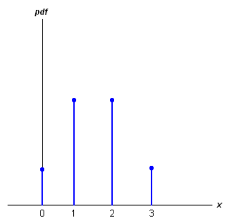
MDM4U: Mathematics of Data Management

Accounting For All Probabilities

Probability Distributions

J. Garvin



Probability Recap

Determine the probability of rolling each possible sum using two six-sided dice.

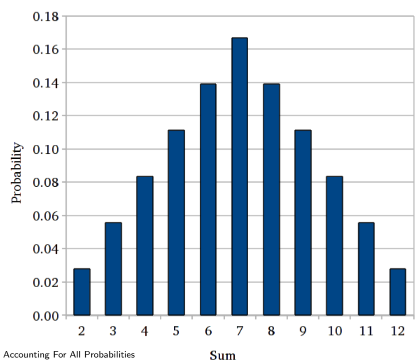
$P(2) = \frac{1}{36}$, $P(3) = \frac{2}{36} = \frac{1}{18}$, $P(4) = \frac{3}{36} = \frac{1}{12}$, $P(5) = \frac{4}{36} = \frac{1}{9}$,

$P(6) = \frac{5}{36}$, $P(7) = \frac{6}{36} = \frac{1}{6}$, $P(8) = \frac{5}{36}$, $P(9) = \frac{4}{36} = \frac{1}{9}$,

$P(10) = \frac{3}{36} = \frac{1}{12}$, $P(11) = \frac{2}{36} = \frac{1}{18}$, $P(12) = \frac{1}{36}$

Probability Recap

We can represent the probabilities for each sum graphically.



Terminology

A *random variable* is an assignment of a numerical value to a real-life occurrence (e.g. the sum of two dice). A random variable is typically denoted by X.

A random variable can take on particular values, denoted by x. Associated with these values are probabilities, $P(X = x)$ or $P(x)$ for short.

A *probability distribution* is a function of the random variable X for all acceptable values of x.

Probability distributions are often represented graphically, like the previous example.

Discrete vs. Continuous Data

Data can be *discrete* or *continuous*.

Discrete data is composed of values that are separate from each other, while continuous data is composed of an infinite number of values, within some interval.

Example

Classify as discrete or continuous data:

- number of coats on a rack (discrete)
- a car's distance from home (continuous)
- shoe sizes (discrete)

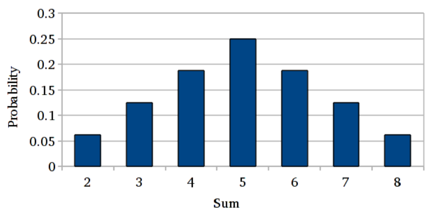
Probability Distributions

Example

Determine the probability distribution for the sum of two tetrahedral dice.

$P(2) = \frac{1}{16}$, $P(3) = \frac{1}{8}$, $P(4) = \frac{3}{16}$, $P(5) = \frac{1}{4}$, $P(6) = \frac{3}{16}$, $P(7) = \frac{1}{8}$, $P(8) = \frac{1}{16}$

Probability Distributions

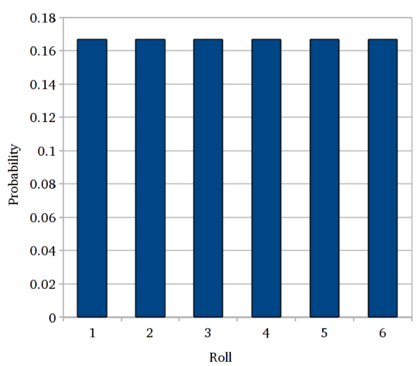


Uniform Probability Distribution

Some probability distributions are *uniform*, in that all probabilities are equally likely.

For example, consider the probability distribution for the roll of a six-sided die.

Uniform Probability Distribution



Uniform Probability Distribution

Probability in a Uniform Probability Distribution

For a uniform probability distribution, $p(x)$, with n possible outcomes, the probability of each outcome is $P(x) = \frac{1}{n}$.

This should be intuitive. Imagine a spinner divided into n equally-sized sectors. Each outcome is equally likely, and a player has a 1 in n chance of landing in any given sector.

Expected Value

In many cases, we are interested in knowing what value we can expect as an outcome.

The *expected value*, denoted $E(X)$, is the predicted average of all possible outcomes in an experiment.

Expected Value for a Discrete Probability Distribution

$E(X) = x_1P(x_1) + x_2P(x_2) + \ldots + x_nP(x_n)$
or using sigma notation...

$$E(X) = \sum_{i=1}^{n} x_iP(x_i)$$

Expected Value

Example

What is the expected sum when two dice are rolled?

Solution: Create a table of values, to calculate the expected value for each roll.

Roll	Prob.	$xP(x)$
2	$\frac{1}{36}$	$\frac{2}{36}$
3	$\frac{2}{36}$	$\frac{6}{36}$
4	$\frac{3}{36}$	$\frac{12}{36}$
5	$\frac{4}{36}$	$\frac{20}{36}$
6	$\frac{5}{36}$	$\frac{30}{36}$
7	$\frac{6}{36}$	$\frac{42}{36}$

Roll	Prob.	$xP(x)$
8	$\frac{5}{36}$	$\frac{40}{36}$
9	$\frac{4}{36}$	$\frac{36}{36}$
10	$\frac{3}{36}$	$\frac{30}{36}$
11	$\frac{2}{36}$	$\frac{22}{36}$
12	$\frac{1}{36}$	$\frac{12}{36}$

So the expected valus is $E(X) = \sum_{i=1}^{n} x_iP(x_i) = 7$.

Fairness

A *fair game* must not be biased toward a particular player.

If the expected value of a game is negative, it may represent a loss for a player, while a positive expected value may represent a win.

The expected value of a fair game is zero.

Fairness

Example

Three coins are tossed. If an even number of heads is tossed, the player wins \$5. If an odd number of heads is tossed, the player loses \$3. Is the game fair?

Create a table of values, to calculate the expected value for each sequence of tosses.

Heads	Outcomes	Probability	Payout	$xP(x)$
0	TTT	$\frac{1}{8}$	5	$\frac{5}{8}$
1	HTT, THT, TTH	$\frac{3}{8}$	-3	$-\frac{9}{8}$
2	HHT, HTH, THH	$\frac{3}{8}$	5	$\frac{15}{8}$
3	HHH	$\frac{1}{8}$	-3	$-\frac{3}{8}$

So $E(X) = \sum_{i=0}^{n} x_i P(x_i) = 1$.

Fairness

On average, a player can expect to win \$1 each round. The game is not fair, but biased toward the player.

To visualize, imagine a game where four rounds are played. Since it is equally likely to get an even number of heads as it is an odd number, both results should occur roughly half the time each.

So for a four round game, a player would be expected to win \$5 twice, and lose \$3 twice, for a net gain of \$4. This is equivalent to \$1 per round.

Questions?