Tommy Norberg
November 15, 2011

HA3) Quality Improvement and Process Control

1. Consider the tensile strength experiment (Section 4.5.1 and Example 4.12). The file `HardwoodANOVA.m` contains Matlab code for doing an analysis of variance of the psi-data in Table 4.4. The model in the ANOVA is

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad \text{for} \quad \begin{cases} i = 1, \ldots, a = 4 \\ j = 1, \ldots, n = 6 \end{cases}$$

   where μ is the overall mean, τ_i is the ith treatment effect, $\epsilon_{ij} \sim \mathrm{N}(0, \sigma^2)$. One of the more crucial prerequisits for the ANOVA is that the variance is the same at each level of the factor. According to the Minitab analysis displayed in Table 4.7, the sample standard deviations are 2.828, 2.805, 1.789 and 2.639 in the four sets of data. Is the fact that one of the four is quite distinct from the other an indication that the equal variance requirement is not fulfilled in this experiment? Explain.

2. In the lecture a possible continuation of the analys of the psi-data is discussed. The file `HardwoodRegression.m` contains Matlab code for doing a regression analysis of the model

$$y \sim \mathrm{N}(\beta_0 + \beta_1 x + \beta_2 x^2, \, \sigma^2)$$

   where y the tensile strength at hardwood concentration x. Familiarize yourself with the code, and

   (a) test at level 5%,

$$H_0 : \beta_2 = 0 \quad \text{vs} \quad H_1 : \beta_2 \neq 0$$

   (b) examine the contribution to the regression sum of squares for the 2nd variable x^2.

3. Control charts for the mean and standard deviation are maintained for an important quality characteristic. The sample size is $n = 6$; $\bar{x}$ and s are computed for each sample. After $m = 30$ samples, $\sum_i \bar{x}_i$ and $\sum_i s_i$ are calculated. Set up mean and standard deviation charts for these data. The process was in statistical control during this phase. Plot the mean and standard deviation of the m samples in the control charts. The data file `dataHA3.m` contains the results for the first m and the 10 further samples. Plot these new data in the control charts. Is the process in statistical control? If not, at which point in time is there a parameter shift?

4. Continuation of Problem 3. Calculate an unbiased estimator of the variance σ^2 based on the data from the first m samples. Regard this estimate as the true value, which we therefore denote by σ_0^2. Then, for each of the 10 subsequent samples, calculate the P-value of the test of

$$H_0 : \sigma = \sigma_0 \quad \text{vs} \quad H_1 : \sigma \neq \sigma_0$$

   and regard the process as being out of control, if the$P$-value is smaller than 0.0027. Compare with the s chart. Which quality control method, if any, seems to be the better?

Deadline for solutions to this assignment is Wednesday, November 23, 2011.

The Matlab script files to be used in Problems 1 and 2 are available on the web as

`http://www.math.chalmers.se/Stat/Grundutb/CTH/mve145/1112/HardwoodANOVA.m`

and

`http://www.math.chalmers.se/Stat/Grundutb/CTH/mve145/1112/HardwoodRegression.m`

Parameters and data to Problems 3 and 4 are published at the course web page

`http://www.math.chalmers.se/Stat/Grundutb/CTH/mve145/1112/`

in the files `dataHA3.pdf` and `dataHA3.m`.