## 1. Using the Normal Curve with Expected Values and Standard Errors (17.3)

*This section ties 17.1 & 17.2 together and it reminds us Chapters 4 and 5. You can borrow the normal curve to make statements about random processes (such as draws from a box, coin tosses, craps, roulette, your commuting time home, the support for the president, whatever).*

*All that is required is that you:*

(a) *calculate the expected value of the box and*

(b) *calculate the Standard Error based on both the number of draws and the SD (standard deviation) of the box*

*Then, you can calculate standard units with a familiar formula that has been modified:*

```
Z = (observed value - expected value)
     -------------------------------
               Standard Error
```

Example. *Let's go back to the 9 heads in 10 tosses of a coin idea. More formally, the Standard Error for the coin toss situation is* $\sqrt{10} * [(1-0)*\sqrt{((.5*.5)}]$ *(read 17.5 for the specifics on calculating the standard deviation for a 1,0 situation). Let's see how likely it is to get 9 heads in 10 tosses.*

*SE = 1.5811 and Z = ((9 -5) / 1.5811) = 2.53 or about 2.55. The area between + and - 2.55 is 98.92% which leaves 1% total outside of the area. So the chance of getting 9 heads or better is about 1/2 of a percent. The chance of getting 6 heads or more is about 25%*

*Your intuitive sense works well. The combination of the expected value, standard error, and normal curve validates your suspicions.*

*This same method can be used to figure out chances in all kinds of situations.*

## 4. SHORT CUTS & COUNTING (17.4 and 17.5)

*Finding Standard Deviations in this chapter can be difficult. Freedman offers you some handy short cuts*

(a) *If you have a situation with only two numbers, a quick formula for the standard deviation is:*
*(big number - small number) ** $\sqrt{\text{fraction with big number} * \text{fraction with small number}}$

(b) *If you have a situation with only two numbers and you can make one of them a zero and the other a one, then the formula becomes:* $\sqrt{\text{fraction with big number} * \text{fraction with small number}}$

(c) *In all other situations, you will need to "expand" the box by converting fractions into real relative values sometimes, for example, suppose you have 3 numbers in a box in the following proportions:*

| -1 | 0 | +1 |
|---|---|---|
| 1/6 | 1/2 | 1/3 |

This is the same as: -1, 0, 0, 0, 1, 1 and the average is .16667 and the S.D. is .6872. So if you had 12 draws from this box, you would expect a sum of 2 and a standard error of $\sqrt{12}$ * .6872 or 2.38

We can compute chances and probabilities for random processes using the approximate normality of the sum of draws. An example: consider the game of roulette in Las Vegas. There are 38 slots for a ball to fall into at random. 18 of the slots are colored red. 18 of the slots are colored black. 2 of the slots are colored green. For the simplest bet, either red or black, you bet $5. You will win $5 if the ball lands in your colored slot. You will lose your $5 if it falls in a color other than yours.

i. Sketch a reasonable box model for the game of roulette.

ii. What is the box average for this game?

iii. What is the box standard deviation for this game?

iv. Suppose you were to play the game 25 times. What is the probability (or chance) that you will break even (win $0) or do better? What is the chance that you will not break even?

v. After playing 25 times, what is the probability (or chance) that your net winnings fall between -$20 and $40?

http://www.engr.csufresno.edu/~patague/cgi-bin/Roulette.html (a roulette simulation)
http://www.vegas.com/gaming/gaming_tips/photos/roulette4.jpg (a picture of a real wheel)

## Standard Deviations of "Boxes"

The standard deviation is just the same standard deviation of Chapter 4, in Chapter 17 you are given slightly different or short cut formulas.

Example.

Suppose I have a list:

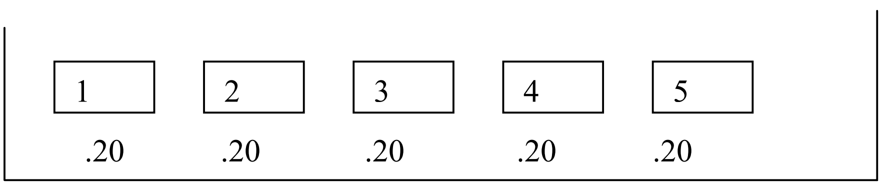
1, 2, 3, 4, 5

The average is 3, the SD is 1.4142 (to get these numbers, just use the formulas in Chapter 4)

Suppose I have the list:

1, 1, 2, 2, 3, 3, 4, 4, 5, 5

Guess what, the average is still 3, the SD is still 1.4142. The reason? The relative frequency of the values didn't change…in other words, in the first list, each number represents 1/5 of the list, in the second list, they still represent 1/5 of the list. So if I had a box that looks like this:




I could treat it in a number of ways (a) recognize that the proportions are the same so it's like a list that is 1, 2, 3, 4, 5 (b) or treat it as if I had 20 1's, 20 2's , 20 3's, 20 4's and 20 5's (c) or some other list that has the same relative frequency like 1,1,2,2,3,3,4,4,5,5

The average is 1+1+2+2+3+3+4+4+5+5 divided by 10 = 3 or you could use a simpler formula:

(.2*1) + (.2*2) + (.2*3) + (.2*4) + (.2*5) = 3. Notice where I get the .2 from?

The standard deviation is (using a chapter 4 formula):

$$\sqrt{\frac{(1-3)^2+(1-3)^2+(2-3)^2+(2-3)^2+(3-3)^2+(3-3)^2+(4-3)^2+(4-3)^2+(5-3)^2+(5-3)^2}{10}}=\sqrt{\frac{20}{10}}=\sqrt{2}$$

but you could simplify it:

$$\sqrt{\frac{2*(1-3)^2+2*(2-3)^2+2*(3-3)^2+2*(4-3)^2+2*(5-3)^2}{10}}=\sqrt{\frac{20}{10}}=\sqrt{2}$$

and you could simplify it further:

$$\sqrt{.2*(1-3)^2+.2*(2-3)^2+.2*(3-3)^2+.2*(4-3)^2+.2*(5-3)^2}=\sqrt{2}$$

notice where the .2 comes from?? The box proportions.

A more complicated box:

| +3 | +1.25 | -20 |
|---|---|---|
| .50 | .40 | .10 |

How should you treat this? Well, the simplest thing to do is probably convert it to a list like: 3,3,3,3,3,1.25,1.25,1.25,1.25,-20 and calculate an average and SD for this list like Chapter 4. Or use the tricks above on the previous page:

Box Average is (.50 * 3) + (.40 * 1.25) + (.10 * -20) = 0
Box Standard Deviation is:

$$\sqrt{.50*(3-0)^2+.40*(1.25-0)^2+.10*(-20-0)^2}=\sqrt{45.125}=6.7175$$

**Boxes with only 2 tickets are simpler, but the same general principles hold.**

Suppose I have a box that looks like this:

| +3 | +1.25 |
|---|---|
| .60 | .40 |

Expanded, this is just like a list that looks like: 3,3,3,3,3,3,1.25, 1.25, 1.25, 1.25

It's average is 2.3 and it's Standard Deviation is: .85732 (using Chapter 4 calculations)

but you can simplify it, according to Freedman (page 298) to

box average = (3 * .60) + (1.25 * .40) = 2.3
box standard deviation = (3 - 1.25) * $\sqrt{.60*.40}$ = .85732

Even simpler boxes are "one-zero" boxes:

| 1 | 0 |
|---|---|
| .90 | .10 |

This is like a list that looks like: 1,1,1,1,1,1,1,1,1,0 it's average is .9 it's standard deviation is: .3
but you can simplify it, according to Freedman (page 298) to:

box average = proportion of 1's in the box or .9
box standard deviation = (1-0)* $\sqrt{.90*.10}$ = .3 but even more simply to $\sqrt{.90*.10}$ = .30

this is how "one-zero" boxes work.