Key Concepts in Descriptive Statistics

James H. Steiger

Department of Psychology and Human Development
Vanderbilt University

P310, 2011

Psychology 310 — Course Goals and Stratey

Introduction

- In statistics, we use numbers to represent quantities.
- Often we approach this process informally, without thinking about it too deeply.
- In this module, we want to take another look at some fundamental questions:
  - What's in a list of numbers?
  - What happens to the information in a list of numbers when we transform a list?
  - Is some information especially vulnerable?
  - Is some information especially robust?
  - Is some information arbitrary?

Introduction

- In statistics, we use numbers to represent quantities.
- Often we approach this process informally, without thinking about it too deeply.
- In this module, we want to take another look at some fundamental questions:
  - What's in a list of numbers?
  - What happens to the information in a list of numbers when we transform a list?
  - Is some information especially vulnerable?
  - Is some information especially robust?
  - Is some information arbitrary?

Introduction

- In statistics, we use numbers to represent quantities.
- Often we approach this process informally, without thinking about it too deeply.
- In this module, we want to take another look at some fundamental questions:
  - What's in a list of numbers?
  - What happens to the information in a list of numbers when we transform a list?
  - Is some information especially vulnerable?
  - Is some information especially robust?
  - Is some information arbitrary?

Introduction

- In statistics, we use numbers to represent quantities.
- Often we approach this process informally, without thinking about it too deeply.
- In this module, we want to take another look at some fundamental questions:
  - What's in a list of numbers?
  - What happens to the information in a list of numbers when we transform a list?
  - Is some information especially vulnerable?
  - Is some information especially robust?
  - Is some information arbitrary?

Introduction

- In statistics, we use numbers to represent quantities.
- Often we approach this process informally, without thinking about it too deeply.
- In this module, we want to take another look at some fundamental questions:
  - What's in a list of numbers?
  - What happens to the information in a list of numbers when we transform a list?
  - Is some information especially vulnerable?
  - Is some information especially robust?
  - Is some information arbitrary?

Introduction

- In statistics, we use numbers to represent quantities.
- Often we approach this process informally, without thinking about it too deeply.
- In this module, we want to take another look at some fundamental questions:
  - What's in a list of numbers?
  - What happens to the information in a list of numbers when we transform a list?
  - Is some information especially vulnerable?
  - Is some information especially robust?
  - Is some information arbitrary?

Introduction

- In statistics, we use numbers to represent quantities.
- Often we approach this process informally, without thinking about it too deeply.
- In this module, we want to take another look at some fundamental questions:
  - What's in a list of numbers?
  - What happens to the information in a list of numbers when we transform a list?
  - Is some information especially vulnerable?
  - Is some information especially robust?
  - Is some information arbitrary?

Introduction

- In statistics, we use numbers to represent quantities.
- Often we approach this process informally, without thinking about it too deeply.
- In this module, we want to take another look at some fundamental questions:
  - What's in a list of numbers?
  - What happens to the information in a list of numbers when we transform a list?
  - Is some information especially vulnerable?
  - Is some information especially robust?
  - Is some information arbitrary?

Introduction

- At the outset, we shall take a simple, visual approach to addressing these questions.
- However, we shall discover that this “simplicity” allows us to see the concepts underlying some familiar statistical formulas.
- This discovery is typical of much of statistics: complex-looking formulas can mask some powerful yet simple underlying concepts.

Introduction

- At the outset, we shall take a simple, visual approach to addressing these questions.
- However, we shall discover that this "simplicity" allows us to see the concepts underlying some familiar statistical formulas.
- This discovery is typical of much of statistics: complex-looking formulas can mask some powerful yet simple underlying concepts.

Introduction

- At the outset, we shall take a simple, visual approach to addressing these questions.
- However, we shall discover that this "simplicity" allows us to see the concepts underlying some familiar statistical formulas.
- This discovery is typical of much of statistics: complex-looking formulas can mask some powerful yet simple underlying concepts.

The Number Line Diagram

- This is a simple device for visual display of one or more lists of numbers.
- The numbers are listed from left to right, with spacing appropriate for a linear scale.
- Using the number line diagram, we can see things that may not otherwise be obvious.
- Here is a diagram of the list of numbers 1,2,4.

1 2   4

- Here is a diagram of two lists.

1 2   4
1 2   6

The Number Line Diagram

- This is a simple device for visual display of one or more lists of numbers.
- The numbers are listed from left to right, with spacing appropriate for a linear scale.
- Using the number line diagram, we can see things that may not otherwise be obvious.
- Here is a diagram of the list of numbers 1,2,4.

1 2   4

- Here is a diagram of two lists.

1 2   4
1 2   6

The Number Line Diagram

- This is a simple device for visual display of one or more lists of numbers.
- The numbers are listed from left to right, with spacing appropriate for a linear scale.
- Using the number line diagram, we can see things that may not otherwise be obvious.
- Here is a diagram of the list of numbers 1,2,4.

1 2 4

- Here is a diagram of two lists.

1 2 4
1 2 6

The Number Line Diagram

- This is a simple device for visual display of one or more lists of numbers.
- The numbers are listed from left to right, with spacing appropriate for a linear scale.
- Using the number line diagram, we can see things that may not otherwise be obvious.
- Here is a diagram of the list of numbers 1,2,4.

```
1 2   4
```

- Here is a diagram of two lists.

```
1 2   4
1 2       6
```

The Number Line Diagram

- This is a simple device for visual display of one or more lists of numbers.
- The numbers are listed from left to right, with spacing appropriate for a linear scale.
- Using the number line diagram, we can see things that may not otherwise be obvious.
- Here is a diagram of the list of numbers 1,2,4.

```
1 2   4
```

- Here is a diagram of two lists.

```
1 2   4
1 2       6
```

Listwise Operations

- We are going to ask a fundamental question: What happens to a list of numbers when we apply the same transformation to every number in the list?
- For example, what happens to a list of numbers if we add 2 to every number in the list?
- Such an operation, applied to every number in the list, is called a *listwise operation*.
- Often, we can use a simple equation to indicate a listwise operation.
- For example,

$$Y = X + 2 \tag{1}$$

means "add 2 to every number in the X list to create a new list, called Y.

Listwise Operations

- We are going to ask a fundamental question: What happens to a list of numbers when we apply the same transformation to every number in the list?
- For example, what happens to a list of numbers if we add 2 to every number in the list?
- Such an operation, applied to every number in the list, is called a *listwise operation*.
- Often, we can use a simple equation to indicate a listwise operation.
- For example,

$$Y = X + 2 \tag{1}$$

means "add 2 to every number in the X list to create a new list, called Y."

Listwise Operations

- We are going to ask a fundamental question: What happens to a list of numbers when we apply the same transformation to every number in the list?
- For example, what happens to a list of numbers if we add 2 to every number in the list?
- Such an operation, applied to every number in the list, is called a *listwise operation*.
- Often, we can use a simple equation to indicate a listwise operation.
- For example,

$$Y = X + 2 \tag{1}$$

means "add 2 to every number in the X list to create a new list, called Y.

Listwise Operations

- We are going to ask a fundamental question: What happens to a list of numbers when we apply the same transformation to every number in the list?
- For example, what happens to a list of numbers if we add 2 to every number in the list?
- Such an operation, applied to every number in the list, is called a *listwise operation*.
- Often, we can use a simple equation to indicate a listwise operation.
- For example,

$$Y = X + 2 \tag{1}$$

means "add 2 to every number in the X list to create a new list, called Y.

Listwise Operations

- We are going to ask a fundamental question: What happens to a list of numbers when we apply the same transformation to every number in the list?
- For example, what happens to a list of numbers if we add 2 to every number in the list?
- Such an operation, applied to every number in the list, is called a *listwise operation*.
- Often, we can use a simple equation to indicate a listwise operation.
- For example,

$$Y = X + 2 \tag{1}$$

means "add 2 to every number in the X list to create a new list, called Y.

Effect of Addition

- If we add a constant to every number in a list, what is the effect on the list?
- Visualization can help us organize our thinking.
- Let's take the simple list $1, 2, 4$ and add 2 to every number in the list. Let's look at the number line diagrams.

1 2 4

3 4 6

- Can you describe what you see?
- What changed about the numbers?
- What did not change?
- What about subtracting a number from every number in a list?

Effect of Addition

- If we add a constant to every number in a list, what is the effect on the list?
- Visualization can help us organize our thinking.
- Let's take the simple list $1, 2, 4$ and add 2 to every number in the list. Let's look at the number line diagrams.

1 2   4

  3 4   6

- Can you describe what you see?
- What changed about the numbers?
- What did not change?
- What about subtracting a number from every number in a list?

Effect of Addition

- If we add a constant to every number in a list, what is the effect on the list?
- Visualization can help us organize our thinking.
- Let's take the simple list $1, 2, 4$ and add 2 to every number in the list. Let's look at the number line diagrams.

```
1  2     4
      3  4     6
```

- Can you describe what you see?
- What changed about the numbers?
- What did not change?
- What about subtracting a number from every number in a list?

Effect of Addition

- If we add a constant to every number in a list, what is the effect on the list?
- Visualization can help us organize our thinking.
- Let's take the simple list $1, 2, 4$ and add 2 to every number in the list. Let's look at the number line diagrams.

```
1  2     4
      3  4     6
```

- Can you describe what you see?
- What changed about the numbers?
- What did not change?
- What about subtracting a number from every number in a list?

Effect of Addition

- If we add a constant to every number in a list, what is the effect on the list?
- Visualization can help us organize our thinking.
- Let's take the simple list $1, 2, 4$ and add 2 to every number in the list. Let's look at the number line diagrams.

```
1  2     4
      3  4     6
```

- Can you describe what you see?
- What changed about the numbers?
- What did not change?
- What about subtracting a number from every number in a list?

Effect of Addition

- If we add a constant to every number in a list, what is the effect on the list?
- Visualization can help us organize our thinking.
- Let's take the simple list $1, 2, 4$ and add 2 to every number in the list. Let's look at the number line diagrams.

```
1  2     4
      3  4     6
```

- Can you describe what you see?
- What changed about the numbers?
- What did not change?
- What about subtracting a number from every number in a list?

Effect of Addition

- If we add a constant to every number in a list, what is the effect on the list?
- Visualization can help us organize our thinking.
- Let's take the simple list $1, 2, 4$ and add 2 to every number in the list. Let's look at the number line diagrams.

```
1  2     4
      3  4     6
```

- Can you describe what you see?
- What changed about the numbers?
- What did not change?
- What about subtracting a number from every number in a list?

Listwise Multiplication (or Division)

- Let's take the same original list, $1, 2, 4$, and now multiply all the numbers by a constant.
- For now, we will restrict ourselves to positive multipliers.
- Suppose the listwise transformation formula is

$$Y = 2X \tag{2}$$

1 2 4
2 4 8

Listwise Multiplication (or Division)

- Let's take the same original list, $1, 2, 4$, and now multiply all the numbers by a constant.
- For now, we will restrict ourselves to positive multipliers.
- Suppose the listwise transformation formula is

$$Y = 2X \tag{2}$$

1 2 4

2 4 8

Listwise Multiplication (or Division)

- Let's take the same original list, $1, 2, 4$, and now multiply all the numbers by a constant.
- For now, we will restrict ourselves to positive multipliers.
- Suppose the listwise transformation formula is

$$Y = 2X \tag{2}$$

```
1  2    4
   2    4           8
```

Listwise Multiplication (or Division)

- What changed?
- What remained the same?

Listwise Multiplication (or Division)

- What changed?
- What remained the same?

An Aside

- The question of *what does not change* during statistical operations occurs during the investigation of what we call *invariance properties.*
- Asking questions about invariance can sometimes produce profound insights.
- For example, Einstein mused that "Relativity Theory" might better have been called "Invariance Theory," since fundamentally, it dealt with what remained invariant in the space-time continuum.

An Aside

- The question of *what does not change* during statistical operations occurs during the investigation of what we call *invariance properties.*
- Asking questions about invariance can sometimes produce profound insights.
- For example, Einstein mused that "Relativity Theory" might better have been called "Invariance Theory," since fundamentally, it dealt with what remained invariant in the space-time continuum.

An Aside

- The question of *what does not change* during statistical operations occurs during the investigation of what we call *invariance properties.*
- Asking questions about invariance can sometimes produce profound insights.
- For example, Einstein mused that "Relativity Theory" might better have been called "Invariance Theory," since fundamentally, it dealt with what remained invariant in the space-time continuum.

Summary

- There are many ways to state what we have observed informally here.
- One way to summarize is to say that:
  - Listwise addition or subtraction moves the numbers as a group, as though they were mounted on a rigid stick, and slid to the left or right.
  - Listwise addition does not change any of the distances between numbers.
  - Listwise multiplication or division by a positive number can move the numbers as a group, but also causes them to "fan in" or "fan out."

Summary

- There are many ways to state what we have observed informally here.
- One way to summarize is to say that:
  - Listwise addition or subtraction moves the numbers as a group, as though they were mounted on a rigid stick, and slid to the left or right.
  - Listwise addition does not change any of the distances between numbers.
  - Listwise multiplication or division *by a positive number* can move the numbers as a group, but also causes them to "fan in" or "fan out."

Summary

- There are many ways to state what we have observed informally here.
- One way to summarize is to say that:
  - Listwise addition or subtraction moves the numbers as a group, as though they were mounted on a rigid stick, and slid to the left or right.
  - Listwise addition does not change any of the distances between numbers.
  - Listwise multiplication or division *by a positive number* can move the numbers as a group, but also causes them to "fan in" or "fan out."

Summary

- There are many ways to state what we have observed informally here.
- One way to summarize is to say that:
  - Listwise addition or subtraction moves the numbers as a group, as though they were mounted on a rigid stick, and slid to the left or right.
  - Listwise addition does not change any of the distances between numbers.
  - Listwise multiplication or division *by a positive number* can move the numbers as a group, but also causes them to "fan in" or "fan out."

Summary

- There are many ways to state what we have observed informally here.
- One way to summarize is to say that:
  - Listwise addition or subtraction moves the numbers as a group, as though they were mounted on a rigid stick, and slid to the left or right.
  - Listwise addition does not change any of the distances between numbers.
  - Listwise multiplication or division *by a positive number* can move the numbers as a group, but also causes them to "fan in" or "fan out."

Introduction

- In this section, we discover that all the information in a list of N numbers can be re-expressed in terms of N new numbers.
- These new numbers contain all the information in the original list, and the original list can be reconstructed perfectly from these new numbers.
- However, by recasting the information in this new form, we can get a better "handle" on what information is really contained in a list of numbers, and what the invariance properties are.

Introduction

- In this section, we discover that all the information in a list of N numbers can be re-expressed in terms of N new numbers.
- These new numbers contain all the information in the original list, and the original list can be reconstructed perfectly from these new numbers.
- However, by recasting the information in this new form, we can get a better "handle" on what information is really contained in a list of numbers, and what the invariance properties are.

Introduction

- In this section, we discover that all the information in a list of N numbers can be re-expressed in terms of N new numbers.
- These new numbers contain all the information in the original list, and the original list can be reconstructed perfectly from these new numbers.
- However, by recasting the information in this new form, we can get a better "handle" on what information is really contained in a list of numbers, and what the invariance properties are.

Introduction

- Our N new numbers will be the following:
  - 1 measure of *Location* (or Central Tendency)
  - 1 measure of *Spread* (or Variability)
  - $N-2$ measures of *Shape*
- We shall use simple measures of Location, Spread, and Shape in deriving a number of important principles.
- But it turns out that these principles hold for *any reasonable measures* of these three quantities.
- These principles will allow us to say all kinds of interesting things about statistics while doing virtually no mathematics!

Introduction

- Our N new numbers will be the following:
  - 1 measure of *Location* (or Central Tendency)
  - 1 measure of *Spread* (or Variability)
  - $N - 2$ measures of *Shape*
- We shall use simple measures of Location, Spread, and Shape in deriving a number of important principles.
- But it turns out that these principles hold for *any reasonable measures* of these three quantities.
- These principles will allow us to say all kinds of interesting things about statistics while doing virtually no mathematics!

Introduction

- Our N new numbers will be the following:
  - 1 measure of *Location* (or Central Tendency)
  - 1 measure of *Spread* (or Variability)
  - $N - 2$ measures of *Shape*
- We shall use simple measures of Location, Spread, and Shape in deriving a number of important principles.
- But it turns out that these principles hold for *any reasonable measures* of these three quantities.
- These principles will allow us to say all kinds of interesting things about statistics while doing virtually no mathematics!

Introduction

- Our N new numbers will be the following:
  - 1 measure of *Location* (or Central Tendency)
  - 1 measure of *Spread* (or Variability)
  - $N - 2$ measures of *Shape*
- We shall use simple measures of Location, Spread, and Shape in deriving a number of important principles.
- But it turns out that these principles hold for *any reasonable measures* of these three quantities.
- These principles will allow us to say all kinds of interesting things about statistics while doing virtually no mathematics!

Introduction

- Our N new numbers will be the following:
  - 1 measure of *Location* (or Central Tendency)
  - 1 measure of *Spread* (or Variability)
  - $N-2$ measures of *Shape*
- We shall use simple measures of Location, Spread, and Shape in deriving a number of important principles.
- But it turns out that these principles hold for *any reasonable measures* of these three quantities.
- These principles will allow us to say all kinds of interesting things about statistics while doing virtually no mathematics!

Introduction

- Our N new numbers will be the following:
  - 1 measure of *Location* (or Central Tendency)
  - 1 measure of *Spread* (or Variability)
  - $N - 2$ measures of *Shape*
- We shall use simple measures of Location, Spread, and Shape in deriving a number of important principles.
- But it turns out that these principles hold for *any reasonable measures* of these three quantities.
- These principles will allow us to say all kinds of interesting things about statistics while doing virtually no mathematics!

Introduction

- Our N new numbers will be the following:
  - 1 measure of *Location* (or Central Tendency)
  - 1 measure of *Spread* (or Variability)
  - $N - 2$ measures of *Shape*
- We shall use simple measures of Location, Spread, and Shape in deriving a number of important principles.
- But it turns out that these principles hold for *any reasonable measures* of these three quantities.
- These principles will allow us to say all kinds of interesting things about statistics while doing virtually no mathematics!

Location

- A measure of *location* or *central tendency* answers questions like the following:
  - In what general region is the list located on the number line?
  - What number is typical of the entire list?
  - What number is in the center of the list?
- Later on, we'll get more formal about measures of location.
- For now, we'll adopt a really simple measure of location — the middle value in the list. We'll call it M.
- If the number of numbers (N) is odd, the median is simply the middle value.
- If the number of numbers is even, we will define the median as the average of the two middle values, i.e., a point halfway between the two middle values on the number line.

Location

- A measure of *location* or *central tendency* answers questions like the following:
  - In what general region is the list located on the number line?
  - What number is typical of the entire list?
  - What number is in the center of the list?
- Later on, we'll get more formal about measures of location.
- For now, we'll adopt a really simple measure of location — the middle value in the list. We'll call it M.
- If the number of numbers (N) is odd, the median is simply the middle value.
- If the number of numbers is even, we will define the median as the average of the two middle values, i.e., a point halfway between the two middle values on the number line.

Location

- A measure of *location* or *central tendency* answers questions like the following:
  - In what general region is the list located on the number line?
  - What number is typical of the entire list?
  - What number is in the center of the list?
- Later on, we'll get more formal about measures of location.
- For now, we'll adopt a really simple measure of location — the middle value in the list. We'll call it M.
- If the number of numbers (N) is odd, the median is simply the middle value.
- If the number of numbers is even, we will define the median as the average of the two middle values, i.e., a point halfway between the two middle values on the number line.

Location

- A measure of *location* or *central tendency* answers questions like the following:
  - In what general region is the list located on the number line?
  - What number is typical of the entire list?
  - What number is in the center of the list?
- Later on, we'll get more formal about measures of location.
- For now, we'll adopt a really simple measure of location — the middle value in the list. We'll call it M.
- If the number of numbers (N) is odd, the median is simply the middle value.
- If the number of numbers is even, we will define the median as the average of the two middle values, i.e., a point halfway between the two middle values on the number line.

Location

- A measure of *location* or *central tendency* answers questions like the following:
    - In what general region is the list located on the number line?
    - What number is typical of the entire list?
    - What number is in the center of the list?
- Later on, we'll get more formal about measures of location.
- For now, we'll adopt a really simple measure of location — the middle value in the list. We'll call it M.
- If the number of numbers (N) is odd, the median is simply the middle value.
- If the number of numbers is even, we will define the median as the average of the two middle values, i.e., a point halfway between the two middle values on the number line.

Location

- A measure of *location* or *central tendency* answers questions like the following:
  - In what general region is the list located on the number line?
  - What number is typical of the entire list?
  - What number is in the center of the list?
- Later on, we'll get more formal about measures of location.
- For now, we'll adopt a really simple measure of location — the middle value in the list. We'll call it M.
- If the number of numbers (N) is odd, the median is simply the middle value.
- If the number of numbers is even, we will define the median as the average of the two middle values, i.e., a point halfway between the two middle values on the number line.

Location

- A measure of *location* or *central tendency* answers questions like the following:
  - In what general region is the list located on the number line?
  - What number is typical of the entire list?
  - What number is in the center of the list?
- Later on, we'll get more formal about measures of location.
- For now, we'll adopt a really simple measure of location — the middle value in the list. We'll call it M.
- If the number of numbers (N) is odd, the median is simply the middle value.
- If the number of numbers is even, we will define the median as the average of the two middle values, i.e., a point halfway between the two middle values on the number line.

Location

- A measure of *location* or *central tendency* answers questions like the following:
  - In what general region is the list located on the number line?
  - What number is typical of the entire list?
  - What number is in the center of the list?
- Later on, we'll get more formal about measures of location.
- For now, we'll adopt a really simple measure of location — the middle value in the list. We'll call it M.
- If the number of numbers (N) is odd, the median is simply the middle value.
- If the number of numbers is even, we will define the median as the average of the two middle values, i.e., a point halfway between the two middle values on the number line.

Spread

- Measures of *spread*, or *variability*, assess how far the list is spread out over the number line.
- In our discussions, we will use a very simple measure of spread, the range, which is the difference between the highest and lowest number in the list.
- We will use the letter S (for spread) to stand for the range.

Spread

- Measures of *spread*, or *variability*, assess how far the list is spread out over the number line.
- In our discussions, we will use a very simple measure of spread, the range, which is the difference between the highest and lowest number in the list.
- We will use the letter S (for spread) to stand for the range.

Spread

- Measures of *spread*, or *variability*, assess how far the list is spread out over the number line.
- In our discussions, we will use a very simple measure of spread, the range, which is the difference between the highest and lowest number in the list.
- We will use the letter S (for spread) to stand for the range.

Shape

- Shape of a list of numbers is the pattern of relative interval sizes, moving from left to right.
- Consider the list $20, 30, 40, 60, 65$
- We can compute the *unscaled* distances between the numbers as 10,10,20,5.
- The relative distances are obtained by dividing all the unscaled distances by the first nonzero value.
- The resulting *Shape parameters* are $1, 1, 2, 0.5$

Shape

- Shape of a list of numbers is the pattern of relative interval sizes, moving from left to right.
- Consider the list $20, 30, 40, 60, 65$
- We can compute the *unscaled* distances between the numbers as 10,10,20,5.
- The relative distances are obtained by dividing all the unscaled distances by the first nonzero value.
- The resulting *Shape parameters* are $1, 1, 2, 0.5$

Shape

- Shape of a list of numbers is the pattern of relative interval sizes, moving from left to right.
- Consider the list $20, 30, 40, 60, 65$
- We can compute the *unscaled* distances between the numbers as 10,10,20,5.
- The relative distances are obtained by dividing all the unscaled distances by the first nonzero value.
- The resulting *Shape parameters* are $1, 1, 2, 0.5$

Shape

- Shape of a list of numbers is the pattern of relative interval sizes, moving from left to right.
- Consider the list $20, 30, 40, 60, 65$
- We can compute the *unscaled* distances between the numbers as 10,10,20,5.
- The relative distances are obtained by dividing all the unscaled distances by the first nonzero value.
- The resulting *Shape parameters* are $1, 1, 2, 0.5$

Shape

- Shape of a list of numbers is the pattern of relative interval sizes, moving from left to right.
- Consider the list $20, 30, 40, 60, 65$
- We can compute the *unscaled* distances between the numbers as 10,10,20,5.
- The relative distances are obtained by dividing all the unscaled distances by the first nonzero value.
- The resulting *Shape parameters* are $1, 1, 2, 0.5$

Shape

- To test yourself, answer the following. What are the Shape parameters for:
  - $1, 2, 3, 6, 10$
  - $2, 4, 6, 12, 20$
  - $13, 16, 19, 28, 40$

Shape

- To test yourself, answer the following. What are the Shape parameters for:
  - $1, 2, 3, 6, 10$
  - $2, 4, 6, 12, 20$
  - $13, 16, 19, 28, 40$

Shape

- To test yourself, answer the following. What are the Shape parameters for:
  - $1, 2, 3, 6, 10$
  - $2, 4, 6, 12, 20$
  - $13, 16, 19, 28, 40$

Shape

- To test yourself, answer the following. What are the Shape parameters for:
  - $1, 2, 3, 6, 10$
  - $2, 4, 6, 12, 20$
  - $13, 16, 19, 28, 40$

Effect of Listwise Operations

- Now I have a serious question for you. Why would we want to re-express a list of numbers in terms of Location, Spread, and Shape?
- That's actually a pretty profound question, so let's jump past it and ask some more basic questions (C.P.):
  - What is the effect of listwise addition(subtraction) on Location, Spread, and Shape?
  - What is the effect of listwise multiplication(division) by a positive number on Location, Spread, and Shape?

Effect of Listwise Operations

- Now I have a serious question for you. Why would we want to re-express a list of numbers in terms of Location, Spread, and Shape?
- That's actually a pretty profound question, so let's jump past it and ask some more basic questions (C.P.):
  - What is the effect of listwise addition(subtraction) on Location, Spread, and Shape?
  - What is the effect of listwise multiplication (division) by a positive number on Location, Spread, and Shape?

Effect of Listwise Operations

- Now I have a serious question for you. Why would we want to re-express a list of numbers in terms of Location, Spread, and Shape?
- That's actually a pretty profound question, so let's jump past it and ask some more basic questions (C.P.):
  - What is the effect of listwise addition(subtraction) on Location, Spread, and Shape?
  - What is the effect of listwise multiplication (division) by a positive number on Location, Spread, and Shape?

Effect of Listwise Operations

- Now I have a serious question for you. Why would we want to re-express a list of numbers in terms of Location, Spread, and Shape?
- That's actually a pretty profound question, so let's jump past it and ask some more basic questions (C.P.):
  - What is the effect of listwise addition(subtraction) on Location, Spread, and Shape?
  - What is the effect of listwise multiplication (division) by a positive number on Location, Spread, and Shape?

The Vulnerability Box

- Let's present our results in a summary table I'll refer to as The Vulnerability Box. (Einstein would probably call it the Invariance Box.)

Operation		Effect on	
	Location	Spread	Shape
$+$	$+$		
$-$	$-$		
$\times$	$\times$	$\times$	
$\div$	$\div$	$\div$	

Exploiting the Vulnerability Box

Some Examples

- Tracking changes in a list of numbers
- Rescaling numbers (Changing the Location and/or Spread without affecting Shape).
- Deriving statistical theory.

Exploiting the Vulnerability Box

Some Examples

- Tracking changes in a list of numbers
- Rescaling numbers (Changing the Location and/or Spread without affecting Shape).
- Deriving statistical theory.

Exploiting the Vulnerability Box
Some Examples

- Tracking changes in a list of numbers
- Rescaling numbers (Changing the Location and/or Spread without affecting Shape).
- Deriving statistical theory.

Tracking Changes
Effect on Location and Spread

- The Vulnerability Box can be used to examine any operation that can be expressed as a sequence of listwise additions, subtractions, multiplications, and/or divisions.
- That takes in a lot more territory than it might seem.
- Suppose you have a list of X's with a Location of 75 and a Spread of 20. What will the Location and Spread become if you convert them to Y's with this formula?

$$Y = 4\left(\frac{2X + 30}{20}\right) + 2 \tag{3}$$

- How about Spread?
- What about Shape?

Tracking Changes
Effect on Location and Spread

- The Vulnerability Box can be used to examine any operation that can be expressed as a sequence of listwise additions, subtractions, multiplications, and/or divisions.
- That takes in a lot more territory than it might seem.
- Suppose you have a list of X's with a Location of 75 and a Spread of 20. What will the Location and Spread become if you convert them to Y's with this formula?

$$Y = 4\left(\frac{2X + 30}{20}\right) + 2 \qquad (3)$$

- How about Spread?
- What about Shape?

Tracking Changes
Effect on Location and Spread

- The Vulnerability Box can be used to examine any operation that can be expressed as a sequence of listwise additions, subtractions, multiplications, and/or divisions.
- That takes in a lot more territory than it might seem.
- Suppose you have a list of X's with a Location of 75 and a Spread of 20. What will the Location and Spread become if you convert them to Y's with this formula?

$$Y = 4\left(\frac{2X + 30}{20}\right) + 2 \tag{3}$$

- How about Spread?
- What about Shape?

Tracking Changes
Effect on Location and Spread

- The Vulnerability Box can be used to examine any operation that can be expressed as a sequence of listwise additions, subtractions, multiplications, and/or divisions.
- That takes in a lot more territory than it might seem.
- Suppose you have a list of X's with a Location of 75 and a Spread of 20. What will the Location and Spread become if you convert them to Y's with this formula?

$$Y = 4\left(\frac{2X + 30}{20}\right) + 2 \tag{3}$$

- How about Spread?
- What about Shape?

Tracking Changes
Effect on Location and Spread

- The Vulnerability Box can be used to examine any operation that can be expressed as a sequence of listwise additions, subtractions, multiplications, and/or divisions.
- That takes in a lot more territory than it might seem.
- Suppose you have a list of X's with a Location of 75 and a Spread of 20. What will the Location and Spread become if you convert them to Y's with this formula?

$$Y = 4\left(\frac{2X + 30}{20}\right) + 2 \tag{3}$$

- How about Spread?
- What about Shape?

Solutions

- Location starts at 75. All listwise operations affect location: Multiply by 2 ($75 \times 2 = 150$), Add 30 (180), Divide by 20 (9), Multiply by 4 (36), Add 2 (38).
- Spread starts at 20. Only multiplication or division listwise operations affect spread: Multiply by 2 (40), Add 30 (no change 40), Divide by 20 (2), Multiply by 4 (8), Add 2 (no change 8).
- Shape will stay the same.

Solutions

- Location starts at 75. All listwise operations affect location: Multiply by 2 ($75 \times 2 = 150$), Add 30 (180), Divide by 20 (9), Multiply by 4 (36), Add 2 (38).
- Spread starts at 20. Only multiplication or division listwise operations affect spread: Multiply by 2 (40), Add 30 (no change 40), Divide by 20 (2), Multiply by 4 (8), Add 2 (no change 8).
- Shape will stay the same.

Solutions

- Location starts at 75. All listwise operations affect location: Multiply by 2 ($75 \times 2 = 150$), Add 30 (180), Divide by 20 (9), Multiply by 4 (36), Add 2 (38).
- Spread starts at 20. Only multiplication or division listwise operations affect spread: Multiply by 2 (40), Add 30 (no change 40), Divide by 20 (2), Multiply by 4 (8), Add 2 (no change 8).
- Shape will stay the same.

Rescaling Numbers

- Often we will find ourselves with a set of numbers that has an appropriate Shape, but an inappropriate Location and/or Spread.
- The classic example is a set of course grades that result from an exam that is fundamentally well structured, but too difficult (or perhaps too easy).
- For example, I give an exam and the Location is $M = 50$, and the Spread is $S = 40$, while a more typical set of grades would be $M = 80$ and $S = 20$.
- What can I do?

Rescaling Numbers

- Often we will find ourselves with a set of numbers that has an appropriate Shape, but an inappropriate Location and/or Spread.
- The classic example is a set of course grades that result from an exam that is fundamentally well structured, but too difficult (or perhaps too easy).
- For example, I give an exam and the Location is $M = 50$, and the Spread is $S = 40$, while a more typical set of grades would be $M = 80$ and $S = 20$.
- What can I do?

Rescaling Numbers

- Often we will find ourselves with a set of numbers that has an appropriate Shape, but an inappropriate Location and/or Spread.
- The classic example is a set of course grades that result from an exam that is fundamentally well structured, but too difficult (or perhaps too easy).
- For example, I give an exam and the Location is $M = 50$, and the Spread is $S = 40$, while a more typical set of grades would be $M = 80$ and $S = 20$.
- What can I do?

Rescaling Numbers

- Often we will find ourselves with a set of numbers that has an appropriate Shape, but an inappropriate Location and/or Spread.
- The classic example is a set of course grades that result from an exam that is fundamentally well structured, but too difficult (or perhaps too easy).
- For example, I give an exam and the Location is $M = 50$, and the Spread is $S = 40$, while a more typical set of grades would be $M = 80$ and $S = 20$.
- What can I do?

Rescaling Numbers

- One thing we have learned from the Vulnerability Box is that so long as we multiply (or divide) by a positive number, and add or subtract any number, the Shape of the grades will not change.
- A bit later, we will discover that when grades are "at an interval level of measurement," the Shape has all the information in the numbers that is not arbitrary, and that the Location and Spread are in fact arbitrary.
- Assuming our grades are at an interval level of measurement, we are now going to adjust the Location and Spread to values that are "culturally appropriate."

Rescaling Numbers

- One thing we have learned from the Vulnerability Box is that so long as we multiply (or divide) by a positive number, and add or subtract any number, the Shape of the grades will not change.
- A bit later, we will discover that when grades are “at an interval level of measurement,” the Shape has all the information in the numbers that is not arbitrary, and that the Location and Spread are in fact arbitrary.
- Assuming our grades are at an interval level of measurement, we are now going to adjust the Location and Spread to values that are “culturally appropriate.”

Rescaling Numbers

- One thing we have learned from the Vulnerability Box is that so long as we multiply (or divide) by a positive number, and add or subtract any number, the Shape of the grades will not change.
- A bit later, we will discover that when grades are "at an interval level of measurement," the Shape has all the information in the numbers that is not arbitrary, and that the Location and Spread are in fact arbitrary.
- Assuming our grades are at an interval level of measurement, we are now going to adjust the Location and Spread to values that are "culturally appropriate."

Rescaling Numbers

- If you stare at the Vulnerability Box long enough, you will notice that the stair-step shape of the filled boxes is telling you something important.
- Specifically, if you first adjust Spread by using multiplication, you can then adjust Location using addition/subtraction without changing Spread or Shape, thereby ending up with numbers with the same Shape you started with, but with exactly the Location and Spread you want.
- Let's see how this works.

Rescaling Numbers

- If you stare at the Vulnerability Box long enough, you will notice that the stair-step shape of the filled boxes is telling you something important.
- Specifically, if you first adjust Spread by using multiplication, you can then adjust Location using addition/subtraction without changing Spread or Shape, thereby ending up with numbers with the same Shape you started with, but with exactly the Location and Spread you want.
- Let's see how this works.

Rescaling Numbers

- If you stare at the Vulnerability Box long enough, you will notice that the stair-step shape of the filled boxes is telling you something important.
- Specifically, if you first adjust Spread by using multiplication, you can then adjust Location using addition/subtraction without changing Spread or Shape, thereby ending up with numbers with the same Shape you started with, but with exactly the Location and Spread you want.
- Let's see how this works.

Rescaling Numbers

- We start with numbers with $M = 50$, and the spread is $S = 40$, while what we want is the "culturally appropriate metric" of $M = 80$ and $S = 20$.
- We want to adjust the Spread first. It is currently 40, and we want 20. The lesson of the Vulnerability Box is that multiplication "comes straight through in the Spread and Location."
- So, if we multiply all the numbers by 1/2, we will multiply both the Spread and Location by 1/2. So if we started with numbers with $M = 50$ and $S = 40$, we will now have numbers with $M = 25$ and $S = 20$, and this set of numbers will have the same Shape as when we started.
- We can then adjust the Location to $M = 80$ by adding 55 to all the numbers. This will not change the Spread, and will result in a set of numbers with the same Shape as the original numbers, but a Location of 80 and a Spread of 20.

Rescaling Numbers

- We start with numbers with $M = 50$, and the spread is $S = 40$, while what we want is the "culturally appropriate metric" of $M = 80$ and $S = 20$.
- We want to adjust the Spread first. It is currently 40, and we want 20. The lesson of the Vulnerability Box is that multiplication "comes straight through in the Spread and Location."
- So, if we multiply all the numbers by 1/2, we will multiply both the Spread and Location by 1/2. So if we started with numbers with $M = 50$ and $S = 40$, we will now have numbers with $M = 25$ and $S = 20$, and this set of numbers will have the same Shape as when we started.
- We can then adjust the Location to $M = 80$ by adding 55 to all the numbers. This will not change the Spread, and will result in a set of numbers with the same Shape as the original numbers, but a Location of 80 and a Spread of 20.

Rescaling Numbers

- We start with numbers with $M = 50$, and the spread is $S = 40$, while what we want is the "culturally appropriate metric" of $M = 80$ and $S = 20$.
- We want to adjust the Spread first. It is currently 40, and we want 20. The lesson of the Vulnerability Box is that multiplication "comes straight through in the Spread and Location."
- So, if we multiply all the numbers by 1/2, we will multiply both the Spread and Location by 1/2. So if we started with numbers with $M = 50$ and $S = 40$, we will now have numbers with $M = 25$ and $S = 20$, and this set of numbers will have the same Shape as when we started.
- We can then adjust the Location to $M = 80$ by adding 55 to all the numbers. This will not change the Spread, and will result in a set of numbers with the same Shape as the original numbers, but a Location of 80 and a Spread of 20.

Rescaling Numbers

- We start with numbers with $M = 50$, and the spread is $S = 40$, while what we want is the "culturally appropriate metric" of $M = 80$ and $S = 20$.
- We want to adjust the Spread first. It is currently 40, and we want 20. The lesson of the Vulnerability Box is that multiplication "comes straight through in the Spread and Location."
- So, if we multiply all the numbers by $1/2$, we will multiply both the Spread and Location by $1/2$. So if we started with numbers with $M = 50$ and $S = 40$, we will now have numbers with $M = 25$ and $S = 20$, and this set of numbers will have the same Shape as when we started.
- We can then adjust the Location to $M = 80$ by adding 55 to all the numbers. This will not change the Spread, and will result in a set of numbers with the same Shape as the original numbers, but a Location of 80 and a Spread of 20.

Rescaling Numbers

Example (Rescaling Numbers)

We start with 30,50,70. By our current primitive measures of Location and Spread, these three evenly spaced numbers have a Location of 50 and a Spread of 40.

After multiplying by 1/2, we have three numbers 15,25,35 that have the desired Spread of 20, and are still evenly spaced.

Notice that the Location has changed, to $(1/2) \times 50 = 25$. We want it to be 80. So we must add the difference between where we are (25) and what we want (80), i.e., $80 - 25 = 55$.

After adding 55, we have 70,80,90.

Developing a Rescaling Formula

- The fundamental idea behind rescaling is to:
  1. Adjust Spread with multiplication/division.
  2. Examine where the Location has moved to, and calculate how far it is from the desired value.
  3. Adjust the Location with addition/subtraction.
- We can turn these informal ideas into a formal "prescription" or "set of formulas" for accomplishing linear rescaling. Let the multiplicative constant be designated as a, the additive constant b.
- Define M_x, S_x to be the current metric, M_y, S_y to be the desired metric.
- We know that, in order to adjust the Spread, we need to multiply by S_y/S_x, the ratio of the desired Spread over the current Spread.
- Once we multiply the X values by $a = S_y/S_x$, the Spread will become $aS_x = (S_y/S_x)S_x = S_y$, and the Location will become aM_x.
- If we then add $M_y - aM_x$, the Location will become $aM_x + (M_y - aM_x) = M_y$, and we will have accomplished our objective.
- So the "prescription for linear rescaling is $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.

Developing a Rescaling Formula

- The fundamental idea behind rescaling is to:
  1. Adjust Spread with multiplication/division.
  2. Examine where the Location has moved to, and calculate how far it is from the desired value.
  3. Adjust the Location with addition/subtraction.
- We can turn these informal ideas into a formal "prescription" or "set of formulas" for accomplishing linear rescaling. Let the multiplicative constant be designated as a, the additive constant b.
- Define M_x, S_x to be the current metric, M_y, S_y to be the desired metric.
- We know that, in order to adjust the Spread, we need to multiply by S_y/S_x, the ratio of the desired Spread over the current Spread.
- Once we multiply the X values by $a = S_y/S_x$, the Spread will become $aS_x = (S_y/S_x)S_x = S_y$, and the Location will become aM_x.
- If we then add $M_y - aM_x$, the Location will become $aM_x + (M_y - aM_x) = M_y$, and we will have accomplished our objective.
- So the "prescription for linear rescaling is $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.

Developing a Rescaling Formula

- The fundamental idea behind rescaling is to:
  1. Adjust Spread with multiplication/division.
  2. Examine where the Location has moved to, and calculate how far it is from the desired value.
  3. Adjust the Location with addition/subtraction.
- We can turn these informal ideas into a formal "prescription" or "set of formulas" for accomplishing linear rescaling. Let the multiplicative constant be designated as a, the additive constant b.
- Define $M_x,\ S_x$ to be the current metric, $M_y,\ S_y$ to be the desired metric.
- We know that, in order to adjust the Spread, we need to multiply by S_y/S_x, the ratio of the desired Spread over the current Spread.
- Once we multiply the X values by $a = S_y/S_x$, the Spread will become $aS_x = (S_y/S_x)S_x = S_y$, and the Location will become aM_x.
- If we then add $M_y - aM_x$, the Location will become $aM_x + (M_y - aM_x) = M_y$, and we will have accomplished our objective.
- So the "prescription for linear rescaling is $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.

Developing a Rescaling Formula

- The fundamental idea behind rescaling is to:
  1. Adjust Spread with multiplication/division.
  2. Examine where the Location has moved to, and calculate how far it is from the desired value.
  3. Adjust the Location with addition/subtraction.
- We can turn these informal ideas into a formal "prescription" or "set of formulas" for accomplishing linear rescaling. Let the multiplicative constant be designated as a, the additive constant b.
- Define M_x, S_x to be the current metric, M_y, S_y to be the desired metric.
- We know that, in order to adjust the Spread, we need to multiply by S_y/S_x, the ratio of the desired Spread over the current Spread.
- Once we multiply the X values by $a = S_y/S_x$, the Spread will become $aS_x = (S_y/S_x)S_x = S_y$, and the Location will become aM_x.
- If we then add $M_y - aM_x$, the Location will become $aM_x + (M_y - aM_x) = M_y$, and we will have accomplished our objective.
- So the "prescription for linear rescaling is $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.

Developing a Rescaling Formula

- The fundamental idea behind rescaling is to:
  1. Adjust Spread with multiplication/division.
  2. Examine where the Location has moved to, and calculate how far it is from the desired value.
  3. Adjust the Location with addition/subtraction.
- We can turn these informal ideas into a formal "prescription" or "set of formulas" for accomplishing linear rescaling. Let the multiplicative constant be designated as a, the additive constant b.
- Define M_x, S_x to be the current metric, M_y, S_y to be the desired metric.
- We know that, in order to adjust the Spread, we need to multiply by S_y/S_x, the ratio of the desired Spread over the current Spread.
- Once we multiply the X values by $a = S_y/S_x$, the Spread will become $aS_x = (S_y/S_x)S_x = S_y$, and the Location will become aM_x.
- If we then add $M_y - aM_x$, the Location will become $aM_x + (M_y - aM_x) = M_y$, and we will have accomplished our objective.
- So the "prescription for linear rescaling is $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.

Developing a Rescaling Formula

- The fundamental idea behind rescaling is to:
  1. Adjust Spread with multiplication/division.
  2. Examine where the Location has moved to, and calculate how far it is from the desired value.
  3. Adjust the Location with addition/subtraction.
- We can turn these informal ideas into a formal "prescription" or "set of formulas" for accomplishing linear rescaling. Let the multiplicative constant be designated as a, the additive constant b.
- Define M_x, S_x to be the current metric, M_y, S_y to be the desired metric.
- We know that, in order to adjust the Spread, we need to multiply by S_y/S_x, the ratio of the desired Spread over the current Spread.
- Once we multiply the X values by $a = S_y/S_x$, the Spread will become $aS_x = (S_y/S_x)S_x = S_y$, and the Location will become aM_x.
- If we then add $M_y - aM_x$, the Location will become $aM_x + (M_y - aM_x) = M_y$, and we will have accomplished our objective.
- So the "prescription for linear rescaling is $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.

Developing a Rescaling Formula

- The fundamental idea behind rescaling is to:
  1. Adjust Spread with multiplication/division.
  2. Examine where the Location has moved to, and calculate how far it is from the desired value.
  3. Adjust the Location with addition/subtraction.
- We can turn these informal ideas into a formal "prescription" or "set of formulas" for accomplishing linear rescaling. Let the multiplicative constant be designated as a, the additive constant b.
- Define M_x, S_x to be the current metric, M_y, S_y to be the desired metric.
- We know that, in order to adjust the Spread, we need to multiply by S_y/S_x, the ratio of the desired Spread over the current Spread.
- Once we multiply the X values by $a = S_y/S_x$, the Spread will become $aS_x = (S_y/S_x)S_x = S_y$, and the Location will become aM_x.
- If we then add $M_y - aM_x$, the Location will become $aM_x + (M_y - aM_x) = M_y$, and we will have accomplished our objective.
- So the "prescription for linear rescaling is $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.

Developing a Rescaling Formula

- The fundamental idea behind rescaling is to:
  1. Adjust Spread with multiplication/division.
  2. Examine where the Location has moved to, and calculate how far it is from the desired value.
  3. Adjust the Location with addition/subtraction.
- We can turn these informal ideas into a formal "prescription" or "set of formulas" for accomplishing linear rescaling. Let the multiplicative constant be designated as a, the additive constant b.
- Define M_x, S_x to be the current metric, M_y, S_y to be the desired metric.
- We know that, in order to adjust the Spread, we need to multiply by S_y/S_x, the ratio of the desired Spread over the current Spread.
- Once we multiply the X values by $a = S_y/S_x$, the Spread will become $aS_x = (S_y/S_x)S_x = S_y$, and the Location will become aM_x.
- If we then add $M_y - aM_x$, the Location will become $aM_x + (M_y - aM_x) = M_y$, and we will have accomplished our objective.
- So the "prescription for linear rescaling is $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.

Developing a Rescaling Formula

- The fundamental idea behind rescaling is to:
  1. Adjust Spread with multiplication/division.
  2. Examine where the Location has moved to, and calculate how far it is from the desired value.
  3. Adjust the Location with addition/subtraction.
- We can turn these informal ideas into a formal "prescription" or "set of formulas" for accomplishing linear rescaling. Let the multiplicative constant be designated as a, the additive constant b.
- Define $M_x,\ S_x$ to be the current metric, $M_y,\ S_y$ to be the desired metric.
- We know that, in order to adjust the Spread, we need to multiply by S_y/S_x, the ratio of the desired Spread over the current Spread.
- Once we multiply the X values by $a = S_y/S_x$, the Spread will become $aS_x = (S_y/S_x)S_x = S_y$, and the Location will become aM_x.
- If we then add $M_y - aM_x$, the Location will become $aM_x + (M_y - aM_x) = M_y$, and we will have accomplished our objective.
- So the "prescription for linear rescaling is $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.

Developing a Rescaling Formula

- The fundamental idea behind rescaling is to:
  1. Adjust Spread with multiplication/division.
  2. Examine where the Location has moved to, and calculate how far it is from the desired value.
  3. Adjust the Location with addition/subtraction.
- We can turn these informal ideas into a formal "prescription" or "set of formulas" for accomplishing linear rescaling. Let the multiplicative constant be designated as a, the additive constant b.
- Define M_x, S_x to be the current metric, M_y, S_y to be the desired metric.
- We know that, in order to adjust the Spread, we need to multiply by S_y/S_x, the ratio of the desired Spread over the current Spread.
- Once we multiply the X values by $a = S_y/S_x$, the Spread will become $aS_x = (S_y/S_x)S_x = S_y$, and the Location will become aM_x.
- If we then add $M_y - aM_x$, the Location will become $aM_x + (M_y - aM_x) = M_y$, and we will have accomplished our objective.
- So the "prescription for linear rescaling is $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.

Deriving Statistical Theory

- The Vulnerability Box rules work for concrete lists of numbers, but they also work in more abstract circumstances.
- Consider the following example:
  - You have a set of numbers X with Location M_x and Spread S_x that is not zero.

Deriving Statistical Theory

- The Vulnerability Box rules work for concrete lists of numbers, but they also work in more abstract circumstances.
- Consider the following example:
  - You have a set of numbers X with Location M_x and Spread S_x that is not zero.
  - You transform them all via the following formula

$$Z_x = \frac{X - M_x}{S_x} \tag{4}$$

  - What will be the Location, Spread, and Shape of the new numbers?

Deriving Statistical Theory

- The Vulnerability Box rules work for concrete lists of numbers, but they also work in more abstract circumstances.
- Consider the following example:
  - You have a set of numbers X with Location M_x and Spread S_x that is not zero.
  - You transform them all via the following formula

$$Z_x = \frac{X - M_x}{S_x} \tag{4}$$

  - What will be the Location, Spread, and Shape of the new numbers?

Deriving Statistical Theory

- The Vulnerability Box rules work for concrete lists of numbers, but they also work in more abstract circumstances.
- Consider the following example:
  - You have a set of numbers X with Location M_x and Spread S_x that is not zero.
  - You transform them all via the following formula

$$Z_x = \frac{X - M_x}{S_x} \tag{4}$$

  - What will be the Location, Spread, and Shape of the new numbers?

Deriving Statistical Theory

- The Vulnerability Box rules work for concrete lists of numbers, but they also work in more abstract circumstances.
- Consider the following example:
  - You have a set of numbers X with Location M_x and Spread S_x that is not zero.
  - You transform them all via the following formula

$$Z_x = \frac{X - M_x}{S_x} \tag{4}$$

  - What will be the Location, Spread, and Shape of the new numbers?

Deriving Statistical Theory

- To derive the answer, we first recognize that the Shape of the numbers will not change, since the transformation can be viewed as a subtraction followed by a division, and the divisor is always positive.
- We can deduce the Location and Spread of the numbers by simply applying the Vulnerability Box rules.
- We start with M_x and S_x, and subtract M_x from all the numbers. The Vulnerability Box tells us that this will not affect the Spread, which will stay at S_x, while the Location will change to $M_x - M_x = 0$.
- We now have numbers with a Location of 0 and a Spread of S_x. If we divide them all by S_x, we will divide both the Location and Spread by S_x. The result is that the Location will be $0/S_x = 0$, and the Spread will be $S_x/S_x = 1$.

Deriving Statistical Theory

- To derive the answer, we first recognize that the Shape of the numbers will not change, since the transformation can be viewed as a subtraction followed by a division, and the divisor is always positive.
- We can deduce the Location and Spread of the numbers by simply applying the Vulnerability Box rules.
- We start with M_x and S_x, and subtract M_x from all the numbers. The Vulnerability Box tells us that this will not affect the Spread, which will stay at S_x, while the Location will change to $M_x - M_x = 0$.
- We now have numbers with a Location of 0 and a Spread of S_x. If we divide them all by S_x, we will divide both the Location and Spread by S_x. The result is that the Location will be $0/S_x = 0$, and the Spread will be $S_x/S_x = 1$.

Deriving Statistical Theory

- To derive the answer, we first recognize that the Shape of the numbers will not change, since the transformation can be viewed as a subtraction followed by a division, and the divisor is always positive.
- We can deduce the Location and Spread of the numbers by simply applying the Vulnerability Box rules.
- We start with M_x and S_x, and subtract M_x from all the numbers. The Vulnerability Box tells us that this will not affect the Spread, which will stay at S_x, while the Location will change to $M_x - M_x = 0$.
- We now have numbers with a Location of 0 and a Spread of S_x. If we divide them all by S_x, we will divide both the Location and Spread by S_x. The result is that the Location will be $0/S_x = 0$, and the Spread will be $S_x/S_x = 1$.

Deriving Statistical Theory

- To derive the answer, we first recognize that the Shape of the numbers will not change, since the transformation can be viewed as a subtraction followed by a division, and the divisor is always positive.
- We can deduce the Location and Spread of the numbers by simply applying the Vulnerability Box rules.
- We start with M_x and S_x, and subtract M_x from all the numbers. The Vulnerability Box tells us that this will not affect the Spread, which will stay at S_x, while the Location will change to $M_x - M_x = 0$.
- We now have numbers with a Location of 0 and a Spread of S_x. If we divide them all by S_x, we will divide both the Location and Spread by S_x. The result is that the Location will be $0/S_x = 0$, and the Spread will be $S_x/S_x = 1$.

Deriving Statistical Theory

- We have proven that for any list of numbers with non-zero spread, the "Z-score transformation" produces numbers with the same Shape as the original numbers, but a Location of 0 and a Spread of 1.

Properties of Z-Scores

- If a set of numbers has a Location of 0 and a Spread of 1, we say that they are (according to whatever the current (fixed) definitions of Location and Spread might be) "in Z-score form."
- Consider the following question. Suppose a set of numbers is in Z-score form. Suppose we define the "metric of the numbers" to be their Location and Spread.
- How can we transform these Z-scores into any other desired metric?
- Let's consider the Spread first. It is currently 1. We want it to be something else. What do we need to do to the numbers to change the Spread to that "Something Else"? (Answer from C.P.)
- Will the Location have changed? No, it is still zero. So what do we need to do to the numbers to adjust the Location to something else? (Answer from C.P.)

Properties of Z-Scores

- If a set of numbers has a Location of 0 and a Spread of 1, we say that they are (according to whatever the current (fixed) definitions of Location and Spread might be) "in Z-score form."
- Consider the following question. Suppose a set of numbers is in Z-score form. Suppose we define the "metric of the numbers" to be their Location and Spread.
- How can we transform these Z-scores into any other desired metric?
- Let's consider the Spread first. It is currently 1. We want it to be something else. What do we need to do to the numbers to change the Spread to that "Something Else"? (Answer from C.P.)
- Will the Location have changed? No, it is still zero. So what do we need to do to the numbers to adjust the Location to something else? (Answer from C.P.)

Properties of Z-Scores

- If a set of numbers has a Location of 0 and a Spread of 1, we say that they are (according to whatever the current (fixed) definitions of Location and Spread might be) "in Z-score form."
- Consider the following question. Suppose a set of numbers is in Z-score form. Suppose we define the "metric of the numbers" to be their Location and Spread.
- How can we transform these Z-scores into any other desired metric?
- Let's consider the Spread first. It is currently 1. We want it to be something else. What do we need to do to the numbers to change the Spread to that "Something Else"? (Answer from C.P.)
- Will the Location have changed? No, it is still zero. So what do we need to do to the numbers to adjust the Location to something else? (Answer from C.P.)

Properties of Z-Scores

- If a set of numbers has a Location of 0 and a Spread of 1, we say that they are (according to whatever the current (fixed) definitions of Location and Spread might be) “in Z-score form.”
- Consider the following question. Suppose a set of numbers is in Z-score form. Suppose we define the “metric of the numbers” to be their Location and Spread.
- How can we transform these Z-scores into any other desired metric?
- Let’s consider the Spread first. It is currently 1. We want it to be something else. What do we need to do to the numbers to change the Spread to that “Something Else”? (Answer from C.P.)
- Will the Location have changed? No, it is still zero. So what do we need to do to the numbers to adjust the Location to something else? (Answer from C.P.)

Properties of Z-Scores

- If a set of numbers has a Location of 0 and a Spread of 1, we say that they are (according to whatever the current (fixed) definitions of Location and Spread might be) "in Z-score form."
- Consider the following question. Suppose a set of numbers is in Z-score form. Suppose we define the "metric of the numbers" to be their Location and Spread.
- How can we transform these Z-scores into any other desired metric?
- Let's consider the Spread first. It is currently 1. We want it to be something else. What do we need to do to the numbers to change the Spread to that "Something Else"? (Answer from C.P.)
- Will the Location have changed? No, it is still zero. So what do we need to do to the numbers to adjust the Location to something else? (Answer from C.P.)

Properties of Z-Scores

- That's right, to transform $Z - scores$ into any desired metric, multiply them by the desired Spread, and then add the desired Location.
- Let's try to get a conceptual handle on what that means.
- First of all, in an algebraic sense, we might say that the result is obvious.
- Algebraically, if

$$Z_x = \frac{X - M_x}{S_x} \tag{5}$$

then, of course

$$X = S_x Z_x + M_x \tag{6}$$

Properties of Z-Scores

- That's right, to transform $Z - scores$ into any desired metric, multiply them by the desired Spread, and then add the desired Location.
- Let's try to get a conceptual handle on what that means.
- First of all, in an algebraic sense, we might say that the result is obvious.
- Algebraically, if

$$Z_x = \frac{X - M_x}{S_x} \tag{5}$$

then, of course

$$X = S_x Z_x + M_x \tag{6}$$

Properties of Z-Scores

- That's right, to transform $Z - scores$ into any desired metric, multiply them by the desired Spread, and then add the desired Location.
- Let's try to get a conceptual handle on what that means.
- First of all, in an algebraic sense, we might say that the result is obvious.
- Algebraically, if

$$Z_x = \frac{X - M_x}{S_x} \tag{5}$$

then, of course

$$X = S_x Z_x + M_x \tag{6}$$

Properties of Z-Scores

- That's right, to transform $Z - scores$ into any desired metric, multiply them by the desired Spread, and then add the desired Location.
- Let's try to get a conceptual handle on what that means.
- First of all, in an algebraic sense, we might say that the result is obvious.
- Algebraically, if

$$Z_x = \frac{X - M_x}{S_x} \tag{5}$$

then, of course

$$X = S_x Z_x + M_x \tag{6}$$

Properties of Z-Scores

- Notice that Z-scores, in an important sense, "remove the metric" from a set of numbers, or at least establish a very convenient fixed metric.
- A somewhat more subtle property is that *Z-scores for a list of numbers are invariant under any linear rescaling of the raw scores.*
- What did I mean by that?

Properties of Z-Scores

- Notice that Z-scores, in an important sense, "remove the metric" from a set of numbers, or at least establish a very convenient fixed metric.
- A somewhat more subtle property is that *Z-scores for a list of numbers are invariant under any linear rescaling of the raw scores.*
- What did I mean by that?

Properties of Z-Scores

- Notice that Z-scores, in an important sense, "remove the metric" from a set of numbers, or at least establish a very convenient fixed metric.
- A somewhat more subtle property is that *Z-scores for a list of numbers are invariant under any linear rescaling of the raw scores.*
- What did I mean by that?

Properties of Z-Scores

The Linear Transformation

- A *positive linear transformation* of a set of scores X into a new set of scores Y can be written as

$$Y = aX + b \tag{7}$$

- *Any sequence of additions, subtractions, multiplications, or divisions by positive numbers can be expressed as a single linear transformation* of the form $Y = aX + b$.
- Suppose for example, we have a set of numbers and multiply them all by d, subtract e from all of them, divide all those numbers by f, and add g to all the resulting numbers.
- We can express that sequence as

$$\begin{aligned} Y &= \frac{dX - e}{f} + g \\ &= \frac{d}{f}X - \frac{e}{f} + g \\ &= \frac{d}{f}X + (g - \frac{e}{f} \end{aligned}$$

- So when I said "linear rescaling" in the previous slide, I simply meant any sequence of additions, subtractions, multiplications, or divisions by positive numbers.

Properties of Z-Scores
The Linear Transformation

- A *positive linear transformation* of a set of scores X into a new set of scores Y can be written as

$$Y = aX + b \tag{7}$$

- *Any sequence of additions, subtractions, multiplications, or divisions by positive numbers can be expressed as a single linear transformation* of the form $Y = aX + b$.
- Suppose for example, we have a set of numbers and multiply them all by d, subtract e from all of them, divide all those numbers by f, and add g to all the resulting numbers.
- We can express that sequence as

$$\begin{aligned} Y &= \frac{dX - e}{f} + g \\ &= \frac{d}{f}X - \frac{e}{f} + g \\ &= \frac{d}{f}X + (g - \frac{e}{f} \end{aligned}$$

- So when I said "linear rescaling" in the previous slide, I simply meant any sequence of additions, subtractions, multiplications, or divisions by positive numbers.

Properties of Z-Scores

The Linear Transformation

- A *positive linear transformation* of a set of scores X into a new set of scores Y can be written as

$$Y = aX + b \tag{7}$$

- *Any sequence of additions, subtractions, multiplications, or divisions by positive numbers can be expressed as a single linear transformation* of the form $Y = aX + b$.
- Suppose for example, we have a set of numbers and multiply them all by d, subtract e from all of them, divide all those numbers by f, and add g to all the resulting numbers.
- We can express that sequence as

$$\begin{aligned} Y &= \frac{dX - e}{f} + g \\ &= \frac{d}{f}X - \frac{e}{f} + g \\ &= \frac{d}{f}X + (g - \frac{e}{f} \end{aligned}$$

- So when I said "linear rescaling" in the previous slide, I simply meant any sequence of additions, subtractions, multiplications, or divisions by positive numbers.

Properties of Z-Scores
The Linear Transformation

- A *positive linear transformation* of a set of scores X into a new set of scores Y can be written as

$$Y = aX + b \tag{7}$$

- *Any sequence of additions, subtractions, multiplications, or divisions by positive numbers can be expressed as a single linear transformation* of the form $Y = aX + b$.
- Suppose for example, we have a set of numbers and multiply them all by d, subtract e from all of them, divide all those numbers by f, and add g to all the resulting numbers.
- We can express that sequence as

$$\begin{aligned} Y &= \frac{dX - e}{f} + g \\ &= \frac{d}{f}X - \frac{e}{f} + g \\ &= \frac{d}{f}X + (g - \frac{e}{f} \end{aligned}$$

- So when I said "linear rescaling" in the previous slide, I simply meant any sequence of additions, subtractions, multiplications, or divisions by positive numbers.

Properties of Z-Scores
The Linear Transformation

- A *positive linear transformation* of a set of scores X into a new set of scores Y can be written as

$$Y = aX + b \tag{7}$$

- *Any sequence of additions, subtractions, multiplications, or divisions by positive numbers can be expressed as a single linear transformation* of the form $Y = aX + b$.
- Suppose for example, we have a set of numbers and multiply them all by d, subtract e from all of them, divide all those numbers by f, and add g to all the resulting numbers.
- We can express that sequence as

$$\begin{aligned} Y &= \frac{dX - e}{f} + g \\ &= \frac{d}{f}X - \frac{e}{f} + g \\ &= \frac{d}{f}X + (g - \frac{e}{f} \end{aligned}$$

- So when I said "linear rescaling" in the previous slide, I simply meant any sequence of additions, subtractions, multiplications, or divisions by positive numbers.

Three Evenly Spaced Numbers

- Remember that all the properties of the Vulnerability Box and Z-scores hold for any reasonable measures of Location, Spread, and Shape, so long as you keep the definition consistent within the discussion.
- For reasons that will become obvious a little later, I want to restrict our next few discussion points to sets of 3 evenly spaced numbers.
- Moreover, for sets of 3 evenly spaced numbers, I'm going to define Location as before (i.e., the middle value), but I'm going to redefine the Spread S to be the inter-number spacing.
- By these new definitions, the set of X numbers 70,80,90 has Location 80 and Spread 10.

Three Evenly Spaced Numbers

- Remember that all the properties of the Vulnerability Box and Z-scores hold for any reasonable measures of Location, Spread, and Shape, so long as you keep the definition consistent within the discussion.
- For reasons that will become obvious a little later, I want to restrict our next few discussion points to sets of 3 evenly spaced numbers.
- Moreover, for sets of 3 evenly spaced numbers, I'm going to define Location as before (i.e., the middle value), but I'm going to redefine the Spread S to be the inter-number spacing.
- By these new definitions, the set of X numbers 70,80,90 has Location 80 and Spread 10.

Three Evenly Spaced Numbers

- Remember that all the properties of the Vulnerability Box and Z-scores hold for any reasonable measures of Location, Spread, and Shape, so long as you keep the definition consistent within the discussion.
- For reasons that will become obvious a little later, I want to restrict our next few discussion points to sets of 3 evenly spaced numbers.
- Moreover, for sets of 3 evenly spaced numbers, I'm going to define Location as before (i.e., the middle value), but I'm going to redefine the Spread S to be the inter-number spacing.
- By these new definitions, the set of X numbers 70,80,90 has Location 80 and Spread 10.

Three Evenly Spaced Numbers

- Remember that all the properties of the Vulnerability Box and Z-scores hold for any reasonable measures of Location, Spread, and Shape, so long as you keep the definition consistent within the discussion.
- For reasons that will become obvious a little later, I want to restrict our next few discussion points to sets of 3 evenly spaced numbers.
- Moreover, for sets of 3 evenly spaced numbers, I'm going to define Location as before (i.e., the middle value), but I'm going to redefine the Spread S to be the inter-number spacing.
- By these new definitions, the set of X numbers 70,80,90 has Location 80 and Spread 10.

Properties of Z-Scores

- So, in effect, linear rescaling "puts the metric into" a list of numbers, while Z-scoring removes it.
- Let's look at a simple example of "invariance of Z-scores under linear transformation."
- Consider the X list 70,80,90.
- If we convert these numbers to Z-scores (using our new definition of spread) by subtracting 80 and dividing by 10, we get $-1, 0, +1$.
- Now, suppose we were to transform the X numbers into Y by multiplying by 1.1 and subtracting 9.
- What would be the Location and Spread of the Y numbers? (C.P.)
- What would be the Z-score values for the Y numbers?

Properties of Z-Scores

- So, in effect, linear rescaling "puts the metric into" a list of numbers, while Z-scoring removes it.
- Let's look at a simple example of "invariance of Z-scores under linear transformation."
- Consider the X list 70,80,90.
- If we convert these numbers to Z-scores (using our new definition of spread) by subtracting 80 and dividing by 10, we get $-1, 0, +1$.
- Now, suppose we were to transform the X numbers into Y by multiplying by 1.1 and subtracting 9.
- What would be the Location and Spread of the Y numbers? (C.P.)
- What would be the Z-score values for the Y numbers?

Properties of Z-Scores

- So, in effect, linear rescaling "puts the metric into" a list of numbers, while Z-scoring removes it.
- Let's look at a simple example of "invariance of Z-scores under linear transformation."
- Consider the X list 70,80,90.
- If we convert these numbers to Z-scores (using our new definition of spread) by subtracting 80 and dividing by 10, we get $-1, 0, +1$.
- Now, suppose we were to transform the X numbers into Y by multiplying by 1.1 and subtracting 9.
- What would be the Location and Spread of the Y numbers? (C.P.)
- What would be the Z-score values for the Y numbers?

Properties of Z-Scores

- So, in effect, linear rescaling "puts the metric into" a list of numbers, while Z-scoring removes it.
- Let's look at a simple example of "invariance of Z-scores under linear transformation."
- Consider the X list 70,80,90.
- If we convert these numbers to Z-scores (using our new definition of spread) by subtracting 80 and dividing by 10, we get $-1, 0, +1$.
- Now, suppose we were to transform the X numbers into Y by multiplying by 1.1 and subtracting 9.
- What would be the Location and Spread of the Y numbers? (C.P.)
- What would be the Z-score values for the Y numbers?

Properties of Z-Scores

- So, in effect, linear rescaling “puts the metric into” a list of numbers, while Z-scoring removes it.
- Let’s look at a simple example of “invariance of Z-scores under linear transformation.”
- Consider the X list 70,80,90.
- If we convert these numbers to Z-scores (using our new definition of spread) by subtracting 80 and dividing by 10, we get $-1, 0, +1$.
- Now, suppose we were to transform the X numbers into Y by multiplying by 1.1 and subtracting 9.
- What would be the Location and Spread of the Y numbers? (C.P.)
- What would be the Z-score values for the Y numbers?

Properties of Z-Scores

- So, in effect, linear rescaling "puts the metric into" a list of numbers, while Z-scoring removes it.
- Let's look at a simple example of "invariance of Z-scores under linear transformation."
- Consider the X list 70,80,90.
- If we convert these numbers to Z-scores (using our new definition of spread) by subtracting 80 and dividing by 10, we get $-1, 0, +1$.
- Now, suppose we were to transform the X numbers into Y by multiplying by 1.1 and subtracting 9.
- What would be the Location and Spread of the Y numbers? (C.P.)
- What would be the Z-score values for the Y numbers?

Properties of Z-Scores

- So, in effect, linear rescaling "puts the metric into" a list of numbers, while Z-scoring removes it.
- Let's look at a simple example of "invariance of Z-scores under linear transformation."
- Consider the X list 70,80,90.
- If we convert these numbers to Z-scores (using our new definition of spread) by subtracting 80 and dividing by 10, we get $-1, 0, +1$.
- Now, suppose we were to transform the X numbers into Y by multiplying by 1.1 and subtracting 9.
- What would be the Location and Spread of the Y numbers? (C.P.)
- What would be the Z-score values for the Y numbers?

Solutions

- The X list is 70,80,90. Since $M_x = 80$, $S_x = 10$.
- The Z_x scores are $-1, 0, +1$, since $(70 - 80)/10 = -1$, $(80 - 80)/10 = 0$, and $(90 - 80)/10 = +1$.
- The Y scores are 68,79,90, since $1.1 \times 70 - 9 = 68$, $1.1 \times 80 - 9 = 79$, and $1.1 \times 90 - 9 = 90$.
- By our revised definitions, these Y scores have a Location of 79, and a spread of 11.
- The scores have changed, but the Z_y scores are the same as the Z_x scores!
- For example, the first Y score is 68, and it has a Z-score of $(68 - 79)/11$, or -1.

Solutions

- The X list is 70,80,90. Since $M_x = 80$, $S_x = 10$.
- The Z_x scores are $-1, 0, +1$, since $(70 - 80)/10 = -1$, $(80 - 80)/10 = 0$, and $(90 - 80)/10 = +1$.
- The Y scores are 68,79,90, since $1.1 \times 70 - 9 = 68$, $1.1 \times 80 - 9 = 79$, and $1.1 \times 90 - 9 = 90$.
- By our revised definitions, these Y scores have a Location of 79, and a spread of 11.
- The scores have changed, but the Z_y scores are the same as the Z_x scores!
- For example, the first Y score is 68, and it has a Z-score of $(68 - 79)/11$, or -1.

Solutions

- The X list is 70,80,90. Since $M_x = 80$, $S_x = 10$.
- The Z_x scores are $-1, 0, +1$, since $(70 - 80)/10 = -1$, $(80 - 80)/10 = 0$, and $(90 - 80)/10 = +1$.
- The Y scores are 68,79,90, since $1.1 \times 70 - 9 = 68$, $1.1 \times 80 - 9 = 79$, and $1.1 \times 90 - 9 = 90$.
- By our revised definitions, these Y scores have a Location of 79, and a spread of 11.
- The scores have changed, but the Z_y scores are the same as the Z_x scores!
- For example, the first Y score is 68, and it has a Z-score of $(68 - 79)/11$, or -1.

Solutions

- The X list is 70,80,90. Since $M_x = 80$, $S_x = 10$.
- The Z_x scores are $-1, 0, +1$, since $(70 - 80)/10 = -1$, $(80 - 80)/10 = 0$, and $(90 - 80)/10 = +1$.
- The Y scores are 68,79,90, since $1.1 \times 70 - 9 = 68$, $1.1 \times 80 - 9 = 79$, and $1.1 \times 90 - 9 = 90$.
- By our revised definitions, these Y scores have a Location of 79, and a spread of 11.
- The scores have changed, but the Z_y scores are the same as the Z_x scores!
- For example, the first Y score is 68, and it has a Z-score of $(68 - 79)/11$, or -1.

Solutions

- The X list is 70,80,90. Since $M_x = 80$, $S_x = 10$.
- The Z_x scores are $-1, 0, +1$, since $(70 - 80)/10 = -1$, $(80 - 80)/10 = 0$, and $(90 - 80)/10 = +1$.
- The Y scores are 68,79,90, since $1.1 \times 70 - 9 = 68$, $1.1 \times 80 - 9 = 79$, and $1.1 \times 90 - 9 = 90$.
- By our revised definitions, these Y scores have a Location of 79, and a spread of 11.
- The scores have changed, but the Z_y scores are the same as the Z_x scores!
- For example, the first Y score is 68, and it has a Z-score of $(68 - 79)/11$, or -1.

Solutions

- The X list is 70,80,90. Since $M_x = 80$, $S_x = 10$.
- The Z_x scores are $-1, 0, +1$, since $(70 - 80)/10 = -1$, $(80 - 80)/10 = 0$, and $(90 - 80)/10 = +1$.
- The Y scores are 68,79,90, since $1.1 \times 70 - 9 = 68$, $1.1 \times 80 - 9 = 79$, and $1.1 \times 90 - 9 = 90$.
- By our revised definitions, these Y scores have a Location of 79, and a spread of 11.
- The scores have changed, but the Z_y scores are the same as the Z_x scores!
- For example, the first Y score is 68, and it has a Z-score of $(68 - 79)/11$, or -1.

Properties of Z-scores

- The preceding example indicates an even broader principle.
- Suppose two lists of numbers of the same length have the same Shape, the same Location, and the same Spread.
- Then, of course, the lists must be identical!
- So if two lists of numbers are the same length and have the same Shape, then if we linearly transform them to have the same metric (i.e., Location and Spread), then the two lists will be made identical.
- This fact is commonly exploited to equalize scores across different sections of a course.

Properties of Z-scores

- The preceding example indicates an even broader principle.
- Suppose two lists of numbers of the same length have the same Shape, the same Location, and the same Spread.
- Then, of course, the lists must be identical!
- So if two lists of numbers are the same length and have the same Shape, then if we linearly transform them to have the same metric (i.e., Location and Spread), then the two lists will be made identical.
- This fact is commonly exploited to equalize scores across different sections of a course.

Properties of Z-scores

- The preceding example indicates an even broader principle.
- Suppose two lists of numbers of the same length have the same Shape, the same Location, and the same Spread.
- Then, of course, the lists must be identical!
- So if two lists of numbers are the same length and have the same Shape, then if we linearly transform them to have the same metric (i.e., Location and Spread), then the two lists will be made identical.
- This fact is commonly exploited to equalize scores across different sections of a course.

Properties of Z-scores

- The preceding example indicates an even broader principle.
- Suppose two lists of numbers of the same length have the same Shape, the same Location, and the same Spread.
- Then, of course, the lists must be identical!
- So if two lists of numbers are the same length and have the same Shape, then if we linearly transform them to have the same metric (i.e., Location and Spread), then the two lists will be made identical.
- This fact is commonly exploited to equalize scores across different sections of a course.

Properties of Z-scores

- The preceding example indicates an even broader principle.
- Suppose two lists of numbers of the same length have the same Shape, the same Location, and the same Spread.
- Then, of course, the lists must be identical!
- So if two lists of numbers are the same length and have the same Shape, then if we linearly transform them to have the same metric (i.e., Location and Spread), then the two lists will be made identical.
- This fact is commonly exploited to equalize scores across different sections of a course.

Linear Transformation Rules Revisited

- If $Y = aX + b$, then

$$
\begin{aligned}
S_y &= |a| S_x \\
M_y &= a M_x + b
\end{aligned}
$$

- These results immediately follow from our Vulnerability Box results, since a accomplishes multiplication (or division) and b accomplishes addition (or subtraction).

Linear Transformation Rules Revisited

- If $Y = aX + b$, then

$$\begin{aligned} S_y &= |a|S_x \\ M_y &= aM_x + b \end{aligned}$$

- These results immediately follow from our Vulnerability Box results, since a accomplishes multiplication (or division) and b accomplishes addition (or subtraction).

- Consider the equations from the previous slide.

$$
\begin{aligned}
S_y &= |a|S_x \\
M_y &= aM_x + b
\end{aligned}
$$

- If we restrict ourselves to positive values of a, and manipulate a to the left of the first equation and b to the left side of the second equation, we obtain the following formulas for calculating a linear transformation to produce a desired metric:
    - If a set of scores X currently has a metric M_x, S_x, and we wish to linearly transform them via $Y = aX + b$ to a "desired metric" M_y, S_Y, the transformation formula must be

$$
\begin{aligned}
a &= \frac{S_y}{S_x} \\
b &= M_y - aM_x
\end{aligned}
$$

- Notice that we actually deduced these formulas earlier in this module by simply expressing our Vulnerability Box rules in mathematical notation.

- Consider the equations from the previous slide.

$$\begin{aligned} S_y &= |a| S_x \\ M_y &= a M_x + b \end{aligned}$$

- If we restrict ourselves to positive values of a, and manipulate a to the left of the first equation and b to the left side of the second equation, we obtain the following formulas for calculating a linear transformation to produce a desired metric:
  - If a set of scores X currently has a metric M_x, S_x, and we wish to linearly transform them via $Y = aX + b$ to a "desired metric" M_y, S_Y, the transformation formula must be

$$\begin{aligned} a &= \frac{S_y}{S_x} \\ b &= M_y - a M_x \end{aligned}$$

- Notice that we actually deduced these formulas earlier in this module by simply expressing our Vulnerability Box rules in mathematical notation.

- Consider the equations from the previous slide.

$$\begin{aligned} S_y &= |a|S_x \\ M_y &= aM_x + b \end{aligned}$$

- If we restrict ourselves to positive values of a, and manipulate a to the left of the first equation and b to the left side of the second equation, we obtain the following formulas for calculating a linear transformation to produce a desired metric:
  - If a set of scores X currently has a metric M_x, S_x, and we wish to linearly transform them via $Y = aX + b$ to a "desired metric" M_y, S_Y, the transformation formula must be

$$\begin{aligned} a &= \frac{S_y}{S_x} \\ b &= M_y - aM_x \end{aligned}$$

- Notice that we actually deduced these formulas earlier in this module by simply expressing our Vulnerability Box rules in mathematical notation.

- Consider the equations from the previous slide.

$$\begin{aligned} S_y &= |a| S_x \\ M_y &= a M_x + b \end{aligned}$$

- If we restrict ourselves to positive values of a, and manipulate a to the left of the first equation and b to the left side of the second equation, we obtain the following formulas for calculating a linear transformation to produce a desired metric:
  - If a set of scores X currently has a metric M_x, S_x, and we wish to linearly transform them via $Y = aX + b$ to a "desired metric" M_y, S_Y, the transformation formula must be

$$\begin{aligned} a &= \frac{S_y}{S_x} \\ b &= M_y - a M_x \end{aligned}$$

- Notice that we actually deduced these formulas earlier in this module by simply expressing our Vulnerability Box rules in mathematical notation.

Summary

- The Vulnerability Box laws are equivalent to the Laws of Linear Transformation:
  1. Addition, subtraction, multiplication and division all "come straight through" in the Location.
  2. Only multiplication and division come straight through in the Spread.
  3. So long as the multiplier/divisor is positive, none of the four basic arithmetic operations affect Shape.
  4. In algebraic notation, if $Y = aX + b$, then $M_y = aM_x + b$, $S_y = |a|S_x$.

Summary

- The Vulnerability Box laws are equivalent to the Laws of Linear Transformation:
  1. Addition, subtraction, multiplication and division all "come straight through" in the Location.
  2. Only multiplication and division come straight through in the Spread.
  3. So long as the multiplier/divisor is positive, none of the four basic arithmetic operations affect Shape.
  4. In algebraic notation, if $Y = aX + b$, then $M_y = aM_x + b$, $S_y = |a|S_x$.

Summary

- The Vulnerability Box laws are equivalent to the Laws of Linear Transformation:
  1. Addition, subtraction, multiplication and division all "come straight through" in the Location.
  2. Only multiplication and division come straight through in the Spread.
  3. So long as the multiplier/divisor is positive, none of the four basic arithmetic operations affect Shape.
  4. In algebraic notation, if $Y = aX + b$, then $M_y = aM_x + b$, $S_y = |a|S_x$.

Summary

- The Vulnerability Box laws are equivalent to the Laws of Linear Transformation:
  1. Addition, subtraction, multiplication and division all “come straight through” in the Location.
  2. Only multiplication and division come straight through in the Spread.
  3. So long as the multiplier/divisor is positive, none of the four basic arithmetic operations affect Shape.
  4. In algebraic notation, if $Y = aX + b$, then $M_y = aM_x + b$, $S_y = |a|S_x$.

Summary

- The Vulnerability Box laws are equivalent to the Laws of Linear Transformation:
  1. Addition, subtraction, multiplication and division all "come straight through" in the Location.
  2. Only multiplication and division come straight through in the Spread.
  3. So long as the multiplier/divisor is positive, none of the four basic arithmetic operations affect Shape.
  4. In algebraic notation, if $Y = aX + b$, then $M_y = aM_x + b$, $S_y = |a|S_x$.

Summary

- There are three equivalent approaches to linear rescaling:
  1. Informally adjust the Spread with multiplication, then adjust the Location with addition.
  2. Convert the X scores to Z scores first, then multiply by the desired Spread and add the desired Location.
  3. Use the linear transformation rule $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.

Summary

- There are three equivalent approaches to linear rescaling:
  1. Informally adjust the Spread with multiplication, then adjust the Location with addition.
  2. Convert the X scores to Z scores first, then multiply by the desired Spread and add the desired Location.
  3. Use the linear transformation rule $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.

Summary

- There are three equivalent approaches to linear rescaling:
  1. Informally adjust the Spread with multiplication, then adjust the Location with addition.
  2. Convert the X scores to Z scores first, then multiply by the desired Spread and add the desired Location.
  3. Use the linear transformation rule $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.

Summary

- There are three equivalent approaches to linear rescaling:
  1. Informally adjust the Spread with multiplication, then adjust the Location with addition.
  2. Convert the X scores to Z scores first, then multiply by the desired Spread and add the desired Location.
  3. Use the linear transformation rule $Y = aX + b$, where $a = S_y/S_x$, and $b = M_y - aM_x$.