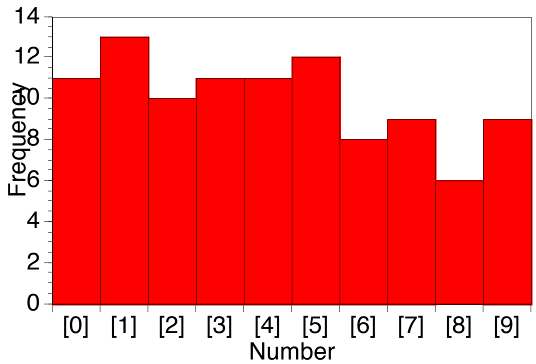
Chapter 12—Hypothesis Tests Applied to Means: One Sample

12.1 Distribution of 100 random digits:



12.3 The mean and standard deviation of the sample are 4.1 and 2.82, respectively, which are reasonably close to the parameters of the population for which the sample was drawn (4.5 and 2.6, respectively). The mean of the distribution of means is 4.28, which is somewhat closer to the population mean, and the standard deviation is 1.22.

a) The Central Limit theorem would predict a sampling distribution of the mean with a mean of 4.5 and a standard deviation of $2.6/\sqrt{5} = 1.16$.
b) These values are close to the values that we would expect.

12.5 If you had drawn 50 samples of size 15, the mean of the sampling distribution should still approximate the mean of the population, but the standard error of that distribution would now be only $2.67/\sqrt{15} = 0.689$.

.

12.7 Why doesn't the previous question address the issue of the state of North Dakota's educational system? These students are certainly not a random sample of high school students in North Dakota or elsewhere. Moreover, they scored above the mean of 500, which would certainly not be expected if North Dakota's system were inadequate. In addition, there is no definition of what is meant by "a terrible state," or a "good state," nor any idea of whether or not the SAT measures such a concept.

> I recognize that this probably sounds like "preaching," but here is a good place to remind students that there has to be a link between their dependent variable and what they think they are studying. It is all too easy for all of us to assume that we have measured something just because we have a number for it. You might get students discussing what kind of data

we would need to collect to answer such a question. They would probably find that it is a lot harder than they thought.

12.9 Unlike the results in the two previous questions, this interval probably is a fair estimate of the confidence interval for P/T ratio across the country. It is not itself biased by the bias in the sampling of SAT scores.

12.11 Weight gain exercise:

For these data the mean weight gain was 3.01 pounds, with a standard deviation of 7.3 pounds. This gives us

$$t = \frac{\overline{X} - \mu}{\frac{s}{\sqrt{N}}} = \frac{3.01 - 0}{\frac{7.3}{\sqrt{29}}} = \frac{3.01}{1.357} = 2.22$$

With 28 *df* the critical value at $\alpha = .05$, two-tailed, is 2.048, which will allow us to reject the null hypothesis and conclude that the girls gained weight at better than chance levels in this experiment.

There is an important movement within statistics right now in the direction of laying a much heavier emphasis on confidence limits than on null hypothesis tests. I think this is a very good example of a place where a behavioral scientist might make good use of a confidence interval. I didn't ask the students to calculate these limits, but they are 0.227 and 5.787. Students should think about what these limits mean and about why they are useful. (The computations in the next exercise give slightly different values, but that is due to rounding.)

12.13 The data are all over the place, with some gains as large as 20.9 lbs and some as low as —9.1. I suspect that we have an effect that works for some participants but not for others.

12.15 Effect size measure for data in Exercise 12.11:

One effect size measure would simply be the mean weight gain of 3.01 pounds. That statistic has real meaning to us, especially if we keep the size of a standard deviation in mind. A dubious alternative method would be to calculate an estimate of $\hat{d}$ using the standard deviation of the gain scores as our base.

$$\hat{d} = \frac{\overline{X}}{s} = \frac{3.01}{7.3} = 0.41$$

If I knew the standard deviation at baseline, that would make a good denominator. Unfortunately that information is not available, and 7.3 is the standard deviation of weight *gains*, and it is difficult to see how that creates a reasonable metric.

12.17 I needed to solve for t in Exercise 12.14 because I did not know the population variance.

12.19 Testing the null hypothesis that children under stress report lower levels of anxiety:

$$t = \frac{\overline{X} - \mu}{\frac{s}{\sqrt{N}}} = \frac{11.00 - 14.55}{\frac{6.085}{\sqrt{36}}} = \frac{-3.55}{1.014} = -3.50$$

With 35 *df* the critical value of t at $\alpha = .05$ (two-tailed) is ±2.03. We can reject H_0 and conclude that children under stress show significantly lower levels of anxiety than normal children.

Here is another situation where the data say that children report lower levels of anxiety, but it was necessary to first verify that their reports could be relied upon.

12.21 Yes, the results in Exercise 12.18 are consistent with the t test in Exercise 12.17. The t test showed that these children showed lower levels of anxiety than the normal population, and the confidence interval did not include 14.55.