Frances Wu
22S: 30 Spring 2004

Statistical Analysis of Data: Detection Thresholds for 0.025 Contrast Level

When collecting data, statistical analysis is necessary to determine the significance of the data. This is used to establish whether or not the data holds for a certain hypothesis. For example, data is being collected in an observational study for Ergonomics 56:147 and must undergo statistical analysis. This study, "Detection Thresholds for Rectangular Targets with Different Aspect Ratios and Orientations," is trying to determine if a difference exists for the detection of stimuli with different contrasts, lengths, widths, and orientations. The research question proposed in this paper is: Given two lines with the same contrast level, is there an effect on the detection threshold if the overall luminance is higher or lower?

Data collection for the study is an ongoing process and has only resulted in eight sets of data at this point. The testing procedure calls for an individual to walk towards a computer screen that is presenting a line at a very low contrast. At the point where the subject can identify the orientation (horizontal, vertical, or diagonal), the distance from the screen is recorded. A total of 81 runs are given from a combination of three widths, three angle orientations, and nine combinations of background and target luminance levels. The data has been dramatically cut back for use in this statistical analysis to only two combinations of variables. Two lines were chosen from the test to be used as the two treatments and the distances from which each subject detected the line as the

response variable. The lines are both horizontal and have a contrast level of 0.025. One set of data is for the line with a background luminance (BG Lum) of 12 cd/m^2 and a target luminance (TG Lum) of 12.3 cd/m^2. The other line has a background luminance of 4 cd/m^2 and a target luminance of 4.1 cd/m^2. The contrast is determined by the equation:

$$Contrast = \frac{BG\,Lum - TG\,Lum}{BG\,Lum} \quad \text{(Equation 1.1)}$$

By using Equation 1.1, the contrast for both lines is determined to be 0.025. The eight sets of data collected are presented in Table 1.1.

TABLE 1.1: Collected Data for 0.025 Contrast Levels

	BG Lum 12 (ft)	BG Lum 4 (ft)
Subject 1	13.50	4.00
Subject 2	3.50	6.75
Subject 3	7.50	1.00
Subject 4	10.00	1.75
Subject 5	12.75	9.25
Subject 6	1.25	5.50
Subject 7	10.00	2.75
Subject 8	4.00	1.75

This is the raw data that was used for further statistical analysis. For simplicity, the distance observations will be further referred to as BG Lum 12 and BG Lum 4 for background luminance 12 cd/m^2 and background luminance 4 cd/m^2, respectively.

From a comparison of a variety of statistical tests such as z-tests, t-tests, and X^2 tests, a paired t-test was chosen to be the best to analyze the data. The data is self-paired because the measurements for BG Lum 12 and BG Lum 4 were

taken for each subject. t-tests allow for inference about the mean without knowing the population standard of deviation. The paired t-test further allows for a comparison of the difference in the distances recorded for each subject. This is beneficial because eyesight abilities vary greatly for each individual and the detection at an overall brighter luminance can be compared with a lower luminance target. The difference was calculated by subtracting the distance for BG Lum 4 from the distance for BG Lum 12. The mean of these distances were then used for calculating the t-statistic. A two-sided test is completed because it is not initially evident whether detection is easier at a higher or lower luminance level. The null hypothesis used was that there exists no difference in the detection thresholds for the BG Lum 12 and BG Lum 4.

$$H_0 : \mu_{BG\ Lum\,12} = \mu_{BG\ Lum\,4}$$
$$H_a : \mu_{BG\ Lum\,12} \neq \mu_{BG\ Lum\,4}$$

The robustness for t-tests does not hold for data sets that contain outliers. The effects of outliers are enhanced even more when the sample size is small. By the central limit theorem, distribution approaches Normal for sample means as the sample size increases. In this case, the sample size is only 8. More data is undeniably needed for this analysis but is not available at this time due to the concurrent progress for both projects. A graphical representation of the data collected is shown on the following page in Figure 1.1. As a rule of thumb, t-tests should only be conducted for sample sizes less than 15 if the data is near Normal, contains no outliers, and is not skewed. For sample sizes greater than 15, the test still cannot be used with outliers. Because more data is not available at this time, t-tests were still conducted as the best test for this situation.