Rounding in WebWork

Important:

For WebWork assignments involving calculations:

- If the number of decimals required is not specified, please include all the decimals you obtain from your calculation in the answer you submit.
- Do not round calculations in intermediate steps. Ideally, you will be using R and storing answers that you will need to reuse so that all decimals are carried to the next step.

Random Variables

Random variable

Variable (typically *X*) whose number value is determined by chance from the outcome of a random procedure

Two Types of Random Variables

Discrete Random Variables

Have numerical values that you can list or count

Examples: Number of siblings, number of contacts stored on a person's phone, number of flips until you see a Head

Continuous Random Variables

Have numerical values that can't be listed or counted because they occur over a range of values

Examples: Height of a person, weight of a bag of chips

Probability Distributions for Discrete RVs

Probability Distribution (or Probability Mass Function)

Keeps track of:

1. All the possible values of a random variable
2. The probability of each value

Example:

Roll a die. The probability distribution of the number showing on top is:

Number	1	2	3	4	5	6
Probability	1/6	1/6	1/6	1/6	1/6	1/6

Probability Distribution: Requirements

1. There is a numerical random variable *X* and its values are associated with probabilities of outcomes of a procedure.
2. The sum of all probabilities must be 1.

$$\sum P(x) = 1$$

3. Each probability value must be between 0 and 1 inclusive.

$$0 \le P(x) \le 1$$

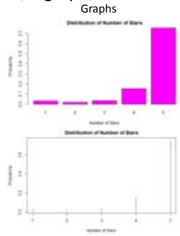
Probability Distributions for Discrete Random Variables

Probability distributions for discrete random variables can be summarized in a table, a graph, or as a formula.

Table

Number of Stars	Probability
1	0.033
2	0.020
3	0.035
4	0.156
5	0.756

Graphs



Formula

The probability of having x children until the first girl is $(1/2)^x$.

Probability distribution for the number of stars given, as a rating, to a particular book on amazon.ca by a randomly selected reviewer.

Example: Constructing a Probability Distribution

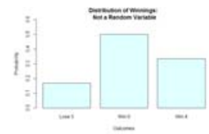
Roll a fair 6-sided die. You win \$4 if you roll a 5 or a 6. You lose \$5 if you roll a 1. For any other outcome, you win or lose nothing.

Construct a table showing the probability distribution for X = amount of money you win playing this game.

You win \$4 if you roll a 5 or 6, so the probability is 2/6.

You lose \$5 if you roll a 1, so the probability is 1/6.

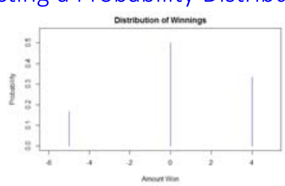
You win \$0 if you roll 2, 3, or 4, so the probability is 1/2.



This plot does not represent the distribution of a random variable, because the outcomes are listed as categories, not as their numerical values.

Example: Constructing a Probability Distribution

Winnings	Probability
−5	1/6
0	1/2
4	2/6



This plot now represents the distribution of the random variable *X,* with the possible numerical values located correctly on the horizontal axis.

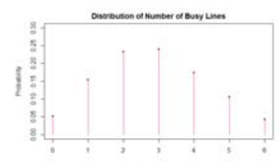
Note: All probability values are between 0 and 1, and their sum is 1.

Example

An accounting office has six incoming telephone lines. The probability distribution of the number of busy lines *X* is

x	0	1	2	3	4	5	6
P(x)	0.052	0.154	0.232	0.240	0.174	0.105	0.043

The distribution of this random variable *X* can be visualized in R as follows:



Example

An accounting office has six incoming telephone lines. The probability distribution of the number of busy lines *X* is

x	0	1	2	3	4	5	6
P(x)	0.052	0.154	0.232	0.240	0.174	0.105	0.043

Find the probability that the number of busy lines is 2 or 3.

P(2 or 3) = P(2) + P(3)

= 0.232 + 0.240 = 0.472

Find the probability that the number of busy lines is at most 2.

P(at most 2) = P(0) + P(1) + P(2)

= 0.052 + 0.154 + 0.232 = 0.438

Mean, Variance and Standard Deviation of a Probability Distribution

$$\mu = \Sigma[x \cdot P(x)]$$ Mean

$$\sigma^2 = \Sigma[(x-\mu)^2 \cdot P(x)]$$ Variance

$$\sigma^2 = \Sigma[x^2 \cdot P(x)] - \mu^2$$ Variance (shortcut)

$$\sigma = \sqrt{\sigma^2}$$ Standard Deviation

As before, round to one more decimal than the original data.

Note that theoretical distributions are populations, so we use μ and σ .

Expected Value

The expected value of a discrete random variable *x* is denoted by *E(X)*, and it represents the mean value of the outcomes. Therefore,

$$E(X) = \Sigma[x \cdot P(x)]$$

The mean can be interpreted as the average value we would get if we repeated the experiment indefinitely.

It represents the centre of mass of the distribution.

Identifying *Unusual* Results

As with samples of data, we can identify "unusual" values by determining if they lie outside these limits:

Maximum usual value = $\mu + 2\sigma$

Minimum usual value = $\mu - 2\sigma$

Example

Back to the telephone lines. The probability distribution of the number of busy lines *X* is

x	0	1	2	3	4	5	6
P(*x*)	0.052	0.154	0.232	0.240	0.174	0.105	0.043

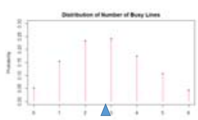
Find μ, the mean of *X*.

$$\mu = \Sigma[x \cdot P(x)]$$

x	P(x)	x·P(x)
0	0.052	0
1	0.154	0.154
2	0.232	0.464
3	0.24	0.72
4	0.174	0.696
5	0.105	0.525
6	0.043	0.258
		2.817

````
```{r}
x<- c(0:6)
problines<- c(0.052, 0.154, 0.232, 0.240,0.174, 0.105, 0.043)
meanx<- sum(x*problines)
meanx
```

[1] 2.817
````



$\mu = 2.817$ (can be rounded to 2.8)

Example

Find σ, the standard deviation of *X*.

x	0	1	2	3	4	5	6
P(*x*)	0.052	0.154	0.232	0.240	0.174	0.105	0.043

$\mu = 2.817$

$$\sigma^2 = \Sigma[(x-\mu)^2 \cdot P(x)]$$

$$\sigma = \sqrt{\sigma^2}$$

x	P(x)	(x-μ)^2	(x-μ)^2·P(x)
0	0.052	7.935489	0.412645
1	0.154	3.301489	0.508429
2	0.232	0.667489	0.154857
3	0.24	0.033489	0.008037
4	0.174	1.399489	0.243511
5	0.105	4.765489	0.500376
6	0.043	10.13149	0.435654
			2.263511

````
```{r}
varx<-sum((x-meanx)^2*problines)
varx
sqrt(varx)
```

[1] 2.263511
[1] 1.504497
````

$\sigma^2 = 2.263511$

$\sigma = \sqrt{2.263511}$
$= 1.5$

Example – continued

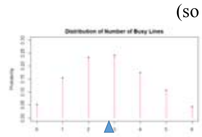
We found that the mean number of lines in use is 2.8, and the standard deviation is 1.5.

What do these values say about the maximum and minimum usual values for the number of lines in use?

$$\text{maximum usual value} = \mu + 2\sigma = 2.8 + 2(1.5) = 5.8$$

$$\text{minimum usual value} = \mu - 2\sigma = 2.8 - 2(1.5) = -0.2$$

(so minimum value is 0)



$\mu - 2\sigma$ $\quad \mu \quad$ $\mu + 2\sigma$

Note: We knew that the number of busy lines is between 0 and 6. Suppose we did not know this. Then knowing only the mean and the standard deviation (not the original distribution) gives us a lot of information about the distribution.

A very Important Discrete Distribution: The Binomial Distribution

Conditions:

1. A **fixed** number of trials (*n*).
2. Only ***two*** outcomes possible at each trial.
3. The trials are ***independent***.
4. The probability (*p*) of success is the same at each trial.

Binomial Distribution:

The binomial distribution counts the number of successes in *n* independent trials of a binomial experiment with fixed probability of success *p*.

Notation

n	denotes the fixed number of trials.
x	denotes a specific number of successes in n trials, so x can be any whole number between 0 and n, inclusive.
p	denotes the probability of success in one of the n trials.
$P(x)$	denotes the probability of getting exactly x successes among the n trials.

Caution

- Be sure that x and p both refer to the same category being called a success.
- When sampling without replacement, recall that we consider trials to be independent if the sample is very small when compared to the population.

Examples

- We will discuss here many random experiments that can be modelled by a binomial distribution.
- We will use R to obtain the answers to the questions easily.

The Language of Events Involing More Than One Value

Flip a fair coin 10 times. List the outcomes in the event that the number of heads is:

a) Exactly 7 (= 7) 0 1 2 3 4 5 6 7 8 9 10

b) At least 7 (≥ 7) 0 1 2 3 4 5 6 7 8 9 10

c) More than 7 (> 7) 0 1 2 3 4 5 6 7 8 9 10

d) Less than 7 (< 7) 0 1 2 3 4 5 6 7 8 9 10

e) At most 7 (≤ 7) 0 1 2 3 4 5 6 7 8 9 10

f) Between 5 and 7 (inclusive) ($5 \leq x \leq 7$)
0 1 2 3 4 5 6 7 8 9 10

Example 1: Guessing on a T/F Test

Suppose you flip a coin to guess the answers of a test with 10 questions, all of which are True/False questions.

You count the number of questions that you guessed correctly.

Is this is an example of a binomial experiment?

1. There is a fixed number of trials (10).
2. Each trial has only two outcomes (the answer is correct or not).
3. Trials are independent (outcome of one sample does not affect the outcome of the other because each question is answered independently).
4. The probability of success (guessing correctly) is constant (1/2).

Find the probability of passing (6 or more correct answers).

Example 2: Guessing on a Multiple Choice Test

Suppose you randomly guess the answers of a test with 10 questions, all of which are multiple choice questions with 4 possible answers.

You count the number of questions that you guessed correctly.

Is this is an example of a binomial experiment?

1. There is a fixed number of trials (10).
2. Each trial has only two outcomes (the answer is correct or not).
3. Trials are independent (outcome of one sample does not affect the outcome of the other because each question is answered independently).
4. The probability of success (guessing correctly) is constant (1/4).

Find the probability of passing (6 or more correct answers).

Sampling with Replacement

Suppose you take a sample of size *n* with replacement from a population where a proportion *p* of the population has a certain characteristic. You count the number of individuals in the sample that have the characteristic.

Is this is an example of a binomial experiment?

1. There is a fixed number of trials *(n).*
2. Each trial has only two outcomes (the individual has the characteristic or not).
3. Trials are independent (outcome of one sample does not affect the outcome of the other because sampling is with replacement).
4. The probability that each individual has the characteristic remains constant: *p* (even if it is unknown).

Remember: as long as *n* is small relative to the size of the population, samples without replacement can be considered as independent.

Example 3: High School Diplomas

According to Statistics Canada's last census data, 86.3% of Canadians between the ages of 25 and 64 had completed a high school diploma or equivalency certificate.

Suppose we randomly select 15 Canadians aged 25 to 64. Obtain the probability distribution of the number of people in the sample that have completed a high school diploma or equivalency certificate. (Assume independence: sample is small.)

Find the probability that at least 8 out of 15 randomly selected Canadians aged 25 to 64 will have a education.

Example 4: Mobile-Only Canadians

According to a study by Media Technology Monitor, 27.5% of Canadians have only cell phones. Assume that 160 people are randomly chosen and asked whether they have only a cell phone. Find the probability that:

At most 20 will have only a cell phone.

http://mediaincanada.com/2017/11/03/just-who-is-the-mobile-only-canadian/

Example 5: Coin Flips

- Flip a coin *n* times. Find the exact probability that 60% or more of your flips are Heads.

So if *n* = 10, find the probability that 6 or more are Heads (6,7, 8, 9, 10)

n = 50,	30	(30, 31, … 50)
n = 100,	60	(60, 61, … 100)

etc.

There are more and more calculations to perform as the number of flips grows.

Example 5: Coin Flips (cont.)

Abraham de Moivre (1667-1754) earned some money as a consultant to gamblers. The need to make long calculations led him to an important result (now known as DeMoivre-Laplace Theorem). The inspiration for the result will be clear when we plot the binomial distribution.

A religious man, he thought that his discovery was evidence of the existence of divine order:

> "And thus it will be found, that although Chance produces irregularities, still the odds will be infinitely great, that in process of time those irregularities will bear no proportion to the recurrency of that Order which naturally results from Original Design."

The Binomial Probability Formula

$$P(x) = \underbrace{\frac{n!}{(n-x)!x!}}_{\text{The number of outcomes with exactly } x \text{ successes among } n \text{ trials}} \cdot \underbrace{p^x \cdot q^{n-x}}_{\text{The probability of } x \text{ successes among } n \text{ trials for any one particular order}}$$

We will be performing all of our calculations using R, so you are not expected to understand or apply this formula.