Xingyu Zhao
4/28/2024

Data Conversion:

The raw data provided is a simple 2D array that gives the angle of refraction in degrees. Each row is a series of measured refraction angles for a set angle of incidence, ranging from 10 to 60 degrees. Unfortunately these values themselves are not included in a separate column of the file, requiring the analysis program to add these itself. When generating the provided histograms for each angle of incidence, a for-loop is used to collect refraction data from only one specific row at a time, allowing the program to generate 6 histograms for the 6 rows of data. The histograms are a visualization of the expected refraction distribution for a specific angle of incidence. In order to compare this data with Snell's Law however, degrees must be converted to radians. This is accomplished by simply multiplying each given degree measurement by $\boldsymbol{\pi}/180$.

Statistics Preparation:

To plot the line of best fit, we first need to generate a scatter plot of the given data. While a program that plots every single measured data value for all 6 angles of incidence could theoretically be made, this process is tedious. A simpler and just-as-effective method is to find the average angle of refraction for each angle of incidence. Using a structure similar to the for-loop used in generating row-specific histograms, such an average can be taken, resulting in 6 data points in the format *(incidence, avg refraction)*. The uncertainty about each point's refraction must then be calculated. While a normal scatter plot would simply use the standard deviation as the uncertainty, we cannot do the same - the refraction values are averaged, not raw data. Instead, we take the standard deviation of the mean, which is done by dividing the standard deviation of the raw data by the square root of the number of data points. In our case, we have 16 data points for each angle of incidence. After taking the mean and standard deviation of these 16 points, we divide the standard deviation by 4. The resulting value is thus the uncertainty, and the mean is the y-value of the data point. This process is done for each angle of incidence.

Generating a Line of Best Fit:

Snell's Law dictates a proportional relationship between the sine of angles of incidence and refraction, with a proportion constant n serving as the relative index of refraction between the two mediums being refracted through. By plotting the sine of incidence against the sine of refraction, the equation can be generalized to a linear $y = x/m$ form, where m is the angle of incidence. Using this information, we can perform a simple one-parameter chi-squared fit to the sine of the provided data (the data given is in angle form, we need the sine of that angle). Our program accomplishes this by testing different variables of m and computing the chi-squared value between the resulting plot and the data, attempting to minimize the resulting value. Here, the chi-squared value is a measure of the correlation between a model and data, essentially describing how much of the variance between the two can be explained by random uncertainty. Our resulting index of refraction is 1.48, with an uncertainty of 0.019. The associated chi-squared value is 0.44, meaning that 44% of the variance in data can be explained by this uncertainty. In addition, the p-value of this fit is 0.99, indicating that 99% of the supplied data is a good fit for our predicted index.

When the line of best fit is plotted using the obtained index of refraction, we see the plot lines up very well with the measured data points. To visualize the minimization of chi-squared and further justify this index, we can plot the chi-squared value itself as a function of refraction index, essentially creating a graph of what our program has done to find this index. A solid vertical line is then placed to mark the predicted refraction index of 1.48, with dashed vertical lines of 1.48 +/- the uncertainty. Dashed horizontal lines are placed at the chi-squared plot's minima and +/- sigma excursions occur. The minima horizontal line intersects the solid vertical line on top of the plot, indicating that the predicted refraction index is A: on the plot and B: the smallest possible chi-squared obtainable, justifying the index of 1.48. In addition, the second horizontal line intersects with the vertical dashed lines on top of the graph as well, indicating that the uncertainty in the refraction index is also valid.