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:

1	2	3	4	5
.20	.20	.20	.20	.20

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.