University of California, Los Angeles
Department of Statistics

Nicolas Christou
Ivo Dinov

Probability

Probability: A measure of the chance that something will occur.

1. Random experiment:
   A process that results in one of possible outcomes. The outcomes cannot be predicted with certainty.

   Examples: Flip a coin, roll a die, roll two dice, draw a card, etc.

2. Sample space of a random experiment:
   It is the set of all possible outcomes of the random experiment, denoted with S.

   Examples:

   a. Flip a coin: $S = \{H, T\}$.
   b. Flip two coins: $S = \{HH, TT, HT, TH\}$.
   c. Roll a die: $S = \{1, 2, 3, 4, 5, 6\}$.
   d. Draw a card: $S = \{A\clubsuit, A\spadesuit, A\diamondsuit, A\heartsuit, \cdots\}$ (all 52 cards).

   See SOCR demos and applets here:

   `http://wiki.stat.ucla.edu/socr/index.php/SOCR_EduMaterials_ExperimentsActivities`
   `http://socr.ucla.edu/htmls/exp/`

3. Event:
   It is the outcome of an experiment, denoted with uppercase letter. It is a subset of the sample space.
   Examples:

   a. Flip a coin: $A = \{H\}$.
   b. Roll a die: $A = \{\text{even number}\}$, $B = \{\text{odd number}\}$, $C = \{1, 2, 3\}$.
   d. Draw a card: $A = \{\text{Ace}\}$, $C = \{\text{Clubs}\}$.

4. Probability of an event A for equally likely outcomes:

$$P(A) = \frac{\text{number of ways in which } A \text{ occurs}}{\text{number of ways in which all outcomes occur}}$$

Examples:

a. Draw a card. Let $A = \{\text{Ace}\}$. Then $P(A) = \frac{4}{52}$.

b. Roll two dice. There are 6×6 possible outcomes. The sum of the two numbers rolled are shown below:

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Let $A = \{\text{sum=5}\}$. Then $P(A) = \frac{4}{36}$.

See SOCR activity and applet:

`http://wiki.stat.ucla.edu/socr/index.php/SOCR_EduMaterials_Activities_DiceExperiment`
`http://www.socr.ucla.edu/htmls/SOCR_Experiments.html`

4. Basic principle of counting: If an experiment has m outcomes and if for every outcome of this experiment there are n outcomes of another experiment then all together there are $m \times n$ outcomes.

Examples:

a. Roll two dice: $6 \times 6 = 36$ outcomes.

b. Flip two coins: $2 \times 2 = 4$ outcomes.

Similarly:

a. Roll three dice: $6 \times 6 \times 6 = 216$ outcomes.

b. Flip three coins: $2 \times 2 \times 2 = 8$ outcomes.

Union of two events A, B

The union of two events A, B, denoted $A \cup B$, is a new event. It is defined as the event containing all outcomes in A or B or BOTH. Similarly the union of n events $A_1, A_2, \cdots, A_n$ is denoted with $A_1 \cup A_2 \cup \cdots \cup A_n$. Key word: *OR*.

Example: Suppose 50 students can be classified by their major and year as follows:

	Sop.	Jun.	Sen.	Total
Econ	10	20	6	36
Math	5	4	5	14
Total	15	24	11	50

A student is selected at random from this group of 50 students. Let $A = \{\text{student is senior}\}$, and $B = \{\text{student is math major}\}$. Find $P(A \cup B)$.

Intersection of two events A, B

The intersection of two events A, B, denoted $A \cap B$, is a new event. It is defined as the event containing all outcomes that belong to both A and B. Similarly the intersection of n events $A_1, A_2, \cdots, A_n$ is denoted with $A_1 \cap A_2 \cap \cdots \cap A_n$. Key word: *AND*.

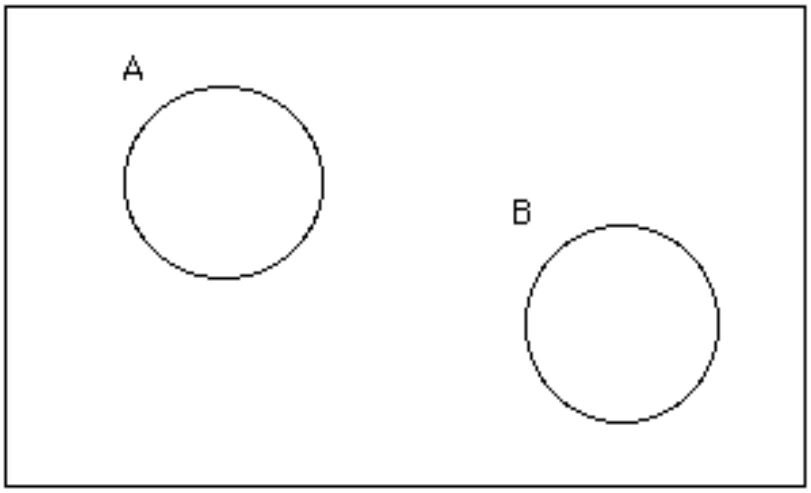
Example: Using the table above find $P(A \cap B)$.

Example: Draw a card at random. Let $A = \{\text{card is clubs}\}$ and $B = \{\text{card is ace or 4}\}$. Find $P(A \cap B)$. Also find $P(A \cup B)$. Use a Venn diagram to show these events.

Mutually exclusive events

It is said that two events A, B are mutually exclusive, (or disjoint) if A, B have no outcomes in common. It follows that A, B are disjoint if $A \cap B = \{\emptyset\}$. Therefore, if A, B are disjoint then $P(A \cap B) = 0$.

Example: Roll a die. Let $A = \{\text{even number}\}$, $B = \{\text{odd number}\}$. The events A, B are mutually exclusive and we can represent them as follows using a Venn diagram.



Axioms of Probability

1. $0 \leq P(A) \leq 1$.
2. $P(S) = 1$.
3. The union for any sequence of mutually exclusive events $A_1, A_2, \cdots$ is equal to $P(A_1 \cup A_2 \cdots) = P(A_1) + P(A_2) + \cdots$
   Let A, B two mutually exclusive event. Then the union of the two events is equal to $P(A \cup B) = P(A) + P(B)$.

Complement events

The complement of an event A, denoted with A' or A^c, is defined to be the event that contains all the outcomes in the sample space that do not belong to A.

Are A, A' mutually exclusive?

Conditional Probability

Let two events A and B. Suppose that event B has already occured. Given this information, what is the probability of event A. This is called "conditional probability" denoted with $P(A|B)$ and we read: the probability of A given B. It is equal to:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Similarly,

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Example: Roll a die. Let $A = \{2\}$, and $B = \{\text{even number}\}$. Draw the two events using a venn diagram. Find $P(A|B)$.

- **Multiplication rule**

  Let events A, B. The probability of A and B (i.e. the probability of the intersection A, B) can be computed as follows:
  $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$.

  Example:
  Draw two cards without replacement. Let's define the events A_1, A_2 as follows:
  $A_1 = \{\text{the first card is an ace}\}$, and $A_2 = \{\text{the second card is an ace}\}$. Find the probability that both cards are aces.

- **Independent events**

  The events A, B are independent if the occurrence of one of them does not affect the occurrence of the other event. If A, B are independent then, $P(A|B) = P(A)$, and the multiplication rule is: $P(A \cap B) = P(A)P(B)$.

  Example: Draw two cards with replacement. Let's define the events A_1, A_2 as follows:
  $A_1 = \{\text{the first card is an ace}\}$, and $A_2 = \{\text{the second card is an ace}\}$. Find the probability that both cards are aces.

- **Addition rule**
  Let events A, B. The probability of A or B or BOTH (i.e. the probability of the union A, B) can be computed as follows:
  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. It follows directly from the Venn diagram.

- If A, B are mutually exclusive events then $P(A \cap B) = 0$, and the addition rule is:
  $P(A \cup B) = P(A) + P(B)$ (this is Axiom 3).

- If A, B are independent with $P(A) > 0$ and $P(B) > 0$, can they be mutually exclusive?

Examples:

1. Draw two cards with replacement. Let's define the events A_1, A_2 as follows: $A_1 = \{\text{the first card is an ace}\}$, and $A_2 = \{\text{the second card is an ace}\}$. Find the probability that the first card is an ace or the second card is an ace, or both cards are aces. Or simply we can ask: Find the probability that we observe at least one ace.

2. Two dice are rolled 10 times. Find the probability that in these 10 rolls we observe the sum of 5 at least once.

De Morgan's Law

Very useful!
Using De Morgan's law we can find the probability of the union of the complements and the intersection of the complements of a number of events. More specifically:

$$P(A' \cap B') = 1 - P(A \cup B)$$

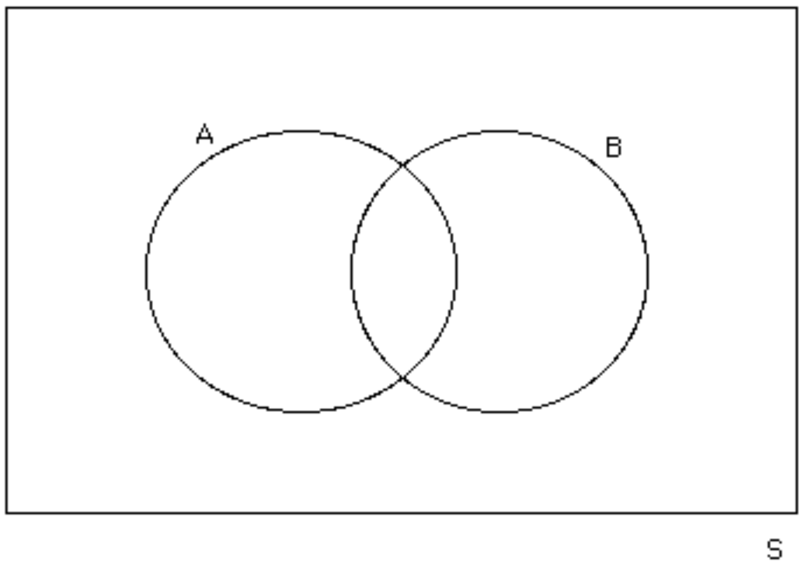
and

$$P(A' \cup B') = 1 - P(A \cap B)$$

Use a Venn diagram to show these results.

Law of total probability
Decomposition rule

Let event A, B as shown in the figure below.



Event A can be expressed as:

$$A = (A \cap B) \cup (A \cap B')$$

The events $A \cap B$ and $A \cap B'$ are mutually exclusive and therefore from Axiom 3 we have:

$$P(A) = P(A \cap B) + P(A \cap B')$$

Using the multiplication rule the above can be expressed as:

$$P(A) = P(A|B)P(B) + P(A|B')P(B')$$

This is called the law of total probability or decomposition rule. It allows us to compute the probability of an event when it is not easy to compute it directly. This rule can be extended to more than two events. If A intersects with the events B, C and D we have: $P(A) = P(A|B)P(B) + P(A|C)P(C) + P(A|D)P(D)$.

Examples:

1. A blood test is 95% effective in detecting a certain disease when the disease in fact is present. The test also yields a false positive result 1% of the healthy persons tested. Suppose that 0.5% (0.005) of the population actually has the disease.

   a. A person will be tested for the disease. Find the probability that the test will be positive.

   b. Suppose that a person was found positive. Find the probability that the person actually has the disease.

2. There are 2 urns. In the first urn there are 5 green balls and 4 blue balls. In the second urn there are 4 white balls, 2 yellow balls, and 2 green balls. One ball is drawn at random from the first urn and it is placed in the second urn. Then a ball is drawn at random from the second urn.

   a. Find the probability that the ball drawn from the second urn is green.

   b. Given that the ball drawn from the second urn is green what is the probability that the ball drawn from urn 1 was also green.

3. A machine that produces parts for cars is in good condition 90% of the time. When the machine is in good condition 1% of the parts are defective. When it is not in good condition 15% of the parts are defective.

   a. Find the probability that a randomly selected part is defective.

   b. Suppose that a randomly selected part was found to be defective. What is the probability that the machine was not in good condition?

4. The marketing manager of a toy manufacturing firm is planning to introduce a new toy into the market. In the past, 40% of the toys introduced by the company have been succesful and 60% have not been succesful. Before the toy is actually marketed, market research is conducted and a report, either favorable or unfavorable, is compiled. In the past, 80% of the succesful toys received favorable reports and 30% of the unsuccesful toys also received favorable reports.

   a. What proportion of the new toys receive favorable market research reports?

   b. Suppose that market research gives a favorable report on a new toy. What is the probability that the new toy will be succesful?

Probability - Summary

- Conditional probability:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

- Multiplication rule:

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

If A, B are independent then $P(A \cap B) = P(A)P(B)$.

- Addition rule:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

If A, B are mutually exclusive then $P(A \cup B) = P(A) + P(B)$.

- De Morgan's Law:

$$P(A' \cap B') = 1 - P(A \cup B)$$

$$P(A' \cup B') = 1 - P(A \cap B)$$

- Useful result (complement events):

$$P(A) + P(A') = 1 \Rightarrow P(A') = 1 - P(A)$$

and

$$P(A|B) + P(A'|B) = 1$$

- If A, B are mutually exclusive then $P(A|B) = 0$ and if A, B are independent then $P(A|B) = A$.

Probability - examples

Example 1:
For each of the following, list the sample space and tell whether you think the events are equally likely.

a. Roll two dice; record the sume of the numbers.

b. A family has 3 children; record the genders in order of birth.

c. Toss 4 coins; record the number of tails.

d. Toss a coin 10 times.; record the longest run of heads.

Example 2:
Suppose the probability that a U.S. resident has traveled to Canada is 0.18, to Mexico is 0.09, and to both countries s 0.04. What is the probability that an American chosen at random has

a. traveled to Canada but not to Mexico?

b. traveled to either Canada or Mexico?

c. not traveled to either country?

Example 3:
Employment data at a large company reveal that 72% of the workers are married, that 44% are college graduates, and that half of the grads are married. What is the probability that a randomly chosen worker

a. is neither married nor a college graduate?

b. is married but not a college graduate?

c. is married or a college graduate?

Example 4:
Seventy percent of kids who visit a doctor have fever, and 30% of kids with fever have sore throats. What is the probability that a kid who goes to the doctor has a fever and a sore throat?

Example 5:
You pick three cards at random from a deck. Find the probability of each event described below.

a. You get no aces.

b. You get all hearts.

c. The third card is your first black card.

d. You have at least one diamond.

Example 6:
The soccer team's shirts have arrived in a big box, and people just start grabbing them, looking for the right size. The box contains 4 medium, 10 large, and 6 extra-large shirts. You want a medium for you and one for your sister. Find the probability of each event described.

a. The first two you grab are the wrong sizes.

b. The first medium shirt you find is the third one you check.

c. The first four shirts you pick are all extra-large.

d. At least one of the first four shirts you check is a medium.

Example 7:
Given the table below, are high blood pressure and high cholesterol independent? Explain.

		Blood Pressure	
		High	OK
	High	0.11	0.21
Cholesterol	OK	0.16	0.52

Example 8:
Suppose that 23% of adults smoke cigarettes. It is known that 57% of smokers and 13% of nonsmokers develop a certain lung condition by age 60.

a. Explain how these statistics indicate that lung condition and smoking are not independent.

b. What is the probability that a randomly selected 60-year-old has this lung condition?

Example 9:
The 60 students of a statistics class can be classified as below according to their academic status and whether they live on or off campus.

	On campus	Off campus	Total
Sophomore	15	5	20
Junior	10	5	15
Senior	15	10	25
Total	40	20	60

A student from this class is selected at random. Let the events A, J, C, S be as follows: $A = \{\text{senior}\}$, $J = \{\text{junior}\}$, $C = \{\text{on campus}\}$, $S = \{\text{sophomore}\}$. Find the probability that the student:

a. lives on campus.

b. is a senior.

c. does not live on campus.

d. lives on campus and is a junior.

e. lives on campus or is a junior.

f. neither lives on campus nor is a junior.

g. does not live on campus given that he is a senior.

h. lives on campus given that he is a senior.

i. is a senior and sophomore.

j. lives on campus or lives off campus.

k. lives on campus and is not senior.

l. does not live on campus or is not a sophomore.

m. Are the events senior and campus independent?

Example 10:
An unbiased die in the shape of regular dodecahedron has twelve faces with the numbers 2,2,4,4,4,6,6,10,10,10,12,12, shown separately on the faces. The result of a throw is the number shown on the uppermost face. Each of four players throws the die twice and scores the sum of the two results.

a. What is the probability that a player will get a score greater than six? (Ans. 0.8888).

b. What is the probability that exactly one of the four players will get a score greater than six? Assume that their scores are independent from one another. (Ans. 0.0049).

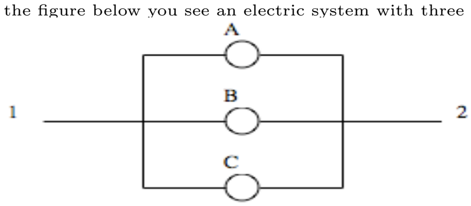
Example 11:
The color of a person's eyes is determined by a single pair of genes. If they are both blue-eyed genes, then the person will have blue eyes; if they are both brown-eyed genes, then the person will have brown eyes; and if one of them is blue-eyed gene and the other a brown-eyed gene, then the person will have brown eyes (because of the latter fact we say that the brown-eyed gene is *dominant* over the blue-eyed one). A newborn child independently receives one eye gene from each of its parents and the gene it receives from a parent is equally likely to be either of the two gnes of that parent. Suppose that person A and both of his parents have brown eyes, but person A's sister has blue eyes.

a. What is the probabilty that person A possesses a blue-eyed gene? (Ans. $\frac{2}{3}$).

Suppose person A's wife has blue eyes.

b. What is the probability that their first child will have blue eyes? (Ans. $\frac{1}{3}$).

Problem 12
On the figure below you see an electric system with three components that operate independently each one with probability 95%.



a. Find the probability that current will flow from point 1 to point 2.

b. Given that current flowed from point 1 to point 2, what is the probability that component A functioned?

c. Let X be the number of paths that allow current to flow from point 1 to point 2. Construct the probability distribution of X. Note: $AB'C'$ is one path because A functions while B and C do not function. Also, X takes the value 0, 1, 2, 3. If $x = 0$ then there is no path that allows current to flow, while if $x = 3$ there are three paths that allow current to flow, etc.

Problem 13
A medical doctor thinks carefully about the following dilemma: "If I am at least 80 percent certain that my patient has this disease, then I always recommend surgery. If am not so certain, then I recommend additional tests that are expensive and sometimes painful. Now, initially I was only 60 percent certain that patient X had the disease, so I ordered test A. Test A always gives positive result if the patient has the disease. The test was positive for patient X and I was ready to recommend surgery when X informed me, for the first time, that he is diabetic. This information complicates matters! I still believe that the probability that X has the disease is 60 percent but the interpretation of the results of test A should be different. This is because test A yields false positive result 30 percent of the time in the case of a diabetic patient. Now what do I do? Should I recommend surgery?" Let denote with D the event that patient X has the disease, and with E the event that test A is positive.

a. Compute the probability that test A is positive.

b. Should the doctor recommend surgery? To help the doctor compute the probability that the patient has the disease given that the test is positive.

The game of craps

A popular gambling game called *craps* is described below. A player rolls two dice, and the sum of the two numbers that appear is recorded. If the sum on the first roll is 7 or 11, the player wins immediately and the game stops. If on the first attempt a sum of 2, or 3, or 12 is scored, the player loses the game immediately. If the sum on the first roll is 4, 5, 6, 8, 9, or 10, then the two dice are rolled again until either the sum of 7 is scored or the original sum is scored. If the original sum is obtained a second time before 7, then the player wins. If the sum of 7 is obtained before the original sum is obtained a second time, then the player loses. Here are some examples:
First roll: sum is 11 then player wins.
First roll: sum is 12 then player loses.
First roll: sum is 5, second roll: sum is 5 then player wins.
First roll: sum is 8, second roll: sum is 2, third roll: sum is 8, then player wins.
First roll: sum is 4, second roll: sum is 11, third roll: sum is 7, then player loses.

Compute the probability that the player will win.

Monty Hall experiment - Let's make a deal...

In this experiment there are three doors. Behind one of the doors there is a prize, a new car! Behind the other two doors there are no prizes. The game begins as follows. The player selects his/her door and then the host - who knows where the prize is located - opens one of the two doors that have no prize. So far we have the player's door closed and one of the other two doors opened. At this point the host makes the offer to the player to stay with his/her initial choice or to switch to the other closed door. What should the player do?

Poker Combinations

A Poker hand consists of five cards. Players try for combinations of two or more cards of a kind, five-card sequences, or five cards of the same suit. Poker is played with a standard 52-card deck in which all suits are of equal value, the cards ranking from the ace high, downward through king, queen, jack, and the numbered cards 10 to the deuce. The ace may also be considered low to form a straight (sequence) ace through five as well as high with king-queen-jack-10.

Here is the traditional ranking of hands with some examples:

1. Royal Flush (ten-jack-queen-king-ace all of the same suit): $10\clubsuit, J\clubsuit, Q\clubsuit, K\clubsuit, A\clubsuit$.
2. Straight Flush (five cards of the same suit in sequence other than royal flush): $2\spadesuit, 3\spadesuit, 4\spadesuit, 5\spadesuit, 6\spadesuit$.
3. Four of a kind, plus any fifth card: $5\clubsuit, 5\spadesuit, 5\diamondsuit, 5\heartsuit, 4\clubsuit$.
4. Full house (three of a kind plus a pair): $5\clubsuit, 5\spadesuit, 5\diamondsuit, 4\heartsuit, 4\clubsuit$.
5. Flush (any five cards of the same suit other than straight flush): $3\clubsuit, 6\clubsuit, 8\clubsuit, 10\clubsuit, J\clubsuit$.
6. Straight (five cards in sequence not all of the same suit): $2\spadesuit, 3\heartsuit, 4\spadesuit, 5\clubsuit, 6\diamondsuit$.
7. Three of a kind plus no pair: $5\clubsuit, 5\spadesuit, 5\diamondsuit, 4\heartsuit, 6\clubsuit$.
8. Two pairs plus any fifth card: $5\clubsuit, 5\spadesuit, 4\diamondsuit, 4\heartsuit, 9\clubsuit$.
9. One pair plus three other cards: $5\clubsuit, 5\spadesuit, 4\diamondsuit, 6\heartsuit, 10\clubsuit$.
10. Nothing

Rank	Poker Hand	# of Combinations
1.	Royal Flush	4
2.	Other Straight Flush	36
3.	Four of a kind	624
4.	Full House	3744
5.	Flush	5108
6.	Straight	10200
7.	Three of a kind	54912
8.	Two Pairs	123552
9.	One Pair	1098240
10.	Nothing	1302540
	Total	2598960

See poker experiment on SOCR:

`http://wiki.stat.ucla.edu/socr/index.php/SOCR_EduMaterials_Activities_PokerExperiment`

A standard 52-card deck consists of the following 13 denominations (4 cards in eac denomination), and 4 suits (13 cards in each suit):

♣	♠	♢	♡
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
9	9	9	9
10	10	10	10
J	J	J	J
Q	Q	Q	Q
K	K	K	K
A	A	A	A

See SOCR poker calculations here:

`http://wiki.stat.ucla.edu/socr/index.php/`
`AP_Statistics_Curriculum_2007_Prob_Count#Poker_Game_Calculations`

There are $\binom{52}{5} = \frac{52!}{(52-5)!5!} = 2598960$ different ways to choose 5 cards from the available 52 cards.

Let's compute the number of combinations of the following poker hand: four of kind plus any fifth card: We need 2 different denominations (for example 4 aces plus an eight). There are $\binom{13}{2}$ different ways to choose 2 denominations from the 13 available denominations. Now, there are $\binom{4}{4}$ ways to choose the 4 aces from the 4 aces, and $\binom{4}{1}$ different ways to choose one eight from the 4 eights. Also, the opposite could have occured (that is, 4 eights and 1 ace). Therefore the number of combinations of a four of a kind plus any fifth card is:

$$\binom{13}{2}\binom{4}{4}\binom{4}{1}2 = 624.$$

Similarly, we can find the number of combinations of a pair plus 3 other cards (for example 2 aces plus 1 two, 1 seven, 1 eight). We need 4 denominations. There are $\binom{13}{4}$ different ways to choose 4 denominations from the available 13 denominations. Now, there are $\binom{4}{2}$ different ways to choose a pair from one of the 4 denominations, and $\binom{4}{1}$ different ways to choose one card from each of the other 3 denominations. Also, we multiply by 4 because the pair could have come from the 2s or the 7s or the 8s. Therefore the number of combinations of a pair plus 3 other cards is:

$$\binom{13}{4}\binom{4}{2}\binom{4}{1}\binom{4}{1}\binom{4}{1}4 = 1098240.$$

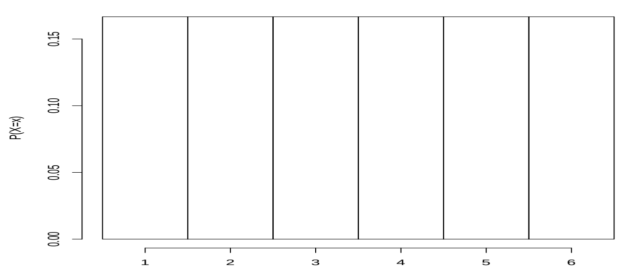
For the examples above, to compute the probability that a player receives a four of kind plus any fifth card or a pair plus 3 other cards we simply divide the number of combinations that this particular poker hands occurs over the total number of ways in which 5 cards can be selected from the available 52 cards.

$$P(\text{four of a kind plus a fifth card}) = \frac{\binom{13}{2}\binom{4}{4}\binom{4}{1}2}{\binom{52}{5}} = \frac{64}{2598960} = 0.0000246.$$

Introduction to random variables

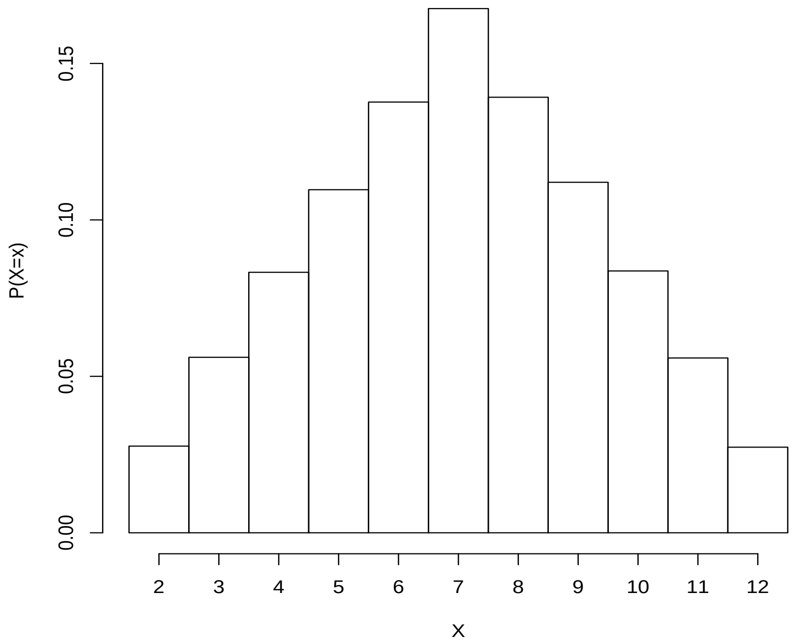
- Discrete random variables.
- Continuous random variables.
- **Discrete random variables**. Denote a discrete random variable with X:
  It is a variable that takes values with some probability.
  Examples:
  a. Roll a die. Let X be the number observed.
  b. Draw 2 cards with replacement. Let X be the number of aces among the 2 cards.
  c. Roll 2 dice. Let X be the sum of the 2 numbers observed.
  d. Toss a coin 5 times. Let X be the number of tails among the 5 tosses.
  e. Randomly select a US household. Let X be the number of people live in this household.
- **Probability distribution of a discrete random variable *X***
  It is the list of all possible values of X with the corresponding probabilities. It can be represented by a table, a graph, or a function.
  Examples:
  a. Roll a die. Let X be the number observed. The probability distribution of X is:

X	$P(X=x)$
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$



b. Roll two dice. Let X be the sum of the two numbers observed. The probability distribution of X is:

X	$P(X = x)$
2	$\frac{1}{36}$
3	$\frac{2}{36}$
4	$\frac{3}{36}$
5	$\frac{4}{36}$
6	$\frac{5}{36}$
7	$\frac{6}{36}$
8	$\frac{5}{36}$
9	$\frac{4}{36}$
10	$\frac{3}{36}$
11	$\frac{2}{36}$
12	$\frac{1}{36}$



We can also represent this distribution with a function: $P(X = x) = \frac{6-|x-7|}{36}$, for $x = 2, 3, \cdots, 12$. This is called probability mass function and returns the probability for each possible value of the random variable X.

- **Expected value (or mean) of a discrete random variable**
  It is denoted with $E(X)$ or μ and it is computed as follows:
  Definition:

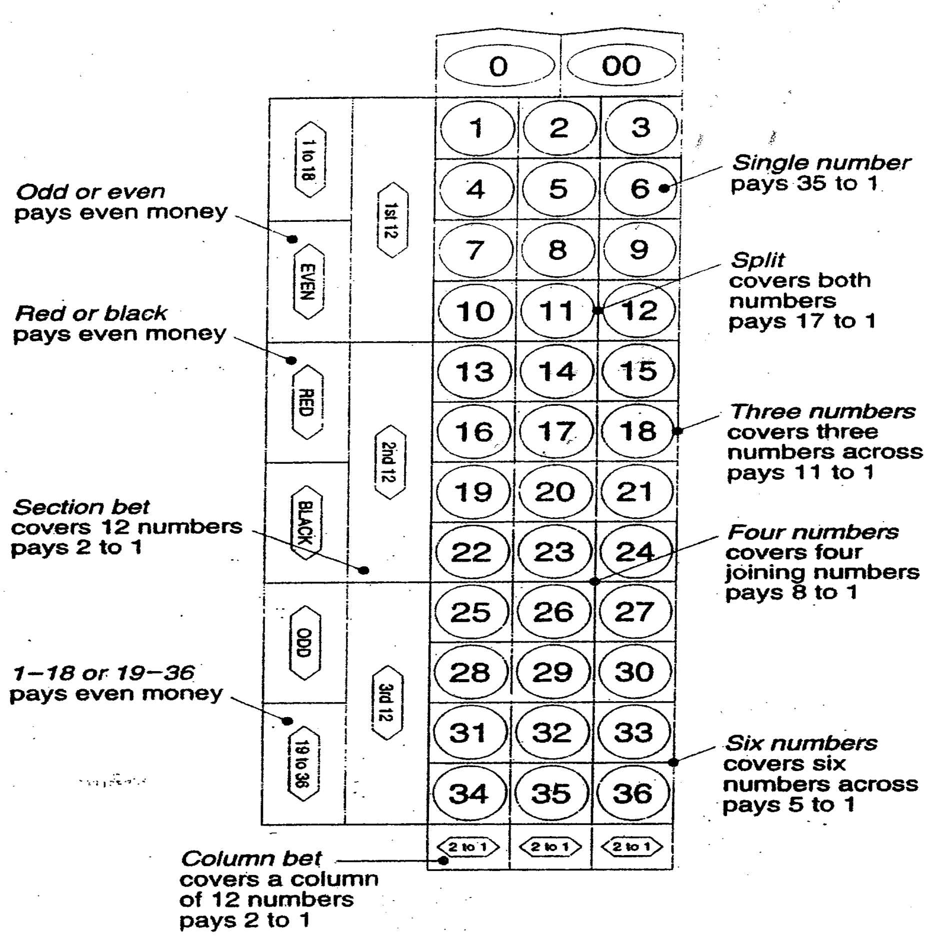
$$\mu = E(X) = \sum_x xP(X = x)$$

It is a weighted average. The weights are the probabilities.

Example:
Roll a die. Let X be the number observed. Find the expected value of X. The $E(X)$ must be somewhere between 1, 6: $E(X) = 1\frac{1}{6} + 2\frac{1}{6} + 3\frac{1}{6} + 4\frac{1}{6} + 5\frac{1}{6} + 6\frac{1}{6} = 3.5$.
What does this number mean?

Example: Casino roulette. Below you see the standard Nevada roulette table:

A NEVADA ROULETTE TABLE



Roulette is a pleasant, relaxed, and highly comfortable way to lose your money.
—JIMMY THE GREEK

FROM: "STATISTICS"
DAVID FREEDMAN
ROBERT PISANI
ROGER PURVES

A player will bet 1$ on four joining numbers. This bet pays 8 : 1. Let X be the player's payoff. Find the player's expected profit. See SOCR roulette experiment:

`http://wiki.stat.ucla.edu/socr/index.php/SOCR_EduMaterials_Activities_RouletteExperiment`

- **Expected value of a sum of random variables**
  Let X and Y be 2 random varables. The expected value of the sum of these 2 random variables is:

  $$E(X+Y) = E(X) + E(Y)$$

  Example:
  Roll 2 dice. Let X be the number observed on the first die and Y be the number observed on the second die. Let W be the sum of the 2 dice. Find the expected value of W. There are 2 ways to solve this problem:

  a. $E(W) = E(X+Y) = E(X) + E(Y) = 3.5 + 3.5 = 7$
  b. Or using the distribution of the sum of the two dice (see page 2):

  $$E(W) = \sum_w wP(W = w) = 2\frac{1}{36} + 3\frac{2}{36} + \cdots + 12\frac{1}{36} = 7.$$

  The expected value of the sum can be extended to more than two random variables:

  $$E(X+Y+Z+\cdots) = E(X) + E(Y) + E(Z) + \cdots$$

SOCR experiments activities:

`http://wiki.stat.ucla.edu/socr/index.php/SOCR_EduMaterials_ExperimentsActivities`

- **Variance and standard deviation of a discrete random variable**
  Consider the following 2 probability distributions:

X	$P(X)$	Y	$P(Y)$
0	$\frac{1}{2}$	5	$\frac{1}{2}$
1	$\frac{1}{2}$	-4	$\frac{1}{2}$

  What do you observe?

  Definition:

$$Var(X) = \sigma^2 = E(X-\mu)^2 = \sum_x^n (x-\mu)^2)P(X=x) = \sum_x x^2 P(X=x) - \mu^2$$

  The standard deviation of a discrete random variable is the square root of the variance:

$$SD(X) = \sqrt{\sigma^2} = \sqrt{\sum_x^n (x-\mu)^2)P(X=x) = \sum_x x^2 P(X=x) - \mu^2}$$

  It follows that:

$$\sigma^2 = EX^2 - \mu^2 \text{ or } EX^2 = \sigma^2 + \mu^2$$

  Example:
  Roll a die. Let X be the number observed. Find the variance of X.

$$Var(X) = 1^2\frac{1}{6} + 2^2\frac{1}{6} + 3^2\frac{1}{6} + 4^2\frac{1}{6} + 5^2\frac{1}{6} + 6^2\frac{1}{6} - 3.5^2 = 2.917.$$

  The standard deviation is:

$$SD(X) = \sqrt{2.917} = 1.708.$$

- **Some properties of expectation and variance**
  Let X, Y random variables and a, b constants.

  a. $E(aX) = aE(X)$
  b. $E(aX+b) = aE(X) + b$
  c. $Var(X+a) = Var(X)$
  d. $Var(aX) = a^2Var(X)$
  e. $Var(aX+b) = a^2Var(X)$.
  f. If X, Y are independent then $Var(X+Y) = Var(X) + Var(Y)$

Example:
An insurance policy costs \$100, and will pay policyholders \$10000 if they suffer a major injury (resulting in hospitalization) or \$3000 if they suffer a minor injury (resulting in lost time from work). The company estimates that each year 1 in every 2000 policyholders may have a major injury, and 1 in 500 a minor injury.

a. Construct the probability distribution for the profit on a policy.

b. What is the company's expected profit on this policy?

c. Do you think the standard deviation is large or small. Why?

d. Compute the standard deviation.

e. Suppose that the company writes (a) 36, (b) 10000 of these policies per year. What are the mean and standard deviation of the annual profit for these 2 cases?

f. Comment!

Solution:

a. Let X the profit of the company on one of these insurance policies. The probability distribution of X is:

X	$P(x)$
-9900	0.0005
-2900	0.002
100	0.9975

b. Expected value of X:
$E(X) = -9900(0.0005) - 2900(0.002) + 100(0.9975) = 89.$

c. The standard deviation will be large.

d. Variance of X:
$var(X) = (-9900)^2(0.0005) + (-2900)^2(0.002) + (100)^2(0.9975) - 89^2 = 67879$. Therefore the standard deviation is: $sd(X) = \sqrt{67879} = 260.54$.

e. Here we need to find the expected value and variance of sum of random variables. In part (i) we have a sum of 36 random variables and in part (ii) a sum of 10000 variables.
Part(i):
$E(Y_1 + \ldots + Y_{36}) = 36(89) = 3204.$
$var(Y_1 + \ldots + Y_{36}) = 36(67879)$. The standard deviation is:
$sd(Y_1 + \ldots + Y_{36}) = \sqrt{36(67879)} = 1563.2.$

Part(ii):
$E(Y_1 + \ldots + Y_{10000}) = 10000(89) = 890000.$
$var(Y_1 + \ldots + Y_{10000}) = 10000(67879)$. The standard deviation is:
$sd(Y_1 + \ldots + Y_{10000}) = \sqrt{10000(67879)} = 26053.6.$