B. Exercises

You should be prepared to present the exercises and at least an attempt to solve them. A complete solution is not required but you should at least have read, understood and attempted to solve them before going to the lecture.

For February 27

22. A person has drawn a SRS-sample with $n=10\ \text{from a population with }N=50.$ Unfortunately one unit becomes partial nonresponse for one of the variables . The observations are

- Y 112, 123, -, 122, 147, 110, 125, 123, 133, 111 X 35, 58, 27, 67, 119, 2, 55, 79, 107, 23
- a) The person does not worry about the non response, but makes the analysis as if he had a SRS with n=9, What will his estimate of the mean for Y_3 be and what will the variance of his estimate be?
- b) Use the theory of linear regression and find the predictive distribution of Y for the two missing values.
- c) The person reads about multiple imputation and he draws a new value from the predictive distribution of Y_3 . He inserts this and makes the analysis as if he had had a complete data set. In his draw he got $Y^*_3 = 120$. What will his mean and variance be?
- d) Repeat this nine more times. And gets nine more means and variances. Do this. Compute the mean of the ten means and the ten variances. Compute the variance of the ten means. Add the variance of the means and the mean of the variances. Compare these with the value that you found in a).

23. In a medical study a population of 12 patients are treated for a special type of heart failure, with a new method. All persons are before treatment classified into five categories after the severity. Half a year after treatment the patients are contacted once again and among other things the stability of the heart rhythm is measured. However, only 9 patients were reached.

Data:	
Severity	Stability measurements
1	23, missing
2	15, 24, 12, missing
3	16, 21, 10
4	8, 13, missing
5	None

Your task is to do a multiple imputation.

- a) Do a linear regression explaining the stability by the severity. Estimate the variance and covariance of the two parameter estimators, slope and intercept. (The variance around the line is from previous trials known to be 5)
- b) Draw a pair of possible parameter values assuming that the true value is normal around the estimates with the variances/covariance from a). (Assume normality)
- c) Draw three possible stability values for the missing assuming normality and that the parameters from b) are correct. Impute these values instead of the missing
- d) Estimate the average stability and its variance for an infinite population assuming that the true values are the true values.
- e) Repeat b-d four more times. You have now five mean value estimates and five variance estimates.
- f) Find the imputation variance by finding the variance of the five repeated imputations.
- g) Find the final estimate as the mean of the five repetitions and the final variance as the sum of the average variance from the repetitions plus the imputation variance.