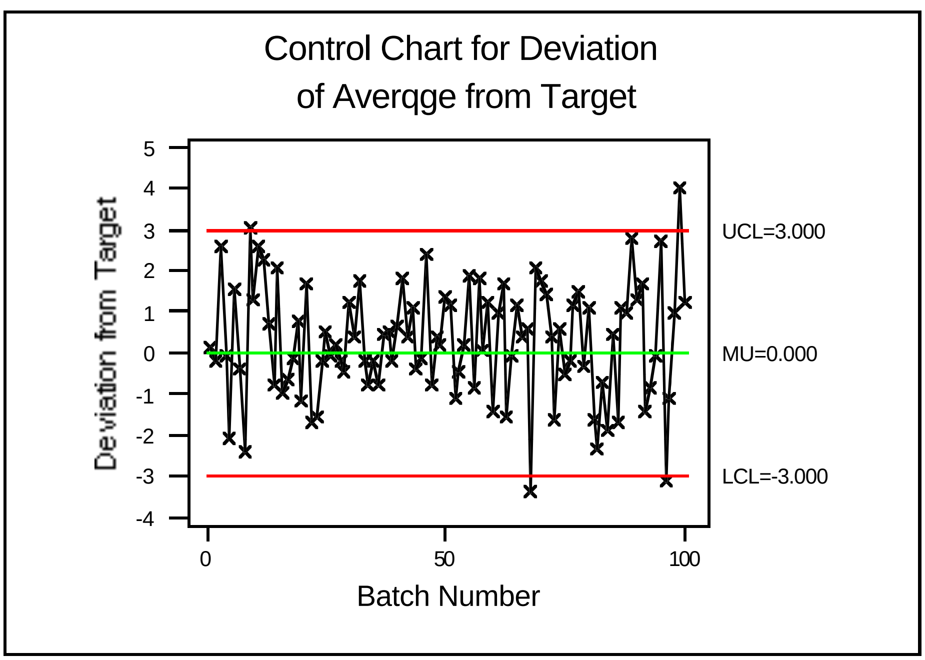
Today: Quality Control (continued)
Earthquake Article and Statistical Estimation
Rest of Assignment # 7 and group formation

Quality Control (continued)

We have already discussed the need for objective measurement in the colour-matching context. In that case, the use of a colour-measuring instrument did a reasonably good job of anticipating whether a customer would judge that a certain painted automobile panel matched a standard colour or not. If you imagine a sequence of such products being produced, the problem is to devise a way to monitor the sequence of panels before they are shipped to the distribution centers (and the retail customer). One hopes to detect any hint of the process variability increasing or decreasing. A method that has general applicability, and is simple to use, is the control chart.

Control charts are used to display quality measurements on a periodic basis (hourly, daily, weekly …) When variability thresholds are crossed, employees discuss the conditions which pertained to the timing of the increased variability, to try to detect the cause of the unusual variability, and to eliminate the cause.

Here is an example of a control chart:



In this example, there appears to be evidence of excess deviation on days 9, 68, 96 and 99. Also, the period from day 33 to 45 appears to be a period of better than usual control of the process. These observations give clues to what aspects of the process can be improved, since the timing of the exceptional events can be linked to external information about the process conditions.

If each point on the chart refers to a days' production, what aspect of the many measurments in a day is actually recorded? If every product can be measured with automated measuring devices, it may be possible to simply average these. Usually, only a random sample of items is measured, and so the sample mean is plotted. The chart uses the 68-95-100 "rule" for normal distributions to plot the Control Limits (UCL and LCL) at ± 3 SDs. The target mean and historical SD are used to draw the green target line and the red Control Limit lines. The idea is that the ± 3 SD lines would not be crossed if everything was as usual. This assumes that the sample average has a normal distribution, since it is a property of normal distributions that ±3 SD is crossed in only a tiny proportion of cases (3/1000 in fact for an exactly normal distribution). So crossing these "3-sigma" lines is evidence that either the mean has shifted, or that the SD has increased. This is the signal to investigate and correct the cause of anomaly.

When the causes of increased varibility are detected and eliminated, the "usual" variability decreases over time.

Why is reduction of variability so important in manufacturing? Consider the sale of lumber, perhaps the familiar "2 x 4". The milling produces boards of various qualities – not only do different species of wood have different characteristics, but some pieces are slightly bent, some have unsightly knots, some are nicked or eaten by insects … If we sell unsorted wood, we can only guarantee the lowest quality and so must accept the lowest price. But if we sort by quality, we can charge higher prices for higher quality. In other words, reducing variability would have a direct impact on improved profits. Similar stories can be told of the profitability of reducing waste, reducing the cost of warranty claims, and improving customer satisfaction. Quality control has been a big part of industrial culture over the past 50 years. Lately the philosophy of quality control has expanded to include more general management practices. A name you will hear in this context is "Six-Sigma" management. This name has its roots in the ± 3 SDs that we have already mentioned.

Here is a (slightly far fetched) example of how quality control philosophy could move out of the realm of manufacturing and into the wider world. Suppose you monitor the marks you get, expressed as a percentage, on each assignment/test/exam you do. The mark is plotted at the date of submission of the assignment or of the test. In monitoring your time series of marks, you note any particularly high or low marks and try to determine what it was that caused the extreme value. The timing of mark in relation to your activities is crucial in this, as is any other information you might have about how you prepared for the test item. By reinforcing successful behavior and reducing unsuccessful behavior, you can increase your average grade and eventually reduce the variability of your grades as well. You learned it here in STAT 100!

Earthquake Article and Statistical Estimation (Tanur pp 249-260):

Living material has a certain concentration of radioactive carbon that is constant while the material is living. But when it dies, the radiocarbon begins to decay, and the rate of decay is known. Thus the concentration of radioactive carbon is dead material can age the material. This technique is used to estimate all the major earthquake dates in a certain location over the last 2000 years.

The Earthquake article describes the use of this technique in a famous study of an area called Pallett Creek near Los Angeles. The San Andreas Fault extends along most of the coast of California. We have similar faults in BC.

The aim of the article is to estimate the chance of an earthquake over the next thirty years. This involves several statistical ideas:

calibration
distribution of sums and averages
shape of distribution via cumulative distribution function
probability plotting
parameters and parameter estimation
confidence intervals for parameter estimation
standard error
Weibull model for durations
insurance premiums

I will give a brief explanation of these ideas in the lecture. But you must read the article to do the second part of assignment #7. Hopefully, the ideas will be made clear by the sequence: lecture, read article, answer assignment.

Assignment #7 second part only (Due Wed, Nov. 27, 2002)

4. Identify, in the Brillinger article on earthquake prediction (Tanur pp 249-260), the sentence in which the following statistical ideas are expressed. You can do this by referencing the page and line number, or else write out the sentence.

a) a probability model can be made specific by the estimation of parameters from data
b) normal distributions often model the spread of sums of independent sources of variation
c) a confidence interval is an interval which has a certain chance of including a parameter value
d) a probability plot is a graphical method for checking whether a certain class of probability distributions (like Normal, Geomtric, or Weibull) fits a sample of data.
e) the risk of an earthquake is quantified by the probability that the earthquake occurs over a certain period of time
f) the insurance premium that one must pay each year for earthquake insurance depends on the amount to be paid out in the event of an earthquake and the probability that the earthquake occurs in the year
g) calibration of a measurement method makes use of the known relationship between the measurement and the quantity that it is supposed to measure.
h) the use of radiocarbon decay to measure the duration since organic material died enables scientists to idetify the age of dead wood fragments

5. Answer question 5. from the Brillinger article (p 260).

Important Note about Assignment #7 (group formation):

You are encouraged to submit this assignment in "groups". Get together with some other students to submit one copy of the assignment 7 solutions – everyone in the group will get the same mark. Groups can be any size up to, but not exceeding, five students. Submissions will be evaluated using the same criteria no matter what the size of the group. If you wish to be assigned to a group, send me an e-mail and I will form groups for those that ask me to. Once a group is formed, members of the group should each send me an e-mail confirming their group membership. I will confirm the group membership on Friday before Friday's class, by sending an e-mail to stat100@sfu.ca. After that, no more grouping will be allowed. To avoid confusion, please make sure you have the agreement of those you wish to group with, and that there are not more than five in the proposed group.

All five Assignment questions are included in this group assignment. Groups of size 1 are OK but then you would miss the learning experience of discussing issues with your peers.