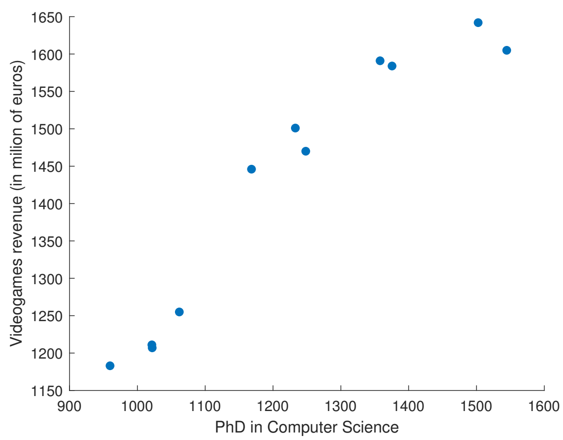
Marco Longfils and Petter Mostad
Applied Mathematics and Statistics
Chalmers and GU

MVE051 / MSG810 Matematisk statistik och diskret matematik

Exam 6 April 2018, 8:30 - 12:30

Allowed aids: Chalmers-approved calculator
and one (two-sided) A4 sheet of paper with your own notes.
Total number of points: 30. To pass, at least 12 points are needed.
Note: All answers should be motivated.

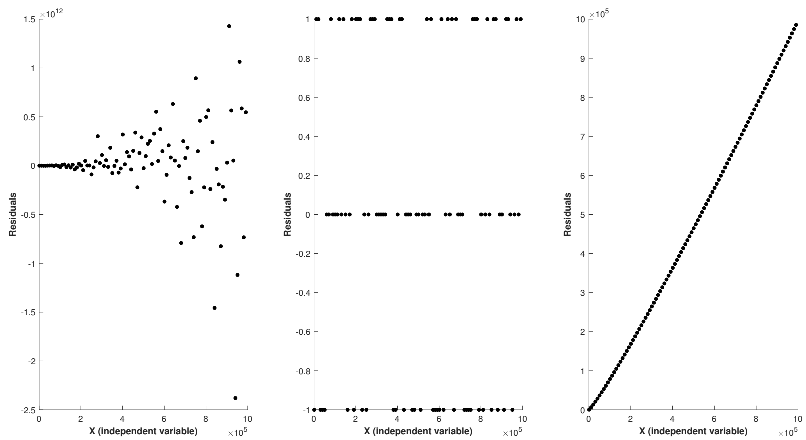
1. **(5 points)** Consider the scatter plot in the figure below, where the number of PhDs awarded in Computer Science in one year in Sweden is shown on the x-axis against revenues from sales of videogames for the same year on the y-axis.



   (a) The sample correlation coefficient is $r = 0.956$. What would be the most reasonable explanation?
   - It is well known that computer scientists like videogames. Thus, the more PhDs are awarded in Computer Science the more videogames will be sold.
   - Playing videogames is likely to cause many students to start a PhD in Computer Science.
   - Both variables are affected by one or more factors not displayed in the graph above.
   - None of the above.

   (b) After performing a linear regression analysis, you should look at the residuals for model diagnostics.

- State the assumptions underlying the linear regression analysis model.
- For each of the residual plots below, state and motivate which of the assumptions it indicates that the data fails to satisfy.



2. **(5 points)** Let X be a discrete random variable which takes values in $\{-2, -1, 0, 1, 2\}$:

   (a) What are the maximum and minimum possible values for $E[X]$? Provide an example of a random variable X as above such that $E[X]$ attains the maximum and mininum possible values (i.e., give two examples, one for the maximum and one for the minimum value).

   (b) What are the maximum and minimum possible values for $Var(X)$? Provide an example of a random variable X as above such that $Var(X)$ attains the maximum and mininum possible values (i.e., give two examples, one for the maximum and one for the minimum value).

   Remember to motivate your answers.

3. **5 points** A student will take the exam of the course MVE051. The night before the exam, he can join his friends to the Pub crawl. If he does go to the party, he has a 99% probability of passing the exam, otherwise his passing probability is reduced to 50%. Since the student can't make up his mind, he decides he will throw a fair coin. If the coin lands on heads, he will go to the Pub crawl.

   If the student has passed the exam, what is the probability that he joined the party?

4. **5 points** Let $\{X_n\}$ be a sequence of independent and identically distributed random variables such that $P(X_n = 1) = 1 - P(X_n = 0) = p$. Define $Y_n = \max\{X_1, ..., X_n\}$.

   (a) Does $\{Y_n\}$ form a Markov chain? If so, provide its transition matrix.

   (b) Compute $P(Y_n = k)$ for $k = 0, 1$.

5. **5 points** Let $X_1, ..., X_n$ be a sequence of independent and identically distributed random variables with mean μ, variance σ^2, and $w_1, ..., w_n \in [0, 1]$ a sequence of weights. Define the estimator $S_n = \sum_{i=1}^{n} w_i X_i$.

   (a) Find the condition on the weights $w_1, ..., w_n$ that ensures S_n is an unbiased estimator of μ.

   (b) Consider now $n = 2$. Find the estimator S_2 with minimum variance among those estimators S_2 that are unbiased.

6. **5 points** Peter is working in tech-support. During the last 50 working hours, he received 113 calls with questions or requests.

   (a) Lisa would like to make predictions about the number of calls Peter will receive in the coming working hours. What are the simplest possible (yet reasonably realistic) assumptions that would enable Lisa to make such predictions, if all the information she has is the statement above?

   (b) Making these assumptions, what is the probability that during the next working hour Peter will receive exactly 2 calls?

   (c) Making these assumptions, what is the (approximate) probability that during the next 50 working hours, Peter will receive more than 120 calls?