
COURSE: MATH 50 DARTMOUTH COLLEGE (MWF 11:15 AM-12:20 PM), FALL 2015
INSTRUCTOR: NISHANT MALIK
HOMEWORK SHEET NUMBER: 6
POSTED ON: 10/23/2015
DUE ON: 10/30/2015

Directions: Any problem marked with asterisk (*) should be completed using IPython Notebook (Jupyter) and can be uploaded at <https://dropitto.me/m50f15> .

1*. **Contract profitability.** A cost analyst for a drilling and blasting examined 84 contracts handled in the last two years and found that the coefficient of correlation between value of contract (Y_1) and profit contribution generated by the contract (Y_2) is $r_{12} = .61$. Assume that bivariate normal model applies.

(a). Test whether or not Y_1 , and Y_2 are statistically independent in the population; use $\alpha = .05$. State the alternatives, decision rule, and conclusion.

(b). Estimate ρ_{12} with a 95 percent confidence interval.

(c). Convert the confidence interval in part (b) to a 95 percent confidence interval for ρ_{12}^2 . Interpret this interval estimate.

Reference: *Kutner et. al. "Applied Linear Regression Models", Ed. 5 (problem no 2.43 on pages 96).*

2*. Refer to **Crime rate** data:

source: <https://netfiles.umn.edu/users/nacht001/www/nachtsheim/Kutner/Chapter%20%201%20Data%20Sets/CH01PR28.txt>

**This data can also be accessed using the function `read_tb_data(ch,pr)` with `ch=1` and `pr=28`.*

Description of the data : Column 1 is the crime rate (crimes reported per 100,000 residents). Column 2 is the percentage of individuals in the county having at least a high-school diploma.

Assume that the normal bivariate model is appropriate.

(a). Compute the Pearson product-moment correlation coefficient r_{12} .

(b). Test whether crime rate and percentage of high school graduates are statistically independent in the population; use $\alpha = .01$. State the alternatives, decision rule, and conclusion.

Reference: *Kutner et. al. "Applied Linear Regression Models", Ed. 5 (problem no 2.48 on pages 97).*

3*. Refer to **Grade point average** data:

Data source: <https://netfiles.umn.edu/users/nacht001/www/nachtsheim/Kutner/Chapter%20%201%20Data%20Sets/CH01PR19.txt>

**This data can also be accessed using the function `read_tb_data(ch,pr)` with $ch=1$ and $pr=19$.*

Description of the data : The director of admissions of a small college selected 120 students at random from the new freshman class in a study to determine whether a student's grade point average (GPA) at the end of the freshman year (Y) can be predicted from the ACT test score (X).

- (a). Prepare a box plot for the ACT scores X_i . Are there any noteworthy features in this plot?
- (b). Prepare a dot plot of the residuals. What information does this plot provide?
- (c). Plot the residual e_i against the fitted values \hat{Y}_i . What departures from normal error regression model can be studied from this plot? What are your findings?
- (d). Prepare a normal probability plot of the residuals. Also obtain the coefficient of correlation between the ordered residuals and their expected values under normality. Test the reasonableness of the normality assumption here with $\alpha = .05$. What do you conclude?
- (e). Conduct the Brown-Forsythe test to determine whether or not the error variance varies with the level of X . Divide the data into the two groups, $X < 26$, $X \geq 26$, and use $\alpha = .01$. State the decision rule and conclusion. Does your conclusion support your preliminary findings in part (c)?

Reference: *Kutner et. al. "Applied Linear Regression Models", Ed. 5 (problem no 3.3 on pages 146).*