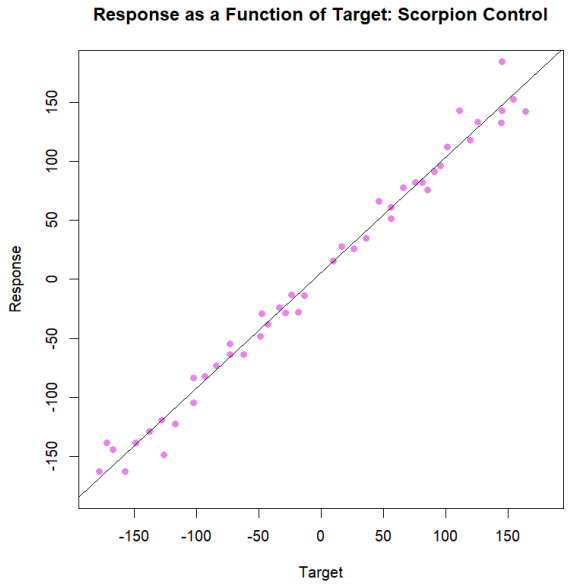
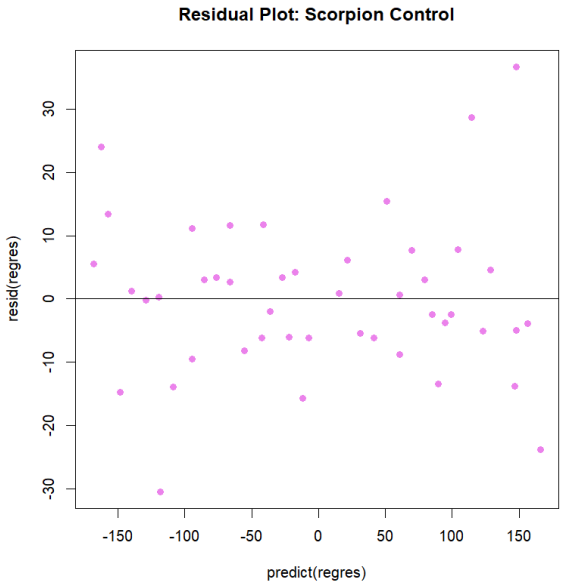
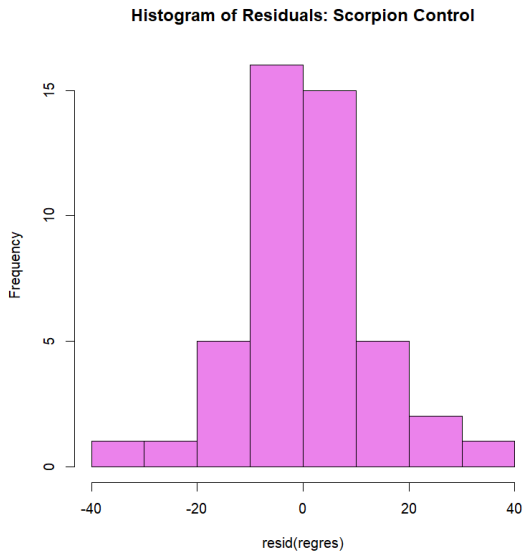
Gabriella Weis
Mr. Lewis
H Applied Statistics
4 November 2023

#12 Data Exploration ll: The Cockroach and the Scorpion

1). Scatter plot, line of best fit, residual plot, and histogram of residuals for control group:







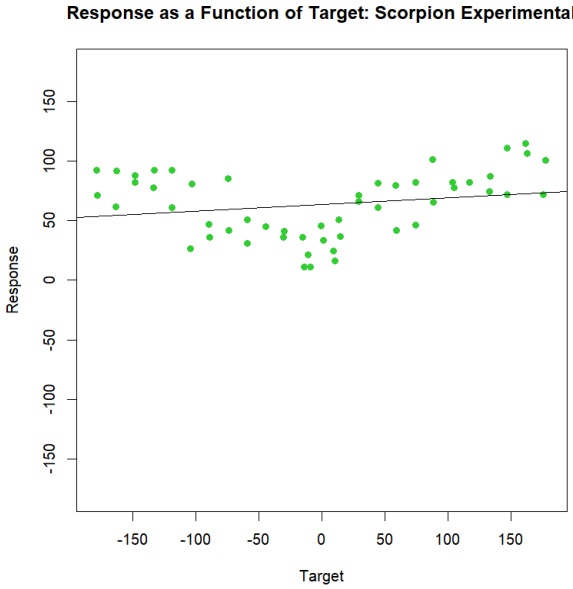
Line of best fit: $\widehat{y} = 5.64927 + 0.97967x$

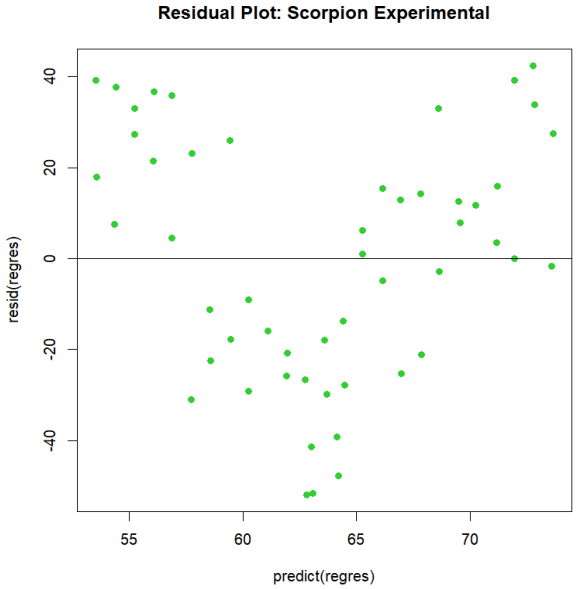
2). The scatter plot has a strong, positive, linear correlation, which is also demonstrated by its line of best fit and coefficient of determination ($R^2 = 0.9853$). The residuals have a mean of 3.451345×10^{-16}, or approximately zero, and a standard deviation of 12.3477. They appear to be randomly distributed above and below the y=0 line on the residual plot, if slightly more scattered than usual on the extremes of the plot, and also appear to be approximately normally distributed on the histogram.

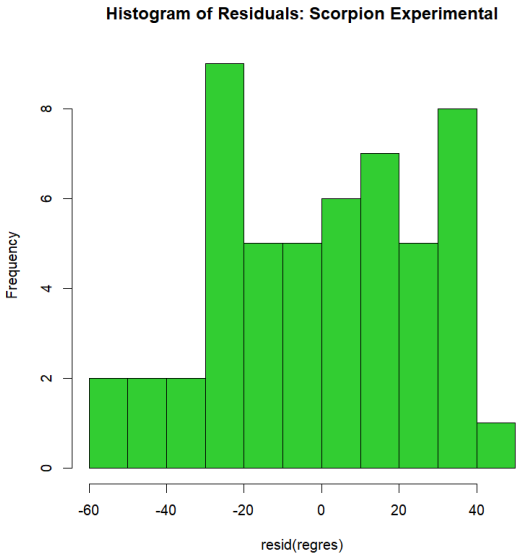
3). The response variable is very accurate according to the regression equation, correlation, and standard deviation of the residuals (SE). The regression equation ($\widehat{y} = 5.64927 + 0.97967x$), possesses a correlation coefficient of 0.9926, signifying a good fit to the data. The standard deviation of residuals (12.3477) also signifies a good fit because it's small in comparison to the data values, meaning the true response differs little from the response predicted by the very accurate regression equation.

4). As the target angle moves farther away from zero, the accuracy of the response variable decreases. Response variable accuracy is at its best when the target angle is closer to zero and at its worst when the target angle is closer to -180 or 180 degrees. I compared the deviation of the points on the scatter plot to the regression line to identify where predictions were the worst and therefore the responses less accurate.

5). Scatter plot, line of best fit, residual plot, and histogram of residuals for experimental group:





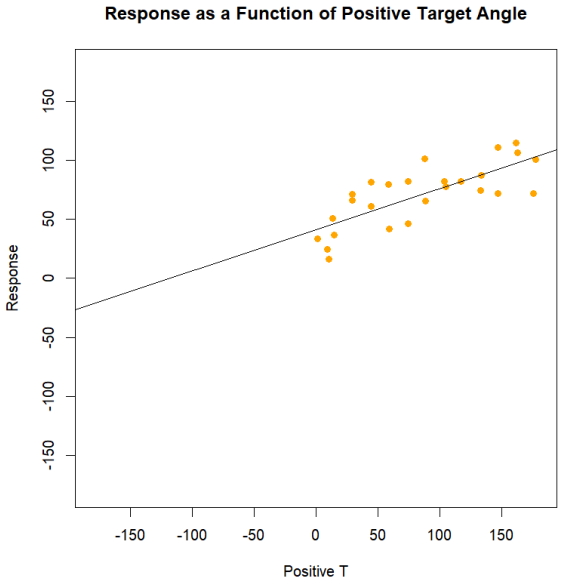


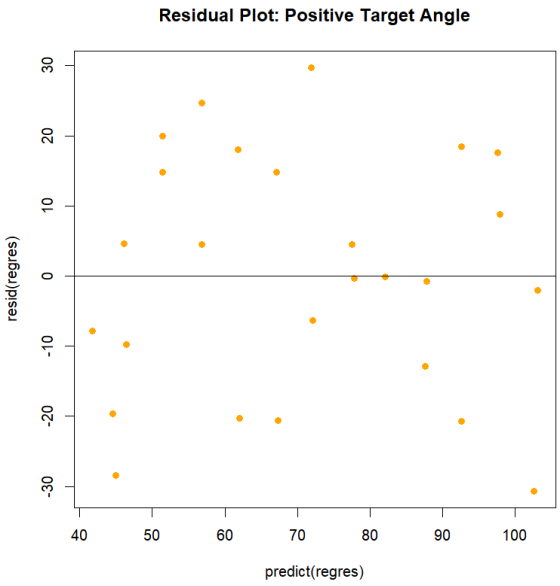
Line of best fit: $\hat{y} = 58.94979 - 0.013131x$

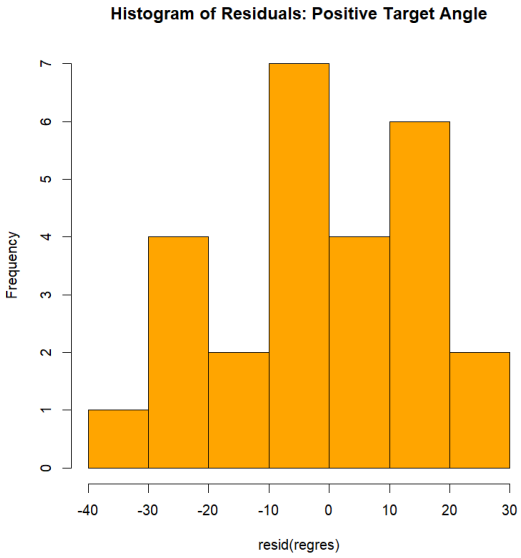
6). The scatter plot has a weak, positive, nonlinear correlation, which is also demonstrated by its line of best fit and coefficient of determination ($R^2 = 0.04691$). The residuals have a mean of -2.490396×10^{-15}, or approximately zero, and a standard deviation of 26.62233. They appear to be distributed disproportionately above and below the y=0 line on the residual plot, and do not appear normally distributed on the histogram either.

7).

a). Positive Target Angle Regression Analysis (angles from experimental group only):



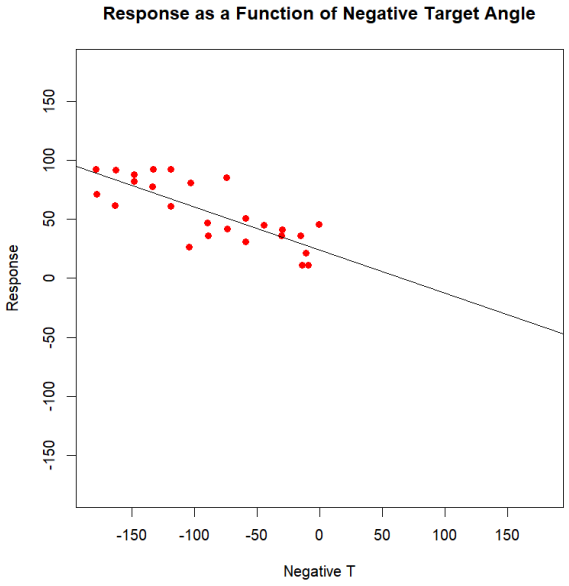


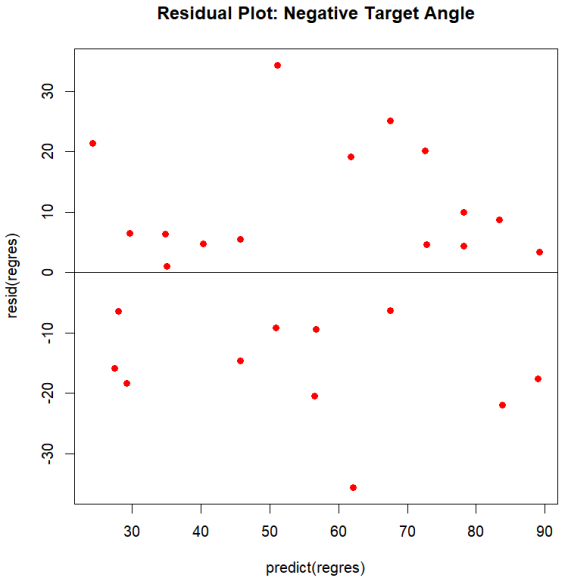


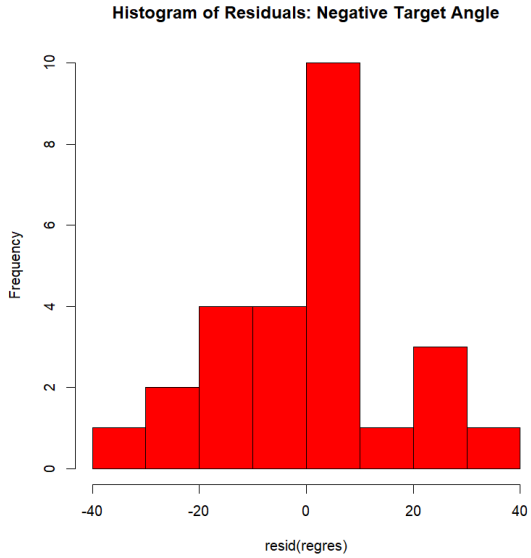
Line of best fit: $\widehat{y} = 39.46513 + 0.38586x$

The scatter plot has a moderate, positive, linear correlation, which is also demonstrated by its line of best fit and coefficient of determination ($R^2 = 0.5806$). The residuals have a mean of 5.214846×10^{-16}, or approximately zero, and a standard deviation of 16.98056. They appear to be relatively randomly distributed above and below the y=0 line on the residual plot, if slightly less varied and closer to the y=0 line between the predicted values (x-axis) 70 and 90, and also appear relatively randomly distributed on the histogram.

b). Negative Target Angle Regression Analysis (angles from experimental group only):







Line of best fit: $\hat{y} = 24.00207 - 0.36573x$

The scatter plot has a moderate, negative, linear correlation, which is also demonstrated by its line of best fit and coefficient of correlation ($R^2 = 0.6139$). The residuals have a mean of $-4.440892 \times 10^{-16}$, or approximately zero, and a standard deviation of 16.65133. They appear to be relatively randomly distributed above and below the y=0 line, and also appear to be relatively randomly distributed on the histogram.

i). Though the regression lines are not approximately the same due to slope, the standard deviations and means of their residuals are very similar, as are their distributions in general.

ii). The residuals for the positive target angles of the experimental group are significantly greater than those of the control group, and the residuals for the negative target angles of the experimental group are significantly greater than those of the control group as well. Through the difference in residuals between the two groups, it is evident that the immobilization of the left BSS severely lessened the accuracy of the scorpions' ability to detect prey.

8). In conclusion, the scorpions' BBSs play a large role in prey detection. A quick glance at the data between the control group and the experimental group makes this observation obvious; with both BBSs, a scorpion's turns are far more accurate. But a deeper dive reveals that in the experimental group, wherein the scorpions' left BBSs are immobilized, the scorpions only ever turn right. Below is a scatter plot comparing what the negative target angle values should look like using the control group as a baseline (below the y=0 line), and what they look like in the experimental group (above the y=0 line). Notice that all of the experimental group's negative target angles have a positive response. This is because the scorpions are only turning right, even when the target angle is negative. Remember once again that in the experimental group, the scorpions' left BBS is inhibited. With this we can conclude that to a scorpion, BBSs are the difference between life and death; without the use of them on either leg, a scorpion would be unable to turn correctly, if at all, and would therefore die of starvation very quickly.