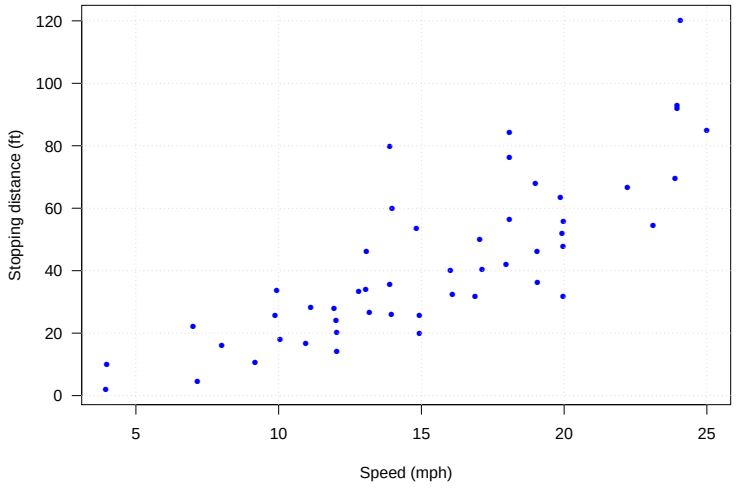
STAT 8020 Statistical Methods II
Practice Exam I

Instructor: Whitney Huang (wkhuang@clemson.edu)

September 20, 2019

Problem 1

A researcher is interested in the relationship between the speed of cars (`speed`) and the distances taken to stop (`dist`). She performed an experimental study (way back in 1920) and the data set is presented in the scatterplot below.



1. Is linear relationship between `speed` and `dist` a reasonable assumption?
2. Let's use x to denote `speed` and use Y to denote `dist`. Write the form of the corresponding simple linear regression.
3. Use the fact that $\sum_{i=1}^{n=50}(X_i-\bar{X})(Y_i-\bar{Y}) = 5384.40$, $\sum_{i=1}^{n=50}(X_i-\bar{X})^2 = 1370.51$, $\bar{X} = 15.40$, and $\bar{Y} = 42.99$ to compute the estimated slope $\hat{\beta}_1$ and intercept $\hat{\beta}_0$
4. Write down the least squares regression line and compute the fitted value with `speed` $= 15$mph.

5. Using the information SSE $= \sum_{i=1}^{50}(Y_i - \hat{Y}_i)^2 = 11362.39$ to compute $\hat{\sigma}$
6. Construct the 95% confidence interval for β_1
7. Test the following hypothesis: $H_0 : \beta_1 = 0$ vs. $H_a : \beta_1 \neq 0$ with $\alpha = 0.05$. You could make use the confidence interval from (6). State your conclusion in plain language in the present context.
8. Construct the 90% prediction interval for a future observation of `dist` with `speed` = 20mph.
9. Fill in the missing values in the ANOVA table below:

Source	df	SS	MS	F
Model	?	SSR =?	MSR = ?	F=?
Error	?	SSE = 11362.39	MSE = ?	
Total	?	SST = 32516.40		

10. Do the residual plot and the Normal Q-Q plot below suggest any regression assumptions may be violated? Explain your answer.