# Statistics for Engineering, 4C3/6C3, 2012

# Assignment 5

**Kevin Dunn, dunnkg@mcmaster.ca** **Due date: 13 February 2012, at noon** (no late handins)

---

Covers paired tests, process capability and least squares modelling. This assignment looks longer than it is, and it is good preparation for the midterm. Most questions are from previous midterms and exams.

## Question 1 [1]

List an advantage of using a paired test over an unpaired test. Give an example, not from the notes, that illustrates your answer.

## Question 2 [3] (600-level students)

An *unpaired* test to distinguish between group A and group B was performed with 18 runs: 9 samples for group A and 9 samples for group B. The pooled variance was 86 units.

Also, a *paired* test on group A and group B was performed with 9 runs. After calculating the paired differences, the variance of these differences was found to be 79 units.

Discuss, in the context of this example, an advantage of paired tests over unpaired tests. Assume 95% confidence intervals, and that the true result was one of "no significant difference between method A and method B". Give numeric values from this example to substantiate your answer.

## Question 3 [2]

A bagging system fills bags with a target weight of 37.4 grams and the lower specification limit is 35.0 grams. Assume the bagging system fills the bags with a standard deviation of 0.8 grams:

1. What is the current Cpk of the process?
2. To what target weight would you have to set the bagging system to obtain Cpk=1.3?
3. How can you adjust the Cpk to 1.3 without adjusting the target weight (i.e. keep the target weight at 37.4 grams)?

## Question 4 [3]

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").

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Temperature** = $T$ [Kelvin] | 441 | 453 | 461 | 470 | 478 | 481 | 483 | 485 | 499 | 500 | 506 | 516 |
| **Melt index** = $m$ [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 = $\overline{T} = 481$ K
- Average melt index, $\overline{m} = 4.925$ g per 10 minutes.
- The summed product, $\sum_i \left(T_i - \overline{T}\right)\left(m_i - \overline{m}\right) = -422.1$
- The sum of squares, $\sum_i \left(T_i - \overline{T}\right)^2 = 5469.0$

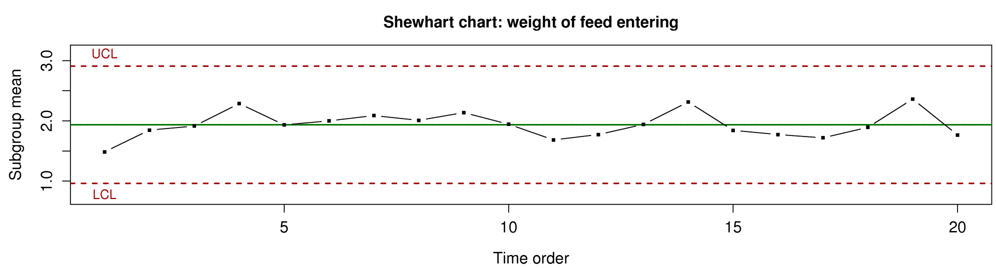
1. Use this information to build a predictive linear model for melt index from the reactor temperature.
2. What is the model's standard error and how do you interpret it in the context of this model? You might find the following software software output helpful, but it is not required to answer the question.

```
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. Quote a confidence interval for the slope coefficient in the model and describe what it means. Again, you may use the above software output to help answer your question.

# Question 5 [2]

The following charts show the weight of feed entering your reactor. The variation in product quality leaving the reactor was unacceptably high during this period of time.

1. What can your group of process engineers learn about the problem, using the time-series plot (100 consecutive measurements, taken 1 minute apart).
2. Why is this variability not seen in the Shewhart chart?
3. Using concepts described in this course, why might this sort of input to the reactor have an effect on the quality of the product leaving the reactor?

## Question 6 [3]

For a distillation column, it is well known that the column temperature directly influences the purity of the product, and this is used in fact for feedback control, to achieve the desired product purity. Use the distillation data set , and build a least squares model that predicts `VapourPressure` from the temperature measurement, `TempC2`. Report the following values:

1. the slope coefficient, and describe what it means in terms of your objective to control the process with a feedback loop
2. the interquartile range and median of the model's residuals
3. the model's standard error
4. a confidence interval for the slope coefficient, and its interpretation.

You may use any computer package to build the model and read these values off the computer output.

---

END