a new element **a**.
- Calculating C(n + 1, k) is equivalent to answering the question: How many subsets of T containing k items are there?
- Case I: The subset contains (k – 1) elements of S plus the element **a**: C(n, k – 1) choices.
- Case II: The subset contains k elements of S and does not contain **a**: C(n, k) choices.
- **Sum Rule:** C(n + 1, k) = C(n, k – 1) + C(n, k).

Pascal's Triangle

- In Pascal's triangle, each number is the sum of the numbers to its upper left and upper right:

With the help of Pascal's triangle, you can find powers of binomial expressions.

For example, the fifth row of Pascal's triangle (1 – 4 – 6 – 4 – 1) helps us to compute $(a + b)^4$:

$(a + b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$

					1					
				1		1				
			1		2		1			
		1		3		3		1		
	1		4		6		4		1	
…		…		…		…		…		…

Pascal's Triangle

- Since we have C(n + 1, k) = C(n, k – 1) + C(n, k) and
C(0, 0) = 1, we can use Pascal's triangle to simplify the computation of C(n, k):

k

n

C(0, 0) = 1

C(1, 0) = 1 C(1, 1) = 1

C(2, 0) = 1 C(2, 1) = 2 C(2, 2) = 1

C(3, 0) = 1 C(3, 1) = 3 C(3, 2) = 3 C(3, 3) = 1

C(4, 0) = 1 C(4, 1) = 4 C(4, 2) = 6 C(4, 3) = 4 C(4, 4) = 1

Some other identities

- Vandermonde's: $\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{r-k}\binom{n}{k}$

- $\binom{2n}{n} = \sum_{k=0}^{n} \binom{n}{k}^2$

- $\binom{n+1}{r+1} = \sum_{j=r}^{n} \binom{j}{r}$

Now it’s time to look at…

Discrete Probability

Discrete Probability

- Everything you have learned about counting constitutes the basis for computing the **probability** of events to happen.
- In the following, we will use the notion **experiment** for a procedure that yields one of a given set of possible outcomes.
- This set of possible outcomes is called the **sample space** of the experiment.
- An **event** is a subset of the sample space.

Discrete Probability

- If all outcomes in the sample space are equally likely, the following definition of probability applies:
- The probability of an event E, which is a subset of a finite sample space S of equally likely outcomes, is given by $p(E) = \frac{|E|}{|S|}$.
- This is the **Laplace's** definition.
- Probability values range from **0** (for an event that will **never** happen) to **1** (for an event that will **always** happen whenever the experiment is carried out).

Discrete Probability

- **Example I:**
- An urn contains four blue balls and five red balls. What is the probability that a ball chosen from the urn is blue?
- **Solution:**
- There are nine possible outcomes, and the event “blue ball is chosen” comprises four of these outcomes. Therefore, the probability of this event is 4/9 or approximately 44.44%.

Discrete Probability

- **Example II:**
- What is the probability of winning the lottery 6/49, that is, picking the correct set of six numbers out of 49?
- **Solution:**
- There are C(49, 6) possible outcomes. Only one of these outcomes will actually make us win the lottery.
- p(E) = 1/C(49, 6) = 1/13,983,816

Discrete Probability

- **Example III:**
- Find the probability that a hand of five cards in poker contains four cards of one kind.
- **Solution:**
- By the product rule, the number of hands of five cards with four cards of one kind is the product of the number of ways to pick one kind, the number of ways to pick the four of this kind out of the four in the deck of this kind, and the number of ways to pick the fifth card. This is $C(13,1)C(4,4)C(48,1)$. And we know there are $C(52,5)$ different hands of 5 cards. So the probability is
- $\frac{C(13,1)C(4,4)C(48,1)}{C(52,5)} \approx 0.0003$

UMASS BOSTON

Complementary Events

- Let E be an event in a sample space S. The probability of an event $\bar{E} = S - E$, the **complementary event** of E, is given by
- p($\bar{E}$) = 1 – p(E).
- This can easily be shown:
- p($\bar{E}$) = (|S| - |E|)/|S| = 1 - |E|/|S| = 1 – p(E).
- This rule is useful if it is easier to determine the probability of the complementary event than the probability of the event itself.

Complementary Events

- **Example I:** A sequence of 10 bits is randomly generated. What is the probability that at least one of these bits is zero?
- **Solution:** There are 2^{10} = 1024 possible outcomes of generating such a sequence. The event $\bar{E}$, **"none of the bits is zero"**, includes only one of these outcomes, namely the sequence 1111111111.
- Therefore, p($\bar{E}$) = 1/1024.
- Now p(E) can easily be computed as p(E) = 1 – p($\bar{E}$) = 1 – 1/1024 = 1023/1024.

Complementary Events

- **Example II:** What is the probability that at least two out of 36 people have the same birthday?
- **Solution:** The sample space S encompasses all possibilities for the birthdays of the 36 people, so $|S| = 365^{36}$.
- Let us consider the event $\bar{E}$ ("no two people out of 36 have the same birthday"). $\bar{E}$ includes P(365, 36) outcomes (365 possibilities for the first person's birthday, 364 for the second, and so on).
- Then $p(\bar{E}) = P(365, 36)/365^{36} = 0.168$, so p(E) = 0.832 or 83.2%

Discrete Probability

- Let E_1 and E_2 be events in the sample space S. Then we have:

- $p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$

Does this remind you of something?

Of course, the principle of **inclusion-exclusion**.

Discrete Probability

- **Example:** What is the probability of a positive integer selected at random from the set of positive integers not exceeding 100 to be divisible by 2 or 5?
- **Solution:**
- E_2: "integer is divisible by 2"
  E_5: "integer is divisible by 5"
- $E_2 = \{2, 4, 6, \ldots, 100\}$
- $|E_2| = 50$
- $p(E_2) = 0.5$

Discrete Probability

- $E_5 = \{5, 10, 15, \ldots, 100\}$
- $|E_5| = 20$
- $p(E_5) = 0.2$

- $E_2 \cap E_5 = \{10, 20, 30, \ldots, 100\}$
- $|E_2 \cap E_5| = 10$
- $p(E_2 \cap E_5) = 0.1$

- $p(E_2 \cup E_5) = p(E_2) + p(E_5) - p(E_2 \cap E_5)$
- $p(E_2 \cup E_5) = 0.5 + 0.2 - 0.1 = 0.6$

Discrete Probability

- What happens if the outcomes of an experiment are **not** equally likely?
- In that case, we assign a probability p(s) to each outcome $s \in S$, where S is the sample space.
- Two conditions have to be met:
- (1): $0 \leq p(s) \leq 1$ for each $s \in S$, and
- (2): $\sum_{s \in S} p(s) = 1$
- This means, as we already know, that (1) each probability must be a value between 0 and 1, and (2) the probabilities must add up to 1, because one of the outcomes is **guaranteed** to occur.

Discrete Probability

- The function p from the set of all outcomes of the sample space S is called a **probability distribution**.

- How can we obtain these probabilities p(s) ?

- The probability p(**s**) assigned to an outcome **s** equals the limit of the number of times **s** occurs divided by the number of times the experiment is performed.

Discrete Probability

- Once we know the probabilities p(s), we can compute the **probability of an event E** as follows:

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

- Suppose that S is a set with n elements. The **uniform distribution** assigns the probability $\frac{1}{n}$ to each element of S.

Discrete Probability

- **Example I:** A die is biased so that the number 3 appears twice as often as each other number.
- What are the probabilities of all possible outcomes?

- **Solution:** There are 6 possible outcomes $s_1, \ldots, s_6$.
- $p(s_1) = p(s_2) = p(s_4) = p(s_5) = p(s_6)$
- $p(s_3) = 2p(s_1)$
- Since the probabilities must add up to 1, we have:
- $5p(s_1) + 2p(s_1) = 1$
- $7p(s_1) = 1$
- $p(s_1) = p(s_2) = p(s_4) = p(s_5) = p(s_6) = 1/7$,
- $p(s_3) = 2/7$

Discrete Probability

- **Example II:** For the biased die from Example I, what is the probability that an odd number appears when we roll the die?
- **Solution:**
- $E_{odd} = \{s_1, s_3, s_5\}$
- Remember the formula $p(E) = \sum_{s \in E} p(s)$.
- $p(E_{odd}) = \sum_{s \in E_{odd}} p(s) = p(s_1) + p(s_3) + p(s_5)$
- $p(E_{odd}) = 1/7 + 2/7 + 1/7 = 4/7 = 57.14\%$

Conditional Probability

- If we toss a coin three times, what is the probability that an odd number of tails appears **(event E)**, if the first toss is a tail **(event F)** ?
- If the first toss is a tail, the possible sequences are TTT, TTH, THT, and THH.
- In two out of these four cases, there is an odd number of tails.
- Therefore, the probability of E, under the condition that F occurs, is 0.5.
- We call this **conditional probability**.

Conditional Probability

- If we want to compute the conditional probability of E given F, we use F as the sample space.
- For any outcome of E to occur under the condition that F also occurs, this outcome must also be in $E \cap F$.

- **Definition:** Let E and F be events with $p(F) > 0$. The conditional probability of E given F, denoted by $p(E \mid F)$, is defined as

- $p(E \mid F) = p(E \cap F)/p(F)$

Conditional Probability

- **Example:** What is the probability of a random bit string of length four to contain at least two consecutive 0s, given that its first bit is a 0 ?

- **Solution:**
- E: "bit string contains at least two consecutive 0s"
- F: "first bit of the string is a 0"

- We know the formula $p(E \mid F) = p(E \cap F)/p(F)$.

- $E \cap F = \{0000, 0001, 0010, 0011, 0100\}$
- $p(E \cap F) = 5/16$
- $p(F) = 8/16 = 1/2$
- $p(E \mid F) = (5/16)/(1/2) = 10/16 = 5/8 = 0.625$

Independence

- Let us return to the example of tossing a coin three times.
- Does the probability of event E (odd number of tails) **depend** on the occurrence of event F (first toss is a tail) ?
- In other words, is it the case that $p(E \mid F) \neq p(E)$?
- We actually find that $p(E \mid F) = 0.5$ and $p(E) = 0.5$, so we say that E and F are **independent events**.

Independence

- Because we have $p(E \mid F) = p(E \cap F)/p(F)$, $p(E \mid F) = p(E)$ if and only if $p(E \cap F) = p(E)p(F)$.

- **Definition:** The events E and F are said to be independent if and only if $p(E \cap F) = p(E)p(F)$.

- Obviously, this definition is **symmetrical** for E and F. If we have $p(E \cap F) = p(E)p(F)$, then it is also true that $p(F \mid E) = p(F)$.

Independence

- **Example:** Suppose E is the event of rolling an even number with an unbiased die. F is the event that the resulting number is divisible by three. Are events E and F independent?
- **Solution:**
- $p(E) = 1/2$, $p(F) = 1/3$.
- $|E \cap F| = 1$ (only 6 is divisible by both 2 and 3)
- $p(E \cap F) = 1/6$
- $p(E \cap F) = p(E)p(F)$
- Conclusion: E and F are **independent**.

UMASS BOSTON

Bernoulli Trials

- Suppose an experiment with **two possible outcomes**, such as tossing a coin.
- Each performance of such an experiment is called a **Bernoulli trial**.
- We will call the two possible outcomes a **success** or a **failure**, respectively.
- If p is the probability of a success and q is the probability of a failure, it is obvious that $p + q = 1$.

Bernoulli Trials

- Often we are interested in the probability of **exactly k successes** when an experiment consists of **n independent Bernoulli trials**.

- **Example:**

A coin is biased so that the probability of head is 2/3. What is the probability of exactly four heads to come up when the coin is tossed seven times?

Bernoulli Trials

- **Solution:**
- There are $2^7 = 128$ possible outcomes.
- The number of possibilities for four heads among the seven trials is C(7, 4).
- The seven trials are independent, so the probability of each of these outcomes is $(2/3)^4(1/3)^3$.
- Consequently, the probability of exactly four heads to appear is
- $C(7, 4)(2/3)^4(1/3)^3 = 560/2187 = 25.61\%$

UMASS BOSTON

Bernoulli Trials

- **Illustration:** Let us denote a success by ‘S’ and a failure by ‘F’. As before, we have a probability of success p and probability of failure q = 1 – p.
- What is the probability of **two** successes in **five** independent Bernoulli trials?
- Let us look at a possible sequence:
- SSFFF
- What is the probability that we will generate exactly this sequence?

Bernoulli Trials

Sequence: S S F F F

Probability: p p q q q = p^2q^3

Another possible sequence:

Sequence: F S F S F

Probability: q p q p q = p^2q^3

- Each sequence with two successes in five trials occurs with probability p^2q^3.

Bernoulli Trials

- And how many possible sequences are there?
- In other words, how many ways are there to pick two items from a list of five?
- We know that there are C(5, 2) = 10 ways to do this, so there are 10 possible sequences, each of which occurs with a probability of p^2q^3.
- Therefore, the probability of **any** such sequence to occur when performing five Bernoulli trials is C(5, 2) p^2q^3.
- In general, for k successes in n Bernoulli trials we have a probability of $\mathbf{C(n,k)p^kq^{n-k}}$.

Random Variables

- In some experiments, we would like to assign a numerical value to each possible outcome in order to facilitate a mathematical analysis of the experiment.
- For this purpose, we introduce **random variables**.
- **Definition:** A random variable is a function from the sample space of an experiment to the set of real numbers. That is, a random variable assigns a real number to each possible outcome.
- Note: Random variables are functions, not variables, and they are not random, but map random results from experiments onto real numbers in a well-defined manner.

Random Variables

- **Example:**

- Let X be the result of a rock-paper-scissors game.
- If player A chooses symbol a and player B chooses symbol b, then

- X(a, b) = 1, if player A wins,
- = 0, if A and B choose the same symbol,
- = -1, if player B wins.

Random Variables

X(rock, rock) =	0
X(rock, paper) =	-1
X(rock, scissors) =	1
X(paper, rock) =	1
X(paper, paper) =	0
X(paper, scissors) =	-1
X(scissors, rock) =	-1
X(scissors, paper) =	1
X(scissors, scissors) =	0

Random Variables

- 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. (The set of pairs in this distribution is determined by the probabilities $p(X = r)$ for $r \in X(S)$.)

- **Example:** Suppose that a coin is flipped three times. Let $X(t)$ be the random variable that equals the number of heads that appear when t is the outcome. So, the distribution of the random variable $X(t)$ is determined by the probabilities $P(X$

Expected Values

- Once we have defined a random variable for our experiment, we can statistically analyze the outcomes of the experiment.

- For example, we can ask: What is the **average value** (called the **expected value**) of a random variable when the experiment is carried out a large number of times?

- Can we just calculate the arithmetic mean across all possible values of the random variable?

Expected Values

- No, we cannot, since it is possible that some outcomes are more likely than others.

- For example, assume the possible outcomes of an experiment are 1 and 2 with probabilities of 0.1 and 0.9, respectively.

- Is the average value 1.5?

- No, since 2 is much more likely to occur than 1, the average must be larger than 1.5.

Expected Values

- Instead, we have to calculate the **weighted sum** of all possible outcomes, that is, each value of the random variable has to be multiplied with its probability before being added to the sum.

- In our example, the average value is given by $0.1 \cdot 1 + 0.9 \cdot 2 = 0.1 + 1.8 = 1.9$.

Expected Values

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

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

- The **deviation** of X at $s \in$ S is $X(s) - E(X)$, the difference between the value of X and the mean of X.

UMASS BOSTON

Expected Values

- **Example:** Let X be the random variable equal to the sum of the numbers that appear when a pair of dice is rolled.
- There are **36 outcomes** (= pairs of numbers from 1 to 6).
- The **range** of X is {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}.
- Are the 36 outcomes equally likely?
- Yes, if the dice are not biased.
- Are the 11 values of X equally likely to occur?
- No, the probabilities vary across values.

Expected Values

$P(X = 2) = 1/36$

$P(X = 3) = 2/36 = 1/18$

$P(X = 4) = 3/36 = 1/12$

$P(X = 5) = 4/36 = 1/9$

$P(X = 6) = 5/36$

$P(X = 7) = 6/36 = 1/6$

$P(X = 8) = 5/36$

$P(X = 9) = 4/36 = 1/9$

$P(X = 10) = 3/36 = 1/12$

$P(X = 11) = 2/36 = 1/18$

$P(X = 12) = 1/36$

Expected Values

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

- $E(X) = 7$

- This means that if we roll two dice many times, sum all the numbers that appear and divide the sum by the number of trials, we expect to find a value of 7.

Expected Values

- **Theorem:**
- If X and Y are random variables on a sample space S, then $E(X + Y) = E(X) + E(Y)$.
- Furthermore, if X_i, $i = 1, 2, \ldots, n$ with a positive integer n, are random variables on S, then $E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$.
- Moreover, if a and b are real numbers, then $E(aX + b) = aE(X) + b$.

Expected Values

- Knowing this theorem, we could now solve the previous example much more easily:

- Let X_1 and X_2 be the numbers appearing on the first and the second die, respectively.

- For each die, there is an equal probability for each of the six numbers to appear. Therefore, $E(X_1) = E(X_2) = (1 + 2 + 3 + 4 + 5 + 6)/6 = 7/2$.

- We now know that $E(X_1 + X_2) = E(X_1) + E(X_2) = 7$.

Expected Values

- We can use our knowledge about expected values to compute the average-case complexity of an algorithm.
- Let the sample space be the set of all possible inputs $a_1, a_2, \ldots, a_n$, and the random variable X assign to each a_j the number of operations that the algorithm executes for that input.
- For each input a_j, the probability that this input occurs is given by $p(a_j)$.
- The algorithm's average-case complexity then is:
- $E(X) = \sum_{j=1,\ldots,n} p(a_j)X(a_j)$

Expected Values

- However, in order to conduct such an average-case analysis, you would need to find out:

  - the number of steps that the algorithms takes for any (!) possible input, and
  - the probability for each of these inputs to occur.

- For most algorithms, this would be a highly complex task, so we will stick with the worst-case analysis.

Independent Random Variables

- **Definition:** The random variables X and Y on a sample space S are **independent** if
- $p(X = r_1 \wedge Y = r_2) = p(X = r_1) \cdot p(Y = r_2)$.

- In other words, X and Y are independent if the probability that $X = r_1 \wedge Y = r_2$ equals the product of the probability that $X = r_1$ and the probability that $Y = r_2$ for all real numbers r_1 and r_2.

Independent Random Variables

- **Example:** Are the random variables X_1 and X_2 from the “pair of dice” example independent?

- **Solution:**
- $p(X_1 = i) = 1/6$
- $p(X_2 = j) = 1/6$
- $p(X_1 = i \wedge X_2 = j) = 1/36$

- Since $p(X_1 = i \wedge X_2 = j) = p(X_1 = i) \cdot p(X_2 = j)$,

the random variables X_1 and X_2 are **independent**.

Independent Random Variables

- **Theorem:** If X and Y are independent random variables on a sample space S, then $E(XY) = E(X)E(Y)$.

- **Note:**
- $E(X + Y) = E(X) + E(Y)$ is true for any X and Y, but
- $E(XY) = E(X)E(Y)$ needs X and Y to be independent.

- How come?

Independent Random Variables

- **Example:** Let X and Y be random variables on some sample space, and each of them assumes the values 1 and 3 with equal probability.
- Then $E(X) = E(Y) = 2$
- If X and Y are **independent**, we have:
- $E(X + Y) = 1/4\cdot(1 + 1) + 1/4\cdot(1 + 3) + 1/4\cdot(3 + 1) + 1/4\cdot(3 + 3) = 4 = E(X) + E(Y)$
- $E(XY) = 1/4\cdot(1\cdot1) + 1/4\cdot(1\cdot3) + 1/4\cdot(3\cdot1) + 1/4\cdot(3\cdot3) = 4 = E(X)\cdot E(Y)$

Independent Random Variables

- Let us now assume that X and Y are **correlated** in such a way that Y = 1 whenever X = 1, and Y = 3 whenever X = 3.
- $E(X + Y) = 1/2\cdot(1 + 1) + 1/2\cdot(3 + 3)$
  $= 4 = E(X) + E(Y)$
- $E(XY) = 1/2\cdot(1\cdot1) + 1/2\cdot(3\cdot3)$
  $= 5 \neq E(X)\cdot E(Y)$

UMASS BOSTON