Chapter 5-Measures of Variability

5.1 Variability of NoPassage group:

Range = 57 – 34 = 23
St. Dev. = 6.83
Variance = 46.62

5.3 The variability of the NoPassage group is much smaller than the variability of the Passage group. If this difference turns out to be reliable, it could possibly be explained by the fact that the questions for the Passage group are asking for more than guessing and test-taking skills, and there may be greater variability due to variability in knowledge. On the other hand, it is not uncommon to find one standard deviation equal to two to three times another in small samples.

5.5 Percentages within two standard deviations in Exercise 5.2

s = 10.61
$\overline{X} \pm 2(10.61) = 70.18 \pm 21.22 = 48.96$ — 91.4
16 scores (or 94%) lie within 2 standard deviations of the mean

5.7 Multiplying or dividing by a constant:

Original	2	3	4	4	5	5	9	$\overline{X}_1 = 4.57$	$s_1 = 2.23$
X * 2	4	6	8	8	10	10	18	$\overline{X}_2 = 9.14$	$s_2 = 4.45$
X / 2	1	1.5	2	2	2.5	2.5	4.5	$\overline{X}_3 = 2.29$	$s_3 = 1.11$

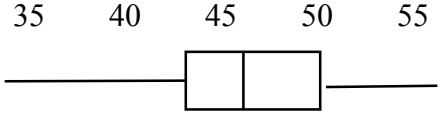
5.9 Convert revised data to mean = 0

Since adding or subtracting a constant will not change the standard deviation, but will change the mean, I can subtract 3.27 from every score for X_2 in Exercise 5.8, making the mean = 0, and keeping $s_2 = 1.0$. the new values are

X_3 -0.889 0.539 -1.842 0.539 -0.413 1.016 1.016 $\overline{X}_1 = 0$ $s_1 = 1$

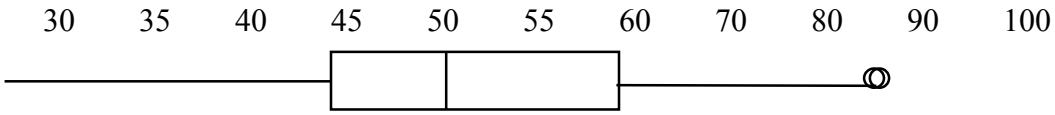
5.11 Boxplot for Exercise 5.1:

Median location = $(N + 1)/2 = 29/2 = 14.5$
Median = 46
Hinge location = (median location +1)/2 = 15/2 = 7.5
Hinge = 43 and 52
H-spread = 52 – 43 = 9
Inner fences = hinges $\pm$ 1.5*H-spread = hinges $\pm$ 1.5*9 = hinges $\pm$ 13.5 = 29.5 and 65.5
Adjacent values = 34 and 57



5.13 Boxplot for ADDSC:

Median location = $(N + 1)/2 = 89/2 = 44.5$
Median = 50
Hinge location = (median location +1)/2 = 45/2 = 22.5
Hinge = 44.5 and 60.5
H-spread = 60.5 – 44.5 = 16
Inner fences = hinges $\pm$ 1.5*H-spread = hinges $\pm$ 1.5*16 = hinges $\pm$ 24
= 20.5 and 85.5
Adjacent values = 26 and 78



5.15 Variance when you add a score equal to the mean.

$$s^2 = \frac{\Sigma(X-\overline{X})^2}{N-1} = \frac{\Sigma(X-\overline{X})^2}{27} = 46.6257$$

$$\Sigma(X-\overline{X})^2 = (N-1)s^2 = 27(46.6257) = 1258.8933$$

Adding a score equal to the mean will not change the sum of the deviations but will increase the denominator to 28.

$$s^2_{new} = \frac{\Sigma(X-\overline{X})^2}{N-1} = \frac{1258.8933}{28} = 44.96$$

Note that the new variance is (1-1/*N*) times the old variance.

in terms of $\sum(X-\overline{X})^2$ when they think of variability.

5.17 Sampling from a small population:

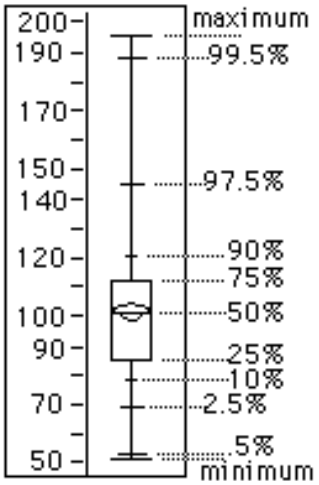
Population 1 2 3 4 5 Sample size = 3 Population variance = 2.00

Sample	Mean	Variance using *N*	Variance using *N*-1
1 1 1	1.00	0.00	0.00
1 1 2	1.33	0.22	0.33
1 1 3	1.67	0.89	1.33
1 1 4	2.00	2.00	3.00
1 1 5	2.33	3.56	5.33
1 2 1	1.33	0.22	0.33
...	...	...	...

5 5 3	4.33	0.89	1.33
5 5 4	4.67	0.22	0.33
5 5 5	5.00	0.00	0.00
Average	3.00	1.33	2.00

I admit that this problem is a pain in the neck, but I want students to see that using *N*-1 is the appropriate way to estimate population variances.

5.19 The following is a cut-and-paste from the JMP help screen. (I don't expect students to make all of these distinctions from what they are given, because many of the lines overlap.)



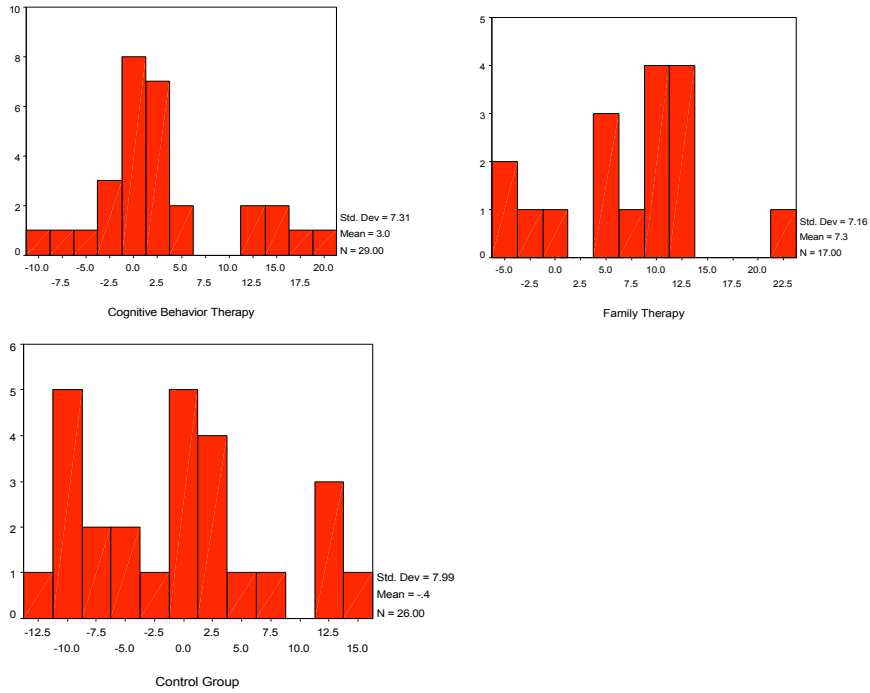
The quantile box plot shows selected quantiles on the response axis. The box shows the median as a line across the middle and the quartiles (25th and 75th percentiles) as its ends. The means diamond identifies the mean of the sample and the 95% confidence interval about the mean.

5.21 Treatment of anorexia:

I would hypothesize that the two treatment groups would show more of a weight gain than the control group, but I have no reason to predict which treatment group would do better. I would assume that the variability would be about the same within each group.

Complete (Before and After) data for the three groups—from which difference scores were derived:

	Cognitive Behavioral	**Family Therapy**	**Control**
Mean	3.01	7.26	-.45
Median	1.40	9.00	-.35
St. Dev.	7.31	7.16	7.99



If we look at the weight gain or loss, it would appear that the Control group remained stable, but the two treatment groups gained weight. The gain is greater for the Family Therapy group.

5.23 The descriptive statistics from SPSS are given below. The variable labels should be clear.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
COGBEHAV	29	-9.10	20.90	3.0069	7.30850	53.414
TRIM	19	-1.40	11.70	1.8000	3.04211	9.254
WINSOR	29	-1.40	11.70	2.9552	4.88851	23.898
Valid N (listwise)	19					

Notice that the Winsorized variance is considerably greater than the trimmed variance, as it should be. However, it is lower than the variance of the original data, reflecting the fact that the extreme values have been replaced. Cognitive behavior scores were positively skewed, with several quite high values and one or two low values. Trimming and Winsorizing reduced the influence of those values. This causes the Winsorized variance to be considerably smaller than the original variance. The trimmed mean is considerably smaller than the original mean, but the Winsorized mean is only slightly smaller.