Random Variables

When we perform an experiment, we are often interested in recording various pieces of numerical data for each trial. For example, when a patient visits the doctor's office, their height, weight, temperature and blood pressure are recorded. These observations vary from patient to patient, hence they are called variables.

Random Variables

When we perform an experiment, we are often interested in recording various pieces of numerical data for each trial. For example, when a patient visits the doctor's office, their height, weight, temperature and blood pressure are recorded. These observations vary from patient to patient, hence they are called variables.

We tend to call these *random variables*, because we cannot predict what their value will be for the next trial of the experiment (for the next patient).

Random Variables

When we perform an experiment, we are often interested in recording various pieces of numerical data for each trial. For example, when a patient visits the doctor's office, their height, weight, temperature and blood pressure are recorded. These observations vary from patient to patient, hence they are called variables.

We tend to call these *random variables*, because we cannot predict what their value will be for the next trial of the experiment (for the next patient).

Rather than repeat and write the words height, weight and blood pressure many times, we tend to give random variables names such as X, Y .... We usually use capital letters to denote the name of the variable and lowercase letters to denote the values.

Random Variables

A **Random Variable** is a rule that assigns a number to each outcome of an experiment.

Random Variables

A **Random Variable** is a rule that assigns a number to each outcome of an experiment.

Example: An experiment consists of rolling a pair of dice, one red and one green, and observing the pair of numbers on the uppermost faces (red first). We let X denote the sum of the numbers on the uppermost faces.

Random Variables

A **Random Variable** is a rule that assigns a number to each outcome of an experiment.

Example: An experiment consists of rolling a pair of dice, one red and one green, and observing the pair of numbers on the uppermost faces (red first). We let X denote the sum of the numbers on the uppermost faces. Below, we show the outcomes on the left and the values of X associated to some of the outcomes on the right:

$$
\begin{array}{llllll}
\{(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\
\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\
\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\
\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\
\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\
\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6)\}
\end{array}
$$

Outcome	X
(1, 1)	2
(2, 1)	3
(3, 1)	4
(4, 1)	5
$\vdots$	$\vdots$

Dice example

(a) What are the possible values of X?

Dice example

(a) What are the possible values of X?

$1 + 1 = 2,\ 1 + 2 = 3,\ \ldots,\ 6 + 6 = 12.$

Dice example

(a) What are the possible values of X?

$1 + 1 = 2$, $1 + 2 = 3$, $\ldots$, $6 + 6 = 12$.

(b) There are 2 outcomes of this experiment for which X has a value of 3, namely $(2, 1)$ and $(1, 2)$. How many outcomes are associated with the remaining values of X?

Dice example

(a) What are the possible values of X?

$1 + 1 = 2$, $1 + 2 = 3$, $\ldots$, $6 + 6 = 12$.

(b) There are 2 outcomes of this experiment for which X has a value of 3, namely $(2, 1)$ and $(1, 2)$. How many outcomes are associated with the remaining values of X?

Value of X	**Number of outcomes**	
2	1	(1,1)
3	2	(1,2), (2,1)
4	3	(1,3), (2,2), (3,1)
5	4	(1,4), (2,3), (3,2), (4,1)
6	5	(1,5), (2,4), (3,3), (4,2), (5,1)
7	6	(1,6), (2,4), (3,4), (4,3), (5,2), (6,1)
8	5	(2,6), (3,5), (4,4), (5,3), (6,2)
9	4	(3,6), (4,5), (5,4), (6,3)
10	3	(4,6), (5,5), (4,6)
11	2	(5,6), (6,5)
12	1	(6,6)

Dice example

(c) We could also define other variables associated to this experiment. Let Y be the product of the numbers on the uppermost faces. What are the values of Y associated to the various outcomes?

Dice example

Outcome	Y
(1, 1)	1
(2, 1), (1, 2)	2
(3, 1), (1, 3)	3
(4, 1), (1, 4)	4
(5, 1), (1, 5)	5
(6, 1), (1, 6)	6

Outcome	Y
(3, 3)	9
(4, 3), (3, 4)	12
(5, 3), (3, 5)	15
(6, 3), (3, 6)	18

Outcome	Y
(2, 2)	4
(3, 2), (2, 3)	6
(4, 2), (2, 4)	8
(5, 2), (2, 5)	10
(6, 2), (2, 6)	12

Outcome	Y
(4, 4)	16
(4, 5), (5, 4)	20
(4, 6), (6, 4)	24
(5, 5)	25
(5, 6), (6, 5)	30
(6, 6)	36

Dice example

(d) What are the possible values of Y?

Dice example

(d) What are the possible values of Y?

$1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 16, 18, 20, 24, 25, 30, 36$

Dice example

(d) What are the possible values of Y?

$1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 16, 18, 20, 24, 25, 30, 36$

(e) Draw up a frequency table for these values.

Dice example

(d) What are the possible values of Y?

$1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 16, 18, 20, 24, 25, 30, 36$

(e) Draw up a frequency table for these values.

Value	1	2	3	4	5	6	8	9	10	12	15	16	18	20	24	25	30	36
Frequency	1	2	2	3	2	4	1	1	2	4	2	1	2	2	2	1	2	1

Coin example

Example: An experiment consists of flipping a coin 4 times and observing the sequence of heads and tails. The random variable X is the number of heads in the observed sequence. Draw a table that shows the possible values of X and the number of outcomes associated to each value.

Coin example

Example: An experiment consists of flipping a coin 4 times and observing the sequence of heads and tails. The random variable X is the number of heads in the observed sequence. Draw a table that shows the possible values of X and the number of outcomes associated to each value.

Value of X	outcomes	no. of outcomes
4	HHHH	1
3	HHHT, HHTH, HTHH, THHH	4
2	HHTT, HTHT, HTTH, THHT, THTH, TTHH	6
1	HTTT, THTT, TTHT, TTTH	4
0	TTTT	1

Discrete vs. continuous random variables

For some random variables, the possible values of the variable can be listed in either a finite or an infinite list. These variables are called **discrete random variables**.

Discrete vs. continuous random variables

For some random variables, the possible values of the variable can be listed in either a finite or an infinite list. These variables are called **discrete random variables**. Some examples:

Experiment	**Random Variable**, X
Roll a pair of six-sided dice	Sum of the numbers
Roll a pair of six-sided dice	Product of the numbers
Toss a coin 10 times	Number of tails
Choose a small pack of M&M's at random	The number of blue M&M's in the pack
Choose a year at random	The number of people who ran the Boston Marathon in that year

Discrete vs. continuous random variables

For some random variables, the possible values of the variable can be listed in either a finite or an infinite list. These variables are called **discrete random variables**. Some examples:

Experiment	**Random Variable**, X
Roll a pair of six-sided dice	Sum of the numbers
Roll a pair of six-sided dice	Product of the numbers
Toss a coin 10 times	Number of tails
Choose a small pack of M&M's at random	The number of blue M&M's in the pack
Choose a year at random	The number of people who ran the Boston Marathon in that year

On the other hand, a **continuous** random variable can assume any value in some interval.

Discrete vs. continuous random variables

For some random variables, the possible values of the variable can be listed in either a finite or an infinite list. These variables are called **discrete random variables**. Some examples:

Experiment	**Random Variable**, X
Roll a pair of six-sided dice	Sum of the numbers
Roll a pair of six-sided dice	Product of the numbers
Toss a coin 10 times	Number of tails
Choose a small pack of M&M's at random	The number of blue M&M's in the pack
Choose a year at random	The number of people who ran the Boston Marathon in that year

On the other hand, a **continuous** random variable can assume any value in some interval. Some examples:

Experiment	**Random Variable**, X
Choose a patient at random	Patient's Height
Choose an apple at random at your local grocery store	Weight of the apple
Choose a customer at random at Subway	The length of time the customer waits to be served

Probability Distributions

For a discrete random variable with finitely many possible values, we can calculate the probability that a particular value of the random variable will be observed by adding the probabilities of the outcomes of our experiment associated to that value of the random variable (assuming that we know those probabilities). This assignment of probabilities to each possible value of X is called the *probability distribution* of X.

Dice example

Example If I roll a pair of fair six sided dice and observe the pair of numbers on the uppermost face, all outcomes are equally likely, each with a probability of $\frac{1}{36}$. Let X denote the sum of the pair of numbers observed. We saw that a value of 3 for X is associated to two outcomes in our sample space: $(2, 1)$ and $(1, 2)$. Therefore the probability that X takes the value 3 or $\mathbf{P}(X = 3)$ is the sum of the probabilities of the two outcomes $(2, 1)$ and $(1, 2)$ which is $\frac{2}{36}$. That is

$$\mathbf{P}(X = 3) = \frac{2}{36}.$$

If X is a discrete random variable with finitely many possible values, we can display the probability distribution of X in a table where the possible values of X are listed alongside their probabilities.

Dice example

I roll a pair of fair six sided dice and observe the pair of numbers on the uppermost face. Let X denote the sum of the pair of numbers observed. Complete the table showing the probability distribution of X below:

$$
\begin{matrix}
\{(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\
(2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6)\}
\end{matrix}
$$

X	$\mathbf{P}$(X)
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	

Dice example

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

This table is an example of a probability distribution associated to a random variable.

Probability Distributions

If a discrete random variable has possible values $x_1, x_2, x_3, \ldots, x_k$, then a **probability distribution** $\mathbf{P}(X)$ is a rule that assigns a probability $\mathbf{P}(x_i)$ to each value x_i. More specifically,

- $0 \leq \mathbf{P}(x_i) \leq 1$ for each x_i.
- $\mathbf{P}(x_1) + \mathbf{P}(x_2) + \cdots + \mathbf{P}(x_k) = 1$.

Probability Distributions

If a discrete random variable has possible values $x_1, x_2, x_3, \ldots, x_k$, then a **probability distribution** $\mathbf{P}(X)$ is a rule that assigns a probability $\mathbf{P}(x_i)$ to each value x_i. More specifically,

- $0 \leq \mathbf{P}(x_i) \leq 1$ for each x_i.
- $\mathbf{P}(x_1) + \mathbf{P}(x_2) + \cdots + \mathbf{P}(x_k) = 1$.

Example An experiment consists of flipping a coin 4 times and observing the sequence of heads and tails. The random variable X is the number of heads in the observed sequence. Fill in probabilities for each possible values of X in the table below.

X	0	1	2	3	4
$\mathbf{P}(X)$	?	?	?	?	?

Coin Example

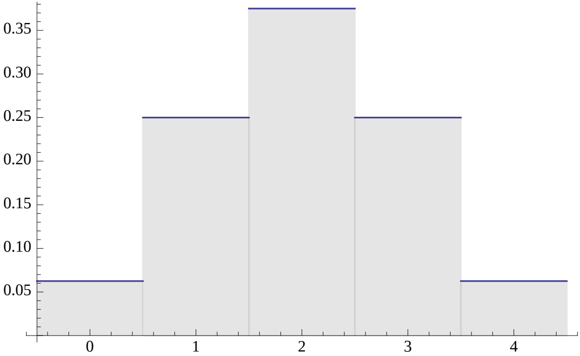
X	0	1	2	3	4
$\mathbf{P}(X)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

Bar graphs of distributions

We can also represent a probability distribution for a discrete random variable with finitely many possible values graphically by constructing a bar graph. We form a category for each value of the random variable (centered at that value) which does not contain any other possible value of the random variable. We make each category of equal width and above each category we draw a bar with height equal to the probability of the corresponding value. If the possible values of the random variable are integers, we can give each bar a base of width 1.

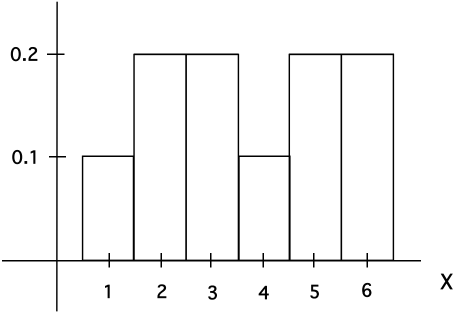
Bar graphs of distributions

Example: An experiment consists of flipping a coin 4 times and observing the sequence of heads and tails. The random variable X is the number of heads in the observed sequence. The following is a graphical representation of the probability distribution of X.



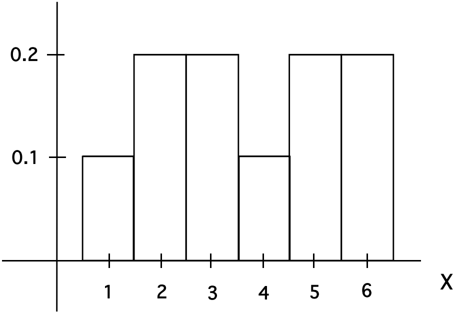
Bar graphs of distributions

Example: The following is a probability distribution histogram for a random variable X.



What is $\mathbf{P}(X \leqslant 5)$?

Bar graphs of distributions



$\mathbf{P}(X \leqslant 5) = \mathbf{P}(X = 5) + \mathbf{P}(X = 4) + \mathbf{P}(X = 3) + \mathbf{P}(X = 1) = 0.2 + 0.1 + 0.2 + 0.2 + 0.1 = 0.8$

OR

$\mathbf{P}(X \leqslant 5) = 1 - \mathbf{P}(X = 6) = 1 - 0.2 = 0.8.$

Some gambling examples

Example: In a carnival game a player flips a coin twice. The player pays \$1 to play. The player then receives \$1 for every head observed and pays \$1, to the game attendant, for every tail observed. Find the probability distribution for the random variable $X =$ the player's (net) earnings.

Some gambling examples

Example: In a carnival game a player flips a coin twice. The player pays \$1 to play. The player then receives \$1 for every head observed and pays \$1, to the game attendant, for every tail observed. Find the probability distribution for the random variable $X =$ the player's (net) earnings.

There are 4 possible outcomes HH, HT, TH, TT. The return to the player in the case HH is 1, the return to the player in the case HT or TH is -1, and the return to the player in the case TT is -3. Hence $\mathbf{P}(X = 1) = \frac{1}{4}$, $\mathbf{P}(X = -1) = \frac{2}{4}$ and $\mathbf{P}(X = -3) = \frac{1}{4}$.

Some gambling examples

A roulette wheel has 18 red numbers, 18 black numbers and 2 green numbers. When the wheel is spun and a ball dropped onto it, the ball is equally likely to land on any of the 38 numbers.

When you bet $1 on red,

- if the ball lands on a red number you get your $1 back plus $1 profit, and
- if the ball lands on a black or a green number, you lose your initial dollar.

What is the probability distribution for your earnings for this game if you bet $1 on red?

Some gambling examples

A roulette wheel has 18 red numbers, 18 black numbers and 2 green numbers. When the wheel is spun and a ball dropped onto it, the ball is equally likely to land on any of the 38 numbers.

When you bet \$1 on red,

- if the ball lands on a red number you get your \$1 back plus \$1 profit, and
- if the ball lands on a black or a green number, you lose your initial dollar.

What is the probability distribution for your earnings for this game if you bet \$1 on red?

There are only two outcomes: you win \$1 or you get $-\$1$.

$$\mathbf{P}(X = 1) = \frac{18}{18 + 18 + 2} = \frac{18}{38}.$$

$$\mathbf{P}(X = -1) = \frac{18 + 2}{18 + 18 + 2} = \frac{20}{38}.$$

Some gambling examples

A roulette wheel has 18 red numbers, 18 black numbers and 2 green numbers. When the wheel is spun and a ball dropped onto it, the ball is equally likely to land on any of the 38 numbers.

When you bet \$1 on red,

- if the ball lands on a red number you get your \$1 back plus \$1 profit, and
- if the ball lands on a black or a green number, you lose your initial dollar.

What is the probability distribution for your earnings for this game if you bet \$1 on red?

There are only two outcomes: you win \$1 or you get −\$1.

$$\mathbf{P}(X = 1) = \frac{18}{18 + 18 + 2} = \frac{18}{38}.$$

$$\mathbf{P}(X = -1) = \frac{18 + 2}{18 + 18 + 2} = \frac{20}{38}.$$

Since $\frac{18}{38}$ is less than 1/2, and $\frac{20}{38}$ is greater than 1/2, you lose more often than you win at Roulette (naturally; otherwise the casino wouldn't offer it!)

A puzzle about roulette

A roulette wheel has 18 red numbers, 18 black numbers and 2 green numbers. When the wheel is spun and a ball dropped onto it, the ball is equally likely to land on any of the 38 numbers.

When you bet \$1 on red,

- if the ball lands on a red number you get your \$1 back plus \$1 profit, and
- if the ball lands on a black or a green number, you lose your initial dollar.

So the probability of winning on a single roll by betting on red is 18/38, which is less that 1/2.

A puzzle about roulette

A roulette wheel has 18 red numbers, 18 black numbers and 2 green numbers. When the wheel is spun and a ball dropped onto it, the ball is equally likely to land on any of the 38 numbers.

When you bet \$1 on red,

- if the ball lands on a red number you get your \$1 back plus \$1 profit, and
- if the ball lands on a black or a green number, you lose your initial dollar.

So the probability of winning on a single roll by betting on red is 18/38, which is less that 1/2.

This seems like a fool's game. But here's a possible strategy for playing it:

1. Begin by betting a dollar on red.
2. If you win, take your winnings and go home.
3. If you lose, place two one-dollar bets in a row on red.
4. Whatever happens on those two rolls, go home (either with your winnings to date, or cutting your losses)

A puzzle about roulette

A roulette wheel has 18 red numbers, 18 black numbers and 2 green numbers. When the wheel is spun and a ball dropped onto it, the ball is equally likely to land on any of the 38 numbers.

When you bet \$1 on red,

- if the ball lands on a red number you get your \$1 back plus \$1 profit, and
- if the ball lands on a black or a green number, you lose your initial dollar.

So the probability of winning on a single roll by betting on red is 18/38, which is less that 1/2.

This seems like a fool's game. But here's a possible strategy for playing it:

1. Begin by betting a dollar on red.
2. If you win, take your winnings and go home.
3. If you lose, place two one-dollar bets in a row on red.
4. Whatever happens on those two rolls, go home (either with your winnings to date, or cutting your losses)

Question: Is this a winning strategy? Specifically, what is the probability that you will leave the roulette wheel with more money than you began with, and is this probability more or less than 1/2?

Solution

Let X be net winnings from this strategy.

Solution

Let X be net winnings from this strategy. Possible outcomes/values for X:

- Win on first roll, probability $18/38 \approx .474$, $X = +1$

Solution

Let X be net winnings from this strategy. Possible outcomes/values for X:

- Win on first roll, probability $18/38 \approx .474$, $X = +1$
- Lose on first, win on next two, probability $(20/38)(18/38)^2 \approx .118$, $X = +1$

Solution

Let X be net winnings from this strategy. Possible outcomes/values for X:

- Win on first roll, probability $18/38 \approx .474$, $X = +1$
- Lose on first, win on next two, probability $(20/38)(18/38)^2 \approx .118$, $X = +1$
- Lose on first, win exactly one of next two, probability $(20/38)2(18/38)(20/38) \approx .262$, $X = -1$

Solution

Let X be net winnings from this strategy. Possible outcomes/values for X:

- Win on first roll, probability $18/38 \approx .474$, $X = +1$
- Lose on first, win on next two, probability $(20/38)(18/38)^2 \approx .118$, $X = +1$
- Lose on first, win exactly one of next two, probability $(20/38)2(18/38)(20/38) \approx .262$, $X = -1$
- Lose all three, probability $(20/38)^3 \approx .146$, $X = -3$.

Solution

Let X be net winnings from this strategy. Possible outcomes/values for X:

- Win on first roll, probability $18/38 \approx .474$, $X = +1$
- Lose on first, win on next two, probability $(20/38)(18/38)^2 \approx .118$, $X = +1$
- Lose on first, win exactly one of next two, probability $(20/38)2(18/38)(20/38) \approx .262$, $X = -1$
- Lose all three, probability $(20/38)^3 \approx .146$, $X = -3$.

So X takes value $+1$ with probability $\approx .592$, value -1 with probability $\approx .262$, and value -3 with probability $\approx .146$.

Solution

Let X be net winnings from this strategy. Possible outcomes/values for X:

- Win on first roll, probability $18/38 \approx .474$, $X = +1$
- Lose on first, win on next two, probability $(20/38)(18/38)^2 \approx .118$, $X = +1$
- Lose on first, win exactly one of next two, probability $(20/38)2(18/38)(20/38) \approx .262$, $X = -1$
- Lose all three, probability $(20/38)^3 \approx .146$, $X = -3$.

So X takes value $+1$ with probability $\approx .592$, value -1 with probability $\approx .262$, and value -3 with probability $\approx .146$.

The strategy *is* winning — you have a net gain more often than a net loss!

More examples

Example (Netty's Scam): Netty the Incredible runs the following scam in her spare time:

She has a business where she forecasts the gender of the unborn child for expectant couples, for a small price. The couple come for a visit to Netty's office and, having met them, Netty retires to her ante-room to gaze into her Crystal Ball. In reality, Netty flips a coin. If the result is "Heads" , she will predict a boy and if the result is "Tails", she will predict a girl. Netty returns to her office and tells the couple of what she saw in her crystal ball. She collects her fee of \$100 from the couple and promises to return \$150 if she was wrong.

More examples

What is the probability distribution for Netty's earnings per consultancy in this business?

More examples

What is the probability distribution for Netty's earnings per consultancy in this business?

Netty will win \$100 if she wins and lose \$50 if she loses. Let X be the random variable which is the amount Netty wins in one consultancy. Hence $\mathbf{P}(X = 100) = 0.5$ and $\mathbf{P}(X = -50) = 0.5$.

More examples

Example: Harold and Maude play a card game as follows. Harold picks a card from a standard deck of 52 cards, and Maude tries to guess its suit without looking at it. If Maude guesses correctly, Harold gives her \$3.00; otherwise, Maude gives Harold \$1.00. What is the probability distribution for Maude's earnings for this game (assuming she is not "psychic")?

More examples

Example: Harold and Maude play a card game as follows. Harold picks a card from a standard deck of 52 cards, and Maude tries to guess its suit without looking at it. If Maude guesses correctly, Harold gives her \$3.00; otherwise, Maude gives Harold \$1.00. What is the probability distribution for Maude's earnings for this game (assuming she is not "psychic")?

Let X be the random variable which is the amount Maude wins in one round. Either Maude wins \$3 or she looses \$1.

Hence $\mathbf{P}(X = 3) = \dfrac{13}{52} = \dfrac{1}{4}$ and $\mathbf{P}(X = -1) = 1 - \dfrac{1}{4} = \dfrac{3}{4}$.

More examples

Example: At a carnival game, the player plays \$1 to play and then rolls a pair of fair six-sided dice. If the sum of the numbers on the uppermost face of the dice is 9 or higher, the game attendant gives the player \$5. Otherwise, the player receives nothing from the attendant. Let X denote the earnings for the player for this game. What is the probability distribution for X?

More examples

Example: At a carnival game, the player plays \$1 to play and then rolls a pair of fair six-sided dice. If the sum of the numbers on the uppermost face of the dice is 9 or higher, the game attendant gives the player \$5. Otherwise, the player receives nothing from the attendant. Let X denote the earnings for the player for this game. What is the probability distribution for X?

The player either wins \$5 - \$1 = \$4 or looses \$1.

$\mathbf{P}(X = 4) = \frac{4+3+2+1}{36} = \frac{10}{36}$ and $\mathbf{P}(X = -1) = \frac{26}{36}$.

More examples

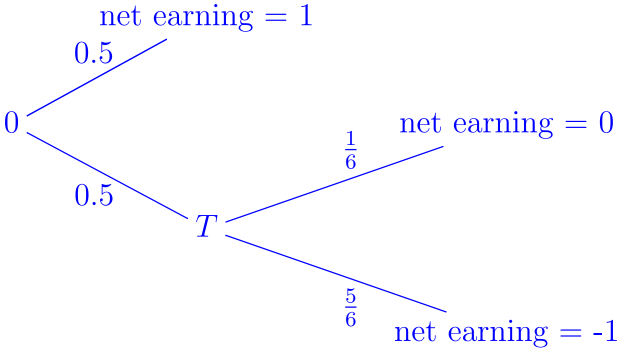
Example: The rules of a carnival game are as follows:

1. The player pays \$1 to play the game.
2. The player then flips a fair coin, if the player gets a head the game attendant gives the player \$2 and the player stops playing.
3. If the player gets a tail on the coin, the player rolls a fair six-sided die. If the player gets a six, the game attendant gives the player \$1 and the game is over.
4. If the player does not get a six on the die, the game is over and the game attendant gives nothing to the player.

Let X denote the player's (net) earnings for this game, what is the probability distribution of X?

More examples

A tree diagram could help.



$\mathbf{P}(X = 1) = 0.5$; $\mathbf{P}(X = 0) = \dfrac{1}{2} \cdot \dfrac{1}{6} = \dfrac{1}{12}$;

$\mathbf{P}(X = -1) = \dfrac{1}{2} \cdot \dfrac{5}{6} = \dfrac{5}{12}$.