Exercises for MSA830, Monday 3rd December 2007

1. Assume that we want to compare two medications used to reduce blood pressure, to see if one reduces blood pressure more than the other. We assume that both reduce blood pressure by a normally distributed amount, and that the standard deviations in these two normal distributions are the same. Assume that a pilot study gives us reason to believe that this standard deviation is around 14 mmHg. Assume that the smallest difference in blood pressure reducing capability that we see as clinically relevant is 10 mmHg. Finally, assume we are going to compare the two medications in a randomized trial where each medication is tried out on n patients. We want to reject the null hypothesis that the two medications have the same capabilities in favour of the alternative that they are different with 90% probability if the actual difference in capabilities is 10 mmHg. How many patients in each group should we include in the study? (Additional question: Can you imagine alternative experimental plans to the one suggested? Discuss possible advantages and disadvantages with various choices of experimental plans)

2. How would the answer to the question above be changed if the standard deviations for the two treatments is assumed to be 10 mmHg? How would it be changed if we would like to have a 95% probability of rejection in the situation above? (A 95% power?)

3. Assume we want to estimate the proportion of Swedes who used an anti-depressive medicine in the last week, by asking n randomly selected persons. If we want to have a proportion with a confidence interval of length at most 10% (so that we could give the proportion as $x \pm 5\%$) how many people would we need to ask? (Additional question: How do we make the selection of n ”randomly selected persons”? Are there things we need to consider in order to avoid getting a bias in our answer?)

4. You are investigating the yield of some process producing a chemical, and you need to find the average yield in a run, with a confidence interval of size ± 3. You start by running the process 10 times, and from this you estimate that the standard deviation of the yield to be 10. How many additional times do you need to run the process in order to get a confidence interval of the required size for the average yield? (Additional question: Your computation depends on the assumption that each run of the process produces a yield that is independent from the yields preceding it. How can you check if this seems to be the case?)