2

Univariate Statistics 2: The Mean and Standard Deviation

The Mean: Language Acquisition

The Variance and Standard Deviation

Example: Fuel Efficiency

Transforming Data

Summary

Exercises

Further Reading

The four main learning outcomes of this chapter are to:

1 be able to calculate the mean of a variable
2 be able to calculate the standard deviation of variable
3 think about possible transformations and
4 remember to use the correct units

Much psychological research concerns calculating and comparing values of one statistic, the *mean.* The mean is also used and referred to in popular culture as the average. Like the median, it is a measure of the centre of a variable's distribution, but it has become more popular. Another important statistic, the standard deviation is also introduced in this chapter. It shows how spread out the data are around the mean. In later chapters we use the standard deviation to help to calculate the precision of the estimate of the mean.

THE MEAN: LANGUAGE ACQUISITION

He was born in Oklahoma in 1973, but at two weeks was uprooted from Oklahoma to live most of his early life in student digs in New York City. At nine months Nim Chimpsky began personalized tuition at Columbia University. Following the apparent acquisition of language by other non-human primates, Herbert Terrace and colleagues (Terrace, 1987; Terrace et al., 1979) tried to teach sign language to Nim under laboratory conditions. The name *Nim Chimpsky*, is with reference to the great linguist Noam Chompsky (see p. 29 of Terrace, 1987, for how they came up with the name).

Language requires using words with meaning and applying rules to combine words, and must go beyond mimicry and cannot be due to even unconscious prompting from the human observers. Terrace's research set strict standards trying to make sure that Nim truly was producing language. Thomas Sebeok, one of the best known semioticians (i.e., people who study signs and meaning), is alleged to have said: 'In my opinion, the alleged language experiments with apes divide into three groups: one, outright fraud; two, self-deception; three, those conducted by Terrace' (http://en.wikipedia.org/wiki/Nim_Chimpsky, accessed 26 June 2006).

Terrace and colleagues documented several of the two- and three-word phrases that Nim used. Table 2.1 (data from Terrace, 1987: Table 7) shows the number of times Nim signed for 18 different types of food, and how often he gave that sign either with the sign for *Nim* versus for *Me* (so, *apple* + *Nim* or *Nim* + *apple* versus *apple* + *Me* or *Me* + *apple*). Terrace called each use of a particular type of phrase a 'token', hence Table 2.1 gives the number of tokens. We are interested in whether Nim used *Nim* or *Me* more. The simplest approach to this is to count the number of *Nim* tokens and the number of *Me* tokens. This is called the *sum* of them. For the *Nim* phrases this is:

$$658 = 90 + 91 + 1 + 24 + 2 + 17 + 26 + 68 + 80 + 4 + 2 + 4 + 24 + 29 + 98 + 21 + 12 + 65$$

The variable for the number of tokens using the word *Nim* for each type of food can be denoted Nim_i, where the subscript i tells the reader that the variable Nim_i can have a different value for each food (note that *Nim* is a variable, Nim is the chimpanzee's name). So, banana is the second food. Thus, $Nim_2 = 91$ tokens. Note that it is always important to include the units that are being measured, here 'tokens' of the particular linguistic type. There is a shorthand way to write 'add up all the values' and that is using the summary sign which is the capital Greek letter sigma: Σ. Write: $\Sigma Nim_i = 658$ tokens to mean the sum of all the values of Nim_i. For the *Me* tokens the sum is: $\Sigma Me_i = 378$ tokens. So, for this set of

18 foods, Nim used his own name about twice as often as the pronoun *Me*. Comparing the sums of each of these variables is informative.

Figure 2.1 shows stem-and-leaf diagrams (see Chapter 1) presented back-to-back to allow the distributions of the two variables to be compared. It is clear that, for most of the food types, Nim uses his own name more than the pronoun *Me*, but there is one outlier for Me_i values. For some reason when referring to *banana* he tends to use his own name. In Chapter 1 we discussed using the median to describe the centre of a distribution. Because there are 18 foods, the median is halfway between the 9th and 10th values; the mid-rank is 9.5. These have been printed in **bold** in the stem-and-leaf diagram. The median for Nim_i is 24 tokens. The median for Me_i is halfway between the 9th and 10th values, halfway between 4 and 9. This is $(4 + 9)/2$ or 6.5 tokens. Because the median is higher for Nim_i, it appears Nim uses his name more than the pronoun *Me*.

An alternative to the median is the mean.[1] It is probably the most used statistic both in science and in the popular press. Like the median, it is a measure of central tendency; it provides a single number for the centre of the distribution. It is sum of all the values divided by the number of items in the sample. In symbols, we write: $\bar{x} = \Sigma x_i/n$, where n is the sample size. The notation $\bar{x}$ is pronounced 'x bar'. To denote the mean of a variable it is common practice to draw a line over the variable name, so the mean of Nim_i is M_{Nim}. There are several alternatives for denoting the mean. Another method is to denote it with the capital letter *M*, so that the mean of Nim_i would be M_{Nim}. These are the two most common ways to denote the mean of a variable. For our example:

$$\overline{Nim} = \frac{\sum Nim_i}{n} = \frac{658}{18} = 36.56 \textit{ tokens and } \overline{Me} = \frac{\sum Me_i}{n} = \frac{378}{18} = 21.00 \textit{ tokens}$$

This shows that the mean number of times Nim used his own name for those 18 foods was 36.56 times per food type compared with only 21.00 times for the pronoun *Me*. So, the mean for Nim_i is higher than the mean for Me_i.

The means for both of these are larger than their respective medians, particularly for the uses of *Me*. This is because of one very large value. For whatever reason, Nim used the pronoun *Me* 131 times to get a banana. The mean is greatly affected by extreme points, while the median is not. This is one of the reasons why people often choose the median, rather than the mean.

---

1 The word 'average' is often used in the press instead of 'mean', for things like a batting average. Statisticians often talk about different types of 'means'. The full name for the mean we describe is the 'arithmetic mean', and it is the most common 'mean' and is what people (even statisticians) are referring to when they say mean. We are often asked if the statistical word 'mean' derives from either the English word that has the synonym cruel or the word with the synonym signify. The answer is: neither. According to the *Oxford English Online Dictionary*, the word 'mean' has several meanings, including sexual intercourse (so if your instructor asks you to find a mean ...). The statistical term is more similar with the word mean as used in music for the middle note than with these other uses (see http://dictionary.oed.com/).

Table 2.1 *The number of times that Nim used his name versus the pronoun Me in response to 18 foods. The foods are in descending order of the total number of tokens used*

Food type	Nim tokens	Me tokens	Total
Banana	91	131	222
Apple	90	44	134
Gum	68	62	130
Sweet	98	31	129
Nut	80	20	100
Yogurt	65	3	68
Tea	21	30	51
Grape	26	14	40
Raisin	29	9	38
Water	12	18	30
Cracker	24	4	28
Pear	24	4	28
Fruit	17	1	18
Peach	4	1	5
Egg	2	2	4
Orange	4	0	4
Pancake	2	2	4
Berry	1	2	3

Nim tokens	Stem	Me tokens
44221	0	0112234**49**
72	10	48
96**44**1	20	0
	30	01
	40	4
85	50	
	60	2
0	70	
810	80	
	90	
	100	
	110	
	120	
	130	1

Figure 2.1 A stem-and-leaf diagram of the distribution of *Nim* tokens and *Me* tokens in response to 18 foods. The values in **bold** are the middle two values

Consider the data from Exercise 1.9 on the amount spent by 10 people at DK's Curry Palace. The values were (all in £, and £1 ≈ $2): 4.50, 13.00, 5.50, 4.50, 6.00, 5.50, 5.00, 5.00, 5.50 and 6.00. The total bill was £60.50. If they decided to split this bill equally, then they would calculate the mean: £60.50/10 people = £6.05 per person. This is different from the median (which you had to find in Exercise 1.9, so we will not give you the answer here). Steve would be paying a lot less than he should since he got £13 worth of food. We say Steve is an outlier and that he has a large positive *residual*. A residual is how far off the observed value is from the estimated value. All the other people have fairly small, but negative residuals. Louise and Susan have negative residuals of £1.50, the largest in magnitude, which means they are subsidizing Steve's gluttony the most. Residual is a concept that is brought up in other chapters, particularly on regression, where more complex methods could be used to estimate how much each person should pay.

THE VARIANCE AND STANDARD DEVIATION

The mean and median are measures of the central tendency of a distribution. In Chapter 1 you were introduced to the *range* and the *IQR* as measures of the spread of the distribution. In this chapter you are introduced to the two most popular statistics to describe the spread of a distribution. These are the *variance* and the *standard deviation*, and they are closely related. We will begin with the variance and then describe how to calculate the standard deviation from this.

When one thinks about how spread out a distribution is, one way to think of it is how far away most of the points are from the mean. If they are all close to the mean, then the distribution is not spread out and we would want a measure of spread to reflect this. If the points are far from the mean then we would want a measure of spread to be large. The variance of a variable can be calculated in four steps, providing that you have already calculated the mean.

1. Subtract the mean of the variable from *each* value, $(x_i - \bar{x})$.
2. Multiply this value by itself; in other words, square it, $(x_i - \bar{x})^2$.
3. Add these values together, $\Sigma (x_i - \bar{x})^2$.
4. Divide by the number of cases minus one, $\Sigma (x_i - \bar{x})^2/(n - 1)$.

The following shows the equation in full:

$$\text{var}\, x_i = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Table 2.2 *Calculating the residuals* $(\bar{x}_i - \bar{x})$, *the squared residuals* $(\bar{x}_i - \bar{x})^2$, *and the sum of the squared residuals* $\Sigma(\bar{x}_i - \bar{x})^2$, *which is often just called the sum of squares. To find the variance divide by n–1 (here 10 – 1 = 9), and you get 1.79*

	$(\bar{x}_i - \bar{x}) =$	$(\bar{x}_i - \bar{x})^2 =$
	(4–4.7) = – 0.7	0.49
	(3–4.7) = – 1.7	2.89
	(4–4.7) = – 0.7	0.49
	(5–4.7) = 0.3	0.09
	(3–4.7) = – 1.7	2.89
	(7–4.7) = 2.3	5.29
	(4–4.7) = – 0.7	0.49
	(6–4.7) = 1.3	1.69
	(6–4.7) = 1.3	1.69
	(5–4.7) = 0.3	0.09
Sum(Σ)	0	16.10 minutes squared

Consider the following example with just 10 data points. There is much interest in how people's ability on various physical and mental tasks declines as they get older. Here 10 young adults and 10 much older adults were asked a series of current affairs questions and the times taken to complete these were recorded. The scores, rounded to the nearest minute, were:

Younger:	4	3	4	5	3	7	4	6	6	5
Older:	5	1	9	6	4	4	10	7	4	10

Table 2.2 shows some of the calculations of the variance for the younger people. Notice that the sum of $(x_i - \bar{x})$, called the *sum of the residuals*, is equal to zero (all the negative values are counterbalanced by the positive ones). This is because of the way that the mean is calculated. Notice also that when each residual is squared it is positive. To get the variance we divide the sum of squared residuals (16.10) by the number of cases minus one (10 –1 = 9) and get 16.10/9 = 1.79.

In other sciences, much care is taken with the units of measurement. For example, physicists are very careful differentiating nine metres per second, which describes a velocity, from nine metres per second squared, which describes a *change* in velocity (an acceleration).

Social scientists sometimes are less careful about this, but it can be important. Consider the data in Table 2.2. The value for x_1 is four minutes. The residual is –0.7 minutes; adding or subtracting does not change that these values are in minutes (if it takes five minutes to preheat the oven, and 10 minutes to cook the pie, it takes 15 minutes in total). You cannot add or subtract items that do not have the same units (e.g., you cannot add five minutes to preheat the oven with two cups of sugar and get seven anything). The only time you can add together values with different units is when you can transform the units of one into the other. For example, 15 minutes to cook a cake plus 45 seconds to eat it, produces a total of: 15 minutes × 60 seconds/minute + 45 seconds = 945 seconds for the entire experience.

When a value is squared, so are the units. Items can be multiplied and divided even if they do not have the same units, as in miles travelled being divided by hours taken to achieve miles per hour. The squared residual of the first person is thus 0.49 minutes squared. Summing all these values does not change the units, so the sum of the squared residuals is 16.10 minutes squared. Dividing by $n - 1$ (9) gives a variance of 1.79 minutes squared. The variance is a kind of average squared residual. This value gives information about the overall amount of spread, or variability, in the data.

If people are thinking about the spread of response times, it is difficult to think in terms of minutes squared. A measure that is closely related to the variance is the *standard deviation*. The standard deviation, denoted *sd*, is simply the square root of the variance.[2] By taking the square root the units return to minutes.

$$sd = \sqrt{var\, x_i} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

For the above data, the variance is 1.79 minutes squared and the standard deviation is 1.34 minutes. The standard deviation is a measure of how far away observations are from the mean in the same units (here minutes) of the original variable. On average, an observation falls 1.34 minutes from the mean. As you can see by the prior sentence, the standard deviation is much easier to interpret compared with the variance that reports the spread in terms of square units.

Because standard deviations are in the same units as the original variables they can be placed onto many graphs. The most common usage, showing standard deviation bars with means, is shown in the next section. Here, consider Figure 2.2. This is a DNA histogram of Figure 1.5 but with a bar showing the mean (the dot) and bars to show how many people are within one standard deviation of the mean. For most distributions there should be at least 50% of the cases within the standard deviation bars. For a particular type of distribution, the Normal distribution (discussed more in Chapter 5), two-thirds of the cases are within this region.

---

2 The square root is the opposite of squaring. For example, if we square 3 minutes we get 9 minutes squared (3 minutes × 3 minutes = 9 minutes squared). So, the square root of 9 minutes squared is 3 minutes. Both the variance and standard deviation will be positive.

OPTIONAL BOX 2.1

When calculating the variance and the standard deviation you have to square the difference between the mean and each individual value (i.e., the residuals). The reason for squaring this difference given in many textbooks is so that all of the values are positive. As said above, simply summing the residuals would make 0 because the positive and the negative values would balance each other. However, it would be much simpler just to take the absolute value (i.e., ignore any minus signs) and sum these as a measure of spread. In fact, this procedure pre-dates squaring the residuals and is known as the 'least absolute values' approach as compared with the 'least squares' approach. The 'least squares' approach caught on about 200 years ago because, when the analyses become more complex, it is much simpler computationally. In fact, it is only with recent computer advances that the 'least absolute values' approach can be used more widely.

As only astute inquisitive students read 'optional' boxes, you will have noticed the word 'least' was added to the names of the two approaches and wondered why. Suppose you had a variable, say age_i, and you wanted to find the value X which meant that $\Sigma(age_i - X)^2$ was as small as possible. It turns out that value is the mean. The mean is the value that makes the sum of the squared deviations from it the smallest. Now, suppose instead you wanted to find the value X which minimized $\Sigma|age_i - X|$, where $|x|$ means the absolute value of x (so $|3| = |-3| = 3$). It turns out the value that minimizes the sum of the absolute values is the median. When problems get more complex in future chapters, it turns out that the method of least squares has a pretty straightforward solution, but the method of least absolute values does not.

EXAMPLE: FUEL EFFICIENCY

Let's consider one more example. Companies building and people driving fuel-efficient automobiles are vital for controlling pollution levels and delaying the more cataclysmic consequences of global warming. We looked up fuel efficiencies on several new makes of cars, mostly produced in 2007, from the site http://www.fueleconomy.gov/feg/findacar.htm. Manufacturers of a few vehicles, like the Hummer H2, are not *required* to list their fuel efficiency because they are so large they are classified as heavy-duty vehicles. General Motors does not list it on their web page because they presumably do not think the consequences of global warming are that bad (why else would they make Hummers?). Various Hummer websites give estimates of about 10mpg. To give Hummer a tiny bit of credit, they have created a run-of-the-mill SUV, the H3, which has slightly better fuel efficiency. Anyway, we sampled the automobiles being bought in two fictitious towns, *Arnieville* and *Baltimore*. *Arnieville* has an odd mixture of people, some who buy automobiles specifically because of good fuel efficiency and some who buy

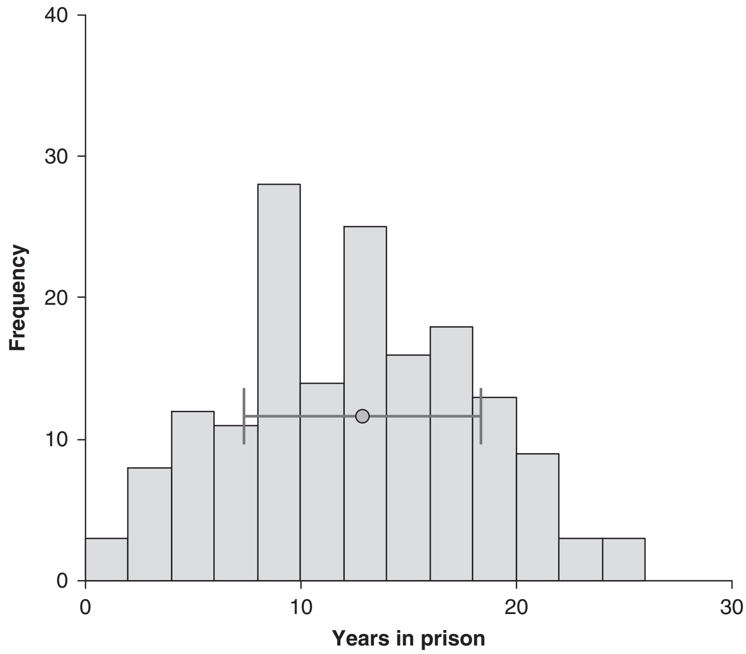

Figure 2.2 A histogram of the amount of time spent in prison before exoneration (like Figure 1.5) with the mean and standard deviation bars added. The circle is the mean and the bars go from one standard deviation below the mean to one standard deviation above the mean

automobiles that tend to have very poor fuel efficiency for unknown reasons, although Wainer's (2005: 94–5) 'penis substitution theory of automobile acquisition' (PEST) might be applicable. *Baltimore* has a mix of people, some nice, some less nice, and they buy a range of automobiles.

The mean mpg for *Arnieville* is:

$$\frac{63+39+33+31+26+21+18+15+11+10+13+13}{12} = \frac{293}{12} = 24.42\ mpg$$

and for *Baltimore* is:

$$\frac{31+28+28+26+26+24+21+18+18+16+35+21}{12} = \frac{292}{12} = 24.33\ mpg$$

So, the means in miles per gallon are approximately the same. This might give the false impression that the two towns have similar automobile buying behaviours. It appears from Table 2.3 that the values for *Arnieville* are more spread out than for *Baltimore*. The

Table 2.3 *Automobiles from two fictitious towns with their fuel efficiencies*

Arnieville automobiles	mpg	gp100m	Baltimore automobiles	mpg	gp100m
Honda Insight 2006	63	1.59	Ford Escape Hybrid	31	3.23
Toyota Camry Hybrid	39	2.56	4WD	28	3.57
Honda Fit	35	2.86	Saturn Ion	28	3.57
Hyundai Accent	33	3.03	Toyota Camry	26	3.85
Ford Focus	31	3.23	Lexus GS 450h	26	3.85
Mitsubishi Eclipse	26	3.85	Volkswagen Passat	24	4.17
Jaguar XK Convertible	21	4.76	Mercury Mariner	21	4.76
Buick Rainier 2WD	18	5.56	FWD	21	4.76
Mercedes-Benz SL65	15	6.67	Ford 500 AWD	18	5.56
AMG	13	7.69	Mazda CX-7 2WD	18	5.56
Jeep Grand Cherokee	11	9.09	BMW 550i	16	6.25
2006	10	10.00	Chevy Trailblazer	13	7.69
Lamborghini Murcielago			4WD		
Hummer H2 2004			Audi Q7		
			Jeep Grand Cherokee		
			2006		

Notes: mpg = miles per gallon; gp100m = gallons per 100 miles.

standard deviation for each town can be calculated by: (1) subtracting the mean from each value to get the residual, (2) squaring each of these residuals, (3) adding all these values together, (4) dividing by n–1 which in this case is 12-1 or 11, and finally (5) taking the square root of this value. For *Arnieville* the standard deviation is: 15.41 mpg; for *Baltimore* it is 5.74 mpg.

As shown in Figure 2.2, standard deviations can be placed onto graphs depicting means to show additional information. This is because, unlike the variance, they are in the same units as the original variable. This extra information is about the spread of the data. Figure 2.3 shows the means and the standard deviations for fuel efficiency for these two fictitious towns. This graph was done on the computer, but it can be made without a computer in six steps:

1. Draw a horizontal line and give a label for each group (here, *Arnieville* and *Baltimore*).
2. Draw a vertical line on the left of the graph. Label it and put some values on this axis. The values should go either from the minimum and maximum values attainable (here 0 mpg and there is no maximum possible) or from some value lower than the mean minus one standard deviation and the mean plus one standard deviation. As with other graphs, it is useful to use round numbers, so we go from 0 mpg to 50 mpg in 10mpg increments.
3. Draw a point where the mean is for each variable.

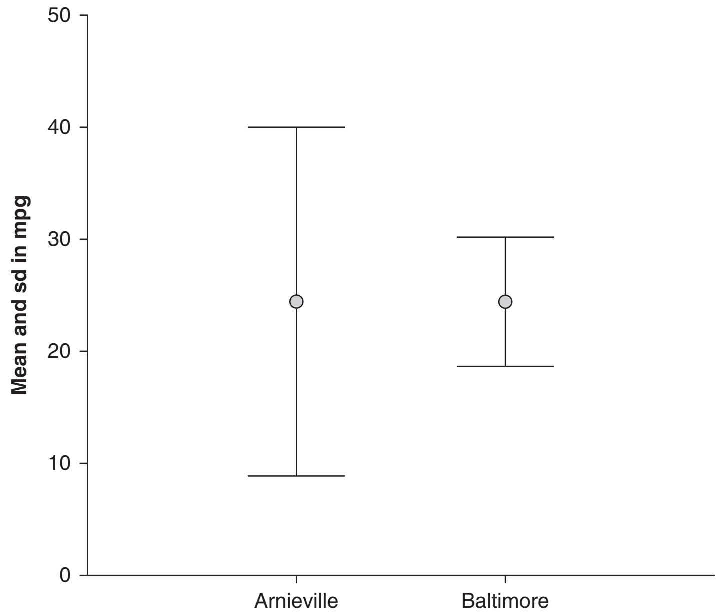

Figure 2.3 The mean fuel efficiency in miles per gallon (mpg) for automobiles from *Arnieville* and *Baltimore*. The bars show one standard deviation (sd) from the mean

4 Subtract the standard deviation of each group from that group's mean. Draw a short horizontal line at this point below the dot. This becomes the lower bound.
5 Add the appropriate standard deviation to each mean and draw a line above the point to make the upper bound.
6 Join these two bounds with a vertical line. The point for the mean should be exactly in the middle of this line.

TRANSFORMING DATA

One of the defining features of a science is that objects are measured. As such, there is much concentration on things like how to measure an attitude or intelligence, and we have stressed the importance of reporting your units. After the data have been collected it is often useful to transform the data either so it makes the data easier to understand (which often makes them more amiable to some of the assumptions of statistical tests discussed in later chapters) or so it changes the data into a form which is more useful for your purposes. The two main examples of this chapter provide examples for each of these.

The stem-and-leaf diagram of Figure 2.1 shows that distributions of tokens for Nim's use of his name and pronoun *Me* are *positively skewed.* This means that most of the values are on the lower end of the scale and there are a few values that are much higher. This is a common situation for many variables within psychology. For example, in testing how much time people take to solve a cognitive task, often a few people get very high scores, taking much longer than most to complete the task.

It is often beneficial to spread out the values at the bottom of the scale and lessen the differences at the top of the scale. Two common methods for positively skewed data like this are to take the square root of the values and to take the natural logarithm of the values. Here we will use the latter approach. The natural logarithm (often denoted *ln* or just *log*) is a mathematical function that is appropriate for values above zero.[3] Because Nim never used some tokens, we have added 0.5 to each value and then taken the natural logarithm. In equations: $\log Nim_i = \ln(Nim_i + 0.5)$. Figure 2.3 shows the effect of taking the logarithm of the values for the histogram for both these variables. Note that the scale changes. The important change is that the shape of the distribution looks more evenly spread. Tukey's (1977) *Exploratory Data Analysis* provides the most lucid rationale for using transformations, though his descriptions for making graphs and calculating some of the transformations are now dated.

It is worth looking specifically at the variable Me_i and its single extremely high value for *banana*. An important question is: what is the impact of this single value on the estimates of the mean and standard deviation? The mean with *banana* is 21.00 tokens, but this drops to 14.53 tokens when this outlier is excluded. The standard deviation is 32.47 tokens with this value, and 17.87 tokens without it. While the mean becomes about a third smaller, the standard deviation drops by about half. The fact that the standard deviation drops more than the mean has important consequences when we discuss inference in Chapter 5. For the logged variable log *Me*, the mean is 2.093 log-tokens. To return to the original units you have to back-transform this value using the inverse of the earlier transformation. Here this means exponentiating the value ($e^{2.093}$) which is 8.11 and subtracting 0.5. Thus, the estimate for the centre of the distribution using this method (which many statisticians would prefer for skewed data) is 7.61 tokens. If *banana* is excluded from these calculations the estimate drops 6.38 tokens. Thus, the impact of this outlier is much less. Throughout this book we will describe other methods that can be used to lessen the impact

3 The ln(x) is the power to which the number e has to be raised to get x. e is a special number within mathematics (another special number is π = 3.14); e is approximately 2.72. If you square (which means taking to the power of 2) this value you get: $2.72^2 = 7.40$. Thus, ln(7.40) = 2. This transformation pulls in high positive values. It does not work for negative values and is equal to negative infinity for 0, so if you have any cases with the value 0 it is common practice to add a small number, like 0.5, to each value before you log the variable. This small value is called a flattening constant.

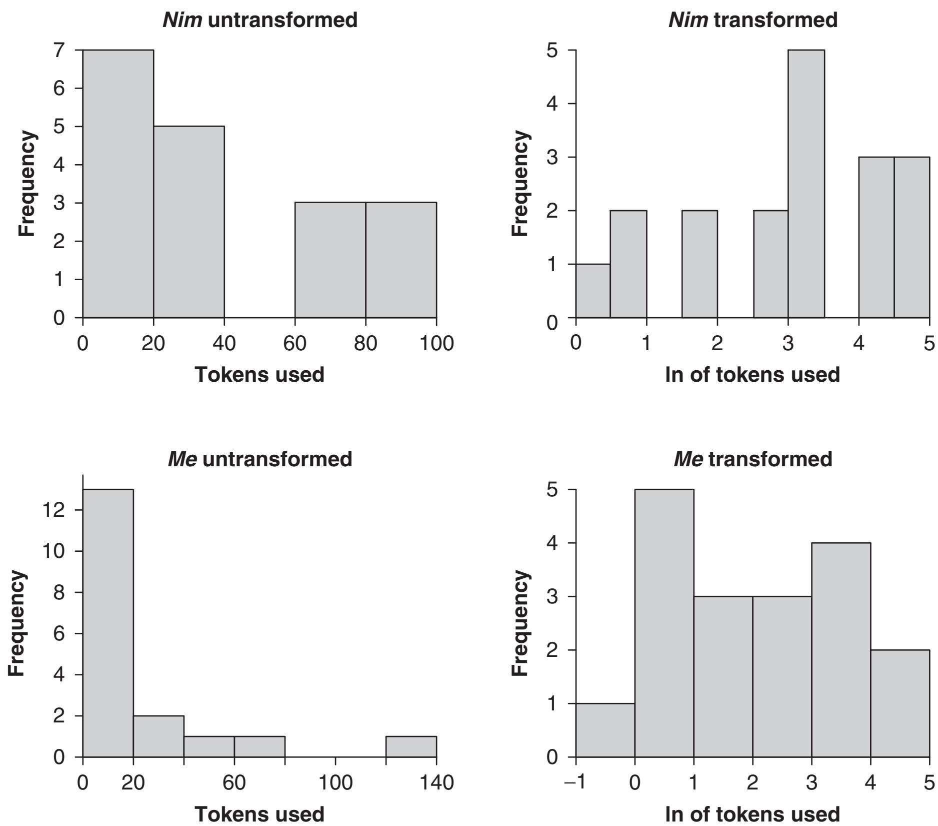

Figure 2.4 Distributions of the amount of times Nim used his own name (*Nim*) and the pronoun *Me*. The untransformed variables are shown to the left, their values after the ln(variable + 0.5) transformation are shown to the right

of outliers. It is important to realize that these methods are often still addressing the same questions as the traditional methods. They are used when traditional methods become less reliable.

For negatively skewed data, which are less common in psychology, the inverse transformations that are used for positively skewed data can be used. Thus, you can square a negatively skewed variable and this usually improves its shape. Similarly, you can exponentiate a variable. There are lots of other methods for transformations to improve the shape of the distribution, but this set is enough for most psychologists.

The second use of transformations is to change the variable into something more useful. The fuel-efficiency data were presented in miles per gallon (mpg) as is the standard in the

USA, the UK, and many other places. This would be the appropriate measurement if you had a certain amount of fuel and wanted to know how far you could travel. Think back to your childhood holidays. Did your mother say: 'We have 30 gallons, where shall we go?' Probably not. In France, fuel efficiency is measured in kilometres per litre. We will not make you completely adopt the French system, but would it make more sense to measure fuel efficiency in gallons per mile? This would equate with your mother saying: 'We are going to Lake Kita which is 100 miles away; how much will the fuel cost?' This is a more useful way to think about fuel efficiency. Table 2.3 shows the number of gallons it takes each of the automobiles to go 100 miles. The question is whether it makes a difference if we look at the data in this way.

Mean *Arnieville* = 5.48 gallons sd *Arnieville* = 2.74 gallons
Mean *Baltimore* = 4.33 gallons sd *Baltimore* = 1.05 gallons

If we assume that everybody has about 100 miles to drive each week, the average *Arnieville* resident uses a gallon more fuel than the average *Baltimore* resident. While there appeared no difference in fuel efficiency between the two fictitious towns using *mpg*, using gallons per 100 miles makes *Baltimore* appear the more efficient town. With this example, the gallons per 100 miles conclusion is more appropriate (Hand, 2004: 45).

SUMMARY

The mean and the standard deviation are two of the most important statistics for psychologists. Whenever you measure any variable you should be interested in where the centre of the distribution is and how spread out the distribution is. While the median and IQR, discussed in Chapter 1, are good measures of this, the mean and standard deviation are much more popular. They are reported in nearly every psychology article. For computational reasons they became the main statistics in the beginning of the twentieth century and many of the statistics designed for more complex problems with multiple variables can be viewed as extensions of these. Therefore, it is vital that you understand them and all introductory statistics books stress them.

The final two points are less often stressed. Many introductory statistics books wait until later chapters to talk about transformations or leave discussion of them out altogether. We decided to bring this topic up in Chapter 2 because transformations are an important tool for statistics and they help to introduce the idea of measurement. As far as units, we are surprised that psychologists seems less concerned about units than other scientists, particularly as many of the scientific concepts in which we are interested present many measurement issues. If you are doing physics, for example, the importance of reporting the correct units will likely be stressed from your first lecture (Tipler & Mosca, 2007: Ch. 1). Reporting the correct units keeps the reader aware of the meaning of any numbers that you present. Omitting the units will leave the numbers as dry uninformative numbers, which is why perhaps statistics courses often seem detached from the exciting discipline of what James (1890) described as 'the science of mental life, both of its phenomena and of their conditions'.

EXERCISES

2.1 Why is it important to report the units of measurement?

2.2 Can two values be added together if they are in different units? Give an example (not from the text) where they can, and an example (not from the text) where they cannot.

2.3 Can two values be added/multiplied together if they are in different units? Illustrate your answer with an example (not from the text).

2.4 Find an example in a recent newspaper of a mean people reported. Try not to use the sports section.

2.5 The median is to the IQR as the mean is to the _________ . Which statistic from this chapter fits best here and why? If you think two fit equally well, think why one may be better.

2.6 The DNA exoneration example from Chapter 1 showed 163 people who had been convicted of, and imprisoned for, horrendous crimes they did not commit. In total, they spent 2095 years in prison. What was the mean time each of these innocent people spent in prison?

2.7 Using the response time data presented earlier in this chapter, Calculate the variance for the older group's response times. How does this compare with the variance of the younger group (see Table 2.2)? Which set of scores is more spread out? Why might this be the case?

2.8 Reeve and Aggleton (1998) were interested in people's memories for *The Archers*, a popular UK radio soap opera. Participants ($n = 48$, 12 in each of four conditions) were given one of two made-up scripts. One depicted a typical day, a visit to a livestock market. The other was of an atypical event (for the Archers), a visit to a boat show. Participants were either Archer novices or experts and were asked 22 questions about the script they were given. The

(Continued)

(Continued)

means and standard deviations for the experts were 15.08 (*sd* = 4.17) for the livestock market, 8.42 (*sd* = 2.11) for the boat show, and were 9.50 (*sd* = 4.48) and 8.42 (*sd* = 3.23), respectively, for the novices. Make an appropriate graph to display these results and comment on what they show.

2.9 Here are rainfall statistics for two cities:

	Sun	Mon	Tue	Wed	Thur	Fri	Sat
Toledo	0	0	1	6	0	0	0
Brighton	1	1	1	1	1	1	1

What are the means, the standard deviations, and variances of the rainfall for each of these two places? What do these tell you about the climates?

2.10 The noon temperatures in Toledo and Brighton for one week are listed below. Because the Toledo we refer to is the one in the US, and the Brighton is the one in the UK, they use degrees Fahrenheit and Celsius respectively.

	Sun	Mon	Tue	Wed	Thur	Fri	Sat	
Toledo	43	40	38	38	42	59	65	in °F
Brighton	8	10	10	14	8	9	11	in °C

Which place has the higher mean temperature? The following may be useful:

°C = 5/9 (°F – 32)
°F = 9/5 °C + 32

2.11 Schkade et al. (2000) collected data from jurors on 15 different cases. They ask for ratings of severity of the crime, on a 0 (no punishment) to 8 (extremely severe punishment) before jury deliberation and after jury deliberation. The values for each case are as given in Table 2.4.

Find the standard deviation and variances of these values. Speculate why any difference in standard deviation between before and after deliberation ratings might have occurred.

2.12 At the beginning of this book it was stated that 'it takes *about* 20 minutes to cook rice' is a statistical phrase. Suppose that you were hired to come up with the time to write on the side of the containers for spicy cheese bread and a vegetable stir-fry. With various different cooking arrangements in different climates and at different altitudes, you cooked these products to perfection. Here are the times, in minutes:

Spicy cheese bread:	8	15	12	16	11	8	14	9	12	15
Veggie stir-fry:	11	14	10	12	12	13	11	13	13	11

Which dish would you give the longer time to, and why? Think about the purpose of the estimate before giving your answer. Think about this both as a statistician and as a cook. The correct answer is not that they should have the same cooking time printed on the side of the box.

Table 2.4 *Severity ratings for 15 cases*

Case	**Before severity**	**After severity**
Reynolds	5.5	.0
Glover	5.0	5.0
Lawson	4.3	4.5
Williams	5.0	5.0
Smith	5.5	6.0
Nelson	5.0	5.0
Hughes	5.0	5.0
West	4.5	5.0
Douglas	4.0	4.0
Crandall	4.0	4.0
Sanders	3.5	3.0
Windsor	3.0	2.0
Stanley	1.0	1.5
Dulworth	0.3	0.0
Newton	0.0	0.0
Means	3.71	3.73

FURTHER READING

The graphing references from the last chapter are all applicable here.

Douglass Altman and Martin Bland have written many clear and focused (i.e., short) articles to help medical researchers understand statistics and these have appeared over the last decade in the *British Medical Journal* (your university computer may have access on www.bmj.com, but are also on Professor Bland's webpage, http://www-sers.york.ac.uk/~mb55/pubs/pbstnote.htm). For example:

Bland, J.M., & Altman, D.G. (1996). Transformations, means, and confidence intervals. *British Medical Journal*, *312*, 1079.

Measurement is a difficult topic and one which is related to statistics. To be a scientist you have to have some measurement theory about your objects of investigation. We assume that most of you have either taken a methodology course where some of these issues are covered or are concurrently taking one. Your textbook for this course will provide an introduction to measurement. Therefore we suggest an excellent book on measurement, but one aimed at students who already have some grasp of the basics.

Hand, D.J. (2004). *Measurement Theory and Practice: The World Through Quantification.* London: Arnold.