Question 1. Concrete

An article on *Concrete Research* from 1989 presented data on compressive strength x and intrinsic permeability y of various concrete mixes and cures. Summary quantities are:

$$n = 15, \sum_{i=1}^{15} y_i = 570, \sum_{i=1}^{15} y_i^2 = 22, \sum_{i=1}^{15} x_i = 45, \sum_{i=1}^{15} x_i^2 = 155, \sum_{i=1}^{15} x_i y_i = 1691.$$

Assume that permeability is linearly related to compressive strength.

(a) Calculate the least squares estimates of the slope and intercept.

(b) Use the equation of the fitted line to predict what permeability would be observed when the compressive strength is $x = 41$.

Question 2. Renewable Energy

The file *"A6-2.csv"* contains information on renewable energy in US States published by the U.S. Energy Information Administration, available on
https://dasl.datadescription.com/datafile/alternative-energy-2016/?_sfm_cases=4+59943&sf_paged=2.
The column "*Ren.Elec.GW.h.*" refers to the percentage of renewable electricity in Gigawatt hours and the column "*Pct.Renewable.incl.Hydro*" refers to the percentage of renewable energy with Hydropower.

(a) Assuming that a simple linear regression model is appropriate, use *R* to obtain the least squares fit estimators relating "*Pct.Renewable.incl.Hydro*" to "*Ren.Elec.GW.h.*".

(b) Plot the data points in a scatter plot and add your linear regression curve. Comment on the appropriateness of the model.

Question 3. Blood Pressure

The data set "*A6-3.csv*" contains the blood pressure (*BP*) and weight (*Weight*) of 20 individuals.

(a) Plot a scatter diagram of the data. Does the straight-line regression model seem to be plausible?

(b) Calculate the error sum of squares, commonly denoted by SS_E. Then use this value to estimate the variance σ^2.

Question 4. Regression without the Intercept Term

Assume that we have n pairs of data $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$.

(a) Suppose that the appropriate model is $Y = \beta x + \epsilon$ (no intercept). Provide an equation to estimate β.

(b) Do you suspect the model $Y = \beta x + \epsilon$ to fit better or worse than $Y = \beta_1 x + \beta_0 + \epsilon$ to a general data set? Explain briefly.

Question 5. Intrinsically Linear

Decide which of the following relations between $Y > 0$ and $x > 0$ are intrinsically linear, where ϵ is a random variable (not necessarily Gaussian). If they are intrinsically linear, provide the function that transforms the equation into a linear relation.

(a) $Y = \frac{\beta_0}{\beta_1 x + \beta_2 + \beta_0 \epsilon}$

(b) $Y = \left(e^{\beta_1 x + \beta_2 + \epsilon}\right) \beta_0$

Question 6. Water Vapor Pressure

The file "*A6-6.csv*" contains the temperature (K) and vapor pressure (mm Hg) of 11 samples.

(a) Plot a scatter diagram of the data. What type of relation between the temperature and vapor pressure do you suspect?

(b) Use an appropriate transformation to fit a linear model to the (transformed) data, relating (transformed) vapor pressure to the (transformed) temperature. Clearly state the transformation you applied as well as the resulting least square estimates $\hat{\beta}_0, \hat{\beta}_1$.

Question 7. t-Test for Regression Models

Consider the following data on the number of pounds of steam (y) used by a chemical plant and the average temperature (x) in Fahrenheit.

Temp	21	24	32	47	50	59
Usage	185.79	214.47	288.03	424.84	454.58	539.03
Temp	68	74	62	50	41	30
Usage	621.5	675.06	562.03	452.93	369.95	273.98

Test the hypothesis $H_0 : \beta_1 = 0$ against $H_1 : \beta_1 \neq 0$ using the t-test with $\alpha = 0.05$.

Question 8. Beauty of a Proof II

Given observations $(y_1, y_2, \ldots, y_n)$ and their predictions $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ $(i = 1, 2, \ldots, n)$, where x_i are observed variables $i = 1, \ldots, n$, $\hat{\beta}_0$ is the least square estimate of the intercept and $\hat{\beta}_1$ is the least square estimate of the slope.

Show that

$$\sum_{i=1}^{n} (y_i - \hat{y}_i) = 0.$$

(Hint: Use the structure of $\hat{y}_i$ and recall the equations for $\hat{\beta}_0$ and $\hat{\beta}_1$.)

Thank you for a great semester!
Thank you for your feedback to improve the STAT2201 lectures and my teaching style, I appreciate it very much.