Lecture 4

Assumptions

In this lecture we will discuss the assumptions made in the development of the t-test, tools available for verifying those assumptions on data, and the relative importance of each assumption. The assumptions are: normality, equal variances for the pooled version of the test, and independence.

Normality

You can make a reasonable assessment about the normality of the data by creating a normal probability plot. This graph plots the data against their corresponding percentiles in a normal distribution. For instance, for the 7 data points for the self-esteem treatment, Minitab ranks them, plots the smallest against the $(0.5)/8^{th}$ percentile of the standard normal distribution, the second against the $(1.5)/8^{th}$ percentile of the standard normal distribution, and so on with the largest being plotted against the $(7.5)/8^{th}$ percentile. Using the standard normal distribution with mean 0 and variance 1 does not matter in the sense that any other normal distribution is just a linear transformation of this one and we ultimately want to assess whether this graph is a line. If the graph is a line or close to a line, then the data is close to being normally distributed. One can also determine how the data fails to be normal by looking at this graph - see the figure below for examples. Normality plots can be created in Minitab using the Stat > Basic Statistics > Normality Test menu.

In general, t-tests are robust to the normality assumption. That is, if the data is not normal but is not too badly skewed, t-tests are still give approximately the correct assessment for the difference of two means. If the samples sizes are small, the data needs not to be very skewed. With increasing sample sizes, the t-test is robust against ever more skewness in the data.

Example 1 For the self-esteem data, we provide the results of the normal probability plots of the two data sets. There is no indication that data is not normal, but the sample sizes are very small and lack of normality would be hard to see in the data.

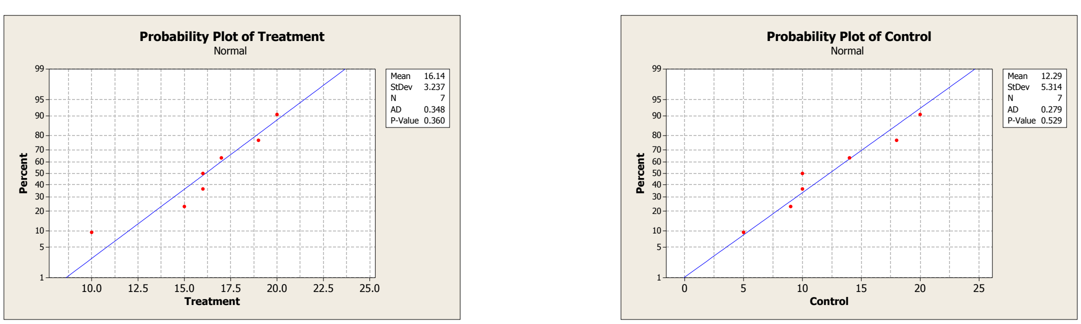

Figure 4.2: Normal probability plots for the self-esteem data.

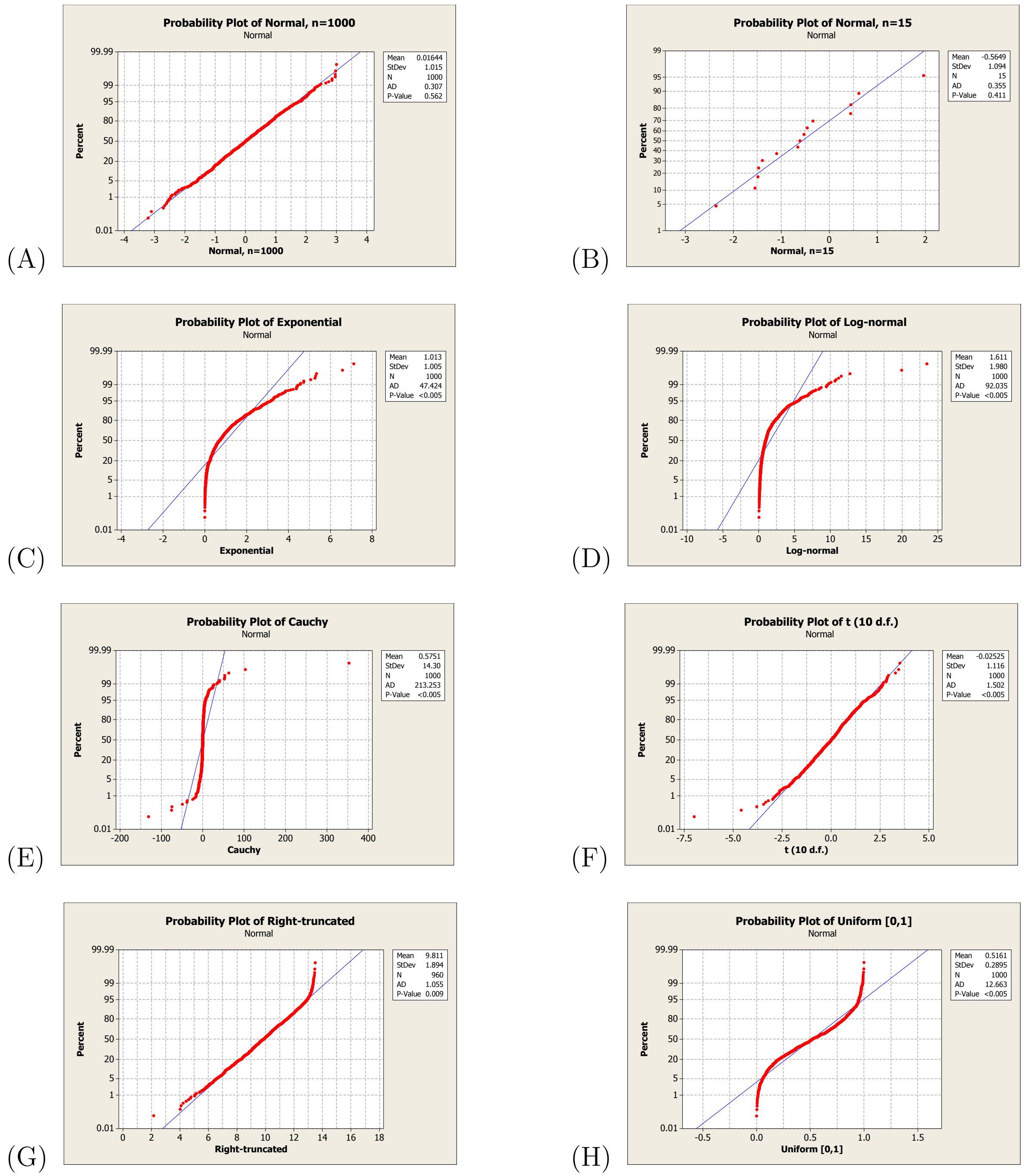

Figure 4.1: Normal probability plots for various samples samples. Plots (A) and (B) are samples from a normal distribution with a large ($n = 1000$) and small ($n = 15$) sample sizes respectively. Plots (C) and (D) are large samples from right-skewed distributions (exponential and log normal, respectively). Plots (E) and (F) are large samples from heavy-tailed distributions (Cauchy and a t-distribution with 10 degrees of freedom, respectively). Plots (G) and (H) are large samples from truncated distributions (a right-truncated normal distribution and the uniform distribution with bounds on both sides, respectively.)

Equal variances

This assumption is most critical for the pooled-variance t-test when the sample sizes are unequal, is less important when the sample sizes are equal (balanced data), and is irrelevant for Welch's t-test. There are two common tests of equal variance - Bartlett's F-test which looks at the ratio of the two variances, is strongly based on the assumption of normality, and is extremely sensitive (non-robust) to violations of that assumption and Levene's test which is the preferred, robust, alternative

The version of Levene's test programmed into Minitab is based on looking at the absolute deviations from the median. If $x_1, x_2, \ldots, x_{n_x}$ is the first sample and $\tilde{x}$ is its median and if If $y_1, y_2, \ldots, y_{n_y}$ is the first sample and $\tilde{y}$ is its median then Levene's test considers two new samples:

sample 1	sample 2
$\lvert x_1 - \tilde{x} \rvert$	$\lvert y_1 - \tilde{y} \rvert$
$\lvert x_2 - \tilde{x} \rvert$	$\lvert y_2 - \tilde{y} \rvert$
$\vdots$	$\vdots$
$\lvert x_{n_x} - \tilde{x} \rvert$	$\vdots$
	$\lvert y_{n_y} - \tilde{y} \rvert$

and tests whether the means of these samples are the same. If so, then the average deviation in the two samples are the same. If not, then the average deviations in the two samples are unequal.

The question that comes up is this: which test of equal means is appropriate for these absolute deviations. The answer is as follows: You ALWAYS form a test under the null hypothesis. The null hypothesis here is that the two (original) populations do have equal variances. Under this assumption the spread of the absolute deviations would be the same too. And so the pooled-variance t-test is the appropriate test.

Minitab provides the results for both Bartlett's test and Levene's test as will as a picture of the variation in two data sets from the Stat > Basic Statistics > 2 Variances menu.

Example 2 For the self-esteem data, we provide the results of the test for equal variances for the two data sets. There is no indication that the variances are unequal. (One can never prove that the variances ARE equal, so statisticians always couch their language and say "no evidence that they are unequal.")

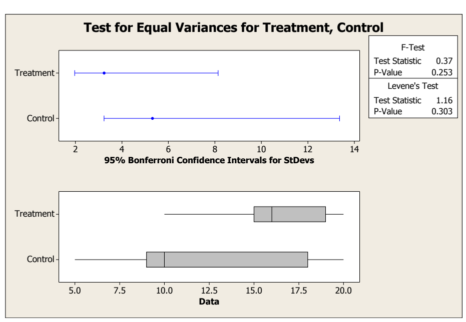

Figure 4.3: Tests for equal variances for the self-esteem data. ∎

Independence

To some extent, violations in the requirement that the data be independent can be assessed by examining the experimental design. If the results are the results from 20 patients where the assessment of the patients can reasonably be considered independent, then there is no reason to test for independence. If the data is given in an arbitrary order, then there is no way to test for independence. On the other hand, if the data are 20 measurements in order from one machine that requires calibration and is known to get off-calibration during use, then it is reasonable to look for violations of the independence assumption by looking for serial correlation in the data.

One way to do this in Minitab is by using the Stat > Nonparametrics > Runs Test menu.

Example 3 For the self-esteem data, the data has clearly been sorted. Thus, the original order has been lost. Further, the results are from individual test scores. Independence is a reasonable assumption and its violation cannot be tested in that data set. ■

Exercises for Lecture 4

1. –

2. –