

# Statistics for Engineering, 4C3/6C3

## Assignment 4

Kevin Dunn, kevin.dunn@mcmaster.ca

Due date: 27 February 2013

### Note: Assignment objectives

- Demonstrate ability to perform phase 1 and phase 2 for monitoring charts
- Understand and interpret Cp and Cpk values
- Build and use least squares models in R

### Question 1 [12]

The Kappa number is a widely used measurement in the pulp and paper industry. It can be measured on-line, and indicates the severity of chemical treatment that must be applied to a wood pulp to obtain a given level of whiteness (i.e. the pulp's bleachability). Data on the [website](#) contain the Kappa values from a pulp mill. Use the first 2000 data points to construct a Shewhart monitoring chart for the Kappa number. In general, you may use any subgroup size you like, but please use  $n = 5$  for ease of grading. Then use the remaining data as your phase 2 (testing) data.

Does the chart perform as expected, i.e. does it capture what you might consider to be unusual operation at the appropriate type-I error rate?

### Question 2 [5]

The following are real data of particle size from the most recent 20 shipments of a speciality powder, taken from the supplier's certificates of analysis:

50.9, 52.9, 51.6, 50.8, 54.6, 52.9, 53.1, 48.4, 51.6, 53.1, 53.8, 52.4, 53.1, 50.8, 54.6, 52.9, 50.0, 53.8, 54.6, 52.2

Calculate the supplier's capability, given their lower specification limit of  $45\ \mu\text{m}$  and their upper limit at  $59\ \mu\text{m}$ . Clearly state all assumptions you make during the calculations (median =  $52.9\ \mu\text{m}$ , sd =  $1.64\ \mu\text{m}$ )

How would you interpret this supplier's capability? Would you purchase your raw material from them if you were making a safety-critical product from this supplier (e.g. air-bags or medication or desiccants for a breaking mask)?

### Question 3 [7]

Plastic sheets are manufactured on your blown film line. The Cp value is 1.7. You sell the plastic sheets to your customers with specification of  $2\ \text{mm} \pm 0.4\ \text{mm}$ .

1. List three important assumptions you must make to interpret the Cp value.
2. What is the theoretical process standard deviation,  $\sigma$ ?
3. What would be the Shewhart chart limits for this system using subgroups of size  $n = 4$ ?

See the last slide in the course slides to see a visual illustration of this question.

**Question 4 [10 = 3 + 2 + 3 + 1 + 1]**

*Parts 1 and 2 from a previous midterm, 2010*

The production of low density polyethylene is carried out in long, thin pipes at high temperature and pressure (1.5 kilometres long, 50mm in diameter, 500 K, 2500 atmospheres). One quality measurement of the LDPE is its melt index. Laboratory measurements of the melt index can take between 2 to 4 hours. Being able to predict this melt index, in real time, allows for faster adjustment to process upsets, reducing the product's variability. There are many variables that are predictive of the melt index, but in this example we only use a temperature measurement that is measured along the reactor's length.

These are the data of temperature (K) and melt index (units of melt index are “grams per 10 minutes”).

<b>Temperature</b> = $T$ [Kelvin]	441	453	461	470	478	481	483	485	499	500	506	516
<b>Melt index</b> = $m$ [g per 10 mins]	9.3	6.6	6.6	7.0	6.1	3.5	2.2	3.6	2.9	3.6	4.2	3.5

The following calculations have already been performed:

- Number of samples,  $n = 12$
- Average temperature =  $\bar{T} = 481$  K
- Average melt index,  $\bar{m} = 4.925$  g per 10 minutes.
- The summed product,  $\sum_i (T_i - \bar{T})(m_i - \bar{m}) = -422.1$
- The sum of squares,  $\sum_i (T_i - \bar{T})^2 = 5469.0$

1. Use this information to build a predictive linear model for melt index from the reactor temperature.
2. What is the model's standard error and how do you interpret it in the context of this model? You might find the following software output helpful, but it is not required to answer the question.

```
Call:
lm(formula = Melt.Index ~ Temperature)

Residuals:
    Min       1Q   Median       3Q      Max
-2.5771 -0.7372  0.1300  1.2035  1.2811

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -----      8.60936   4.885 0.000637
Temperature  -----      0.01788  -4.317 0.001519

Residual standard error: 1.322 on 10 degrees of freedom
Multiple R-squared:  0.6508,    Adjusted R-squared:  0.6159
F-statistic: 18.64 on 1 and 10 DF,  p-value: 0.001519
```

3. Calculate the ANOVA table entries:

- RegSS
- RSS
- TSS

4. Prove to yourself that  $R^2 = \frac{\text{RegSS}}{\text{TSS}}$ . Ensure that the value agrees with the R output above. Also ensure you can reproduce the R output.
5. Prove to yourself that you get the identical  $R^2$  value *from only using the raw data* by calculating  $R^2 = [r(x, y)]^2$ .

---

END