STATS 528: Exam 1

John Sherrill

October 20, 2015

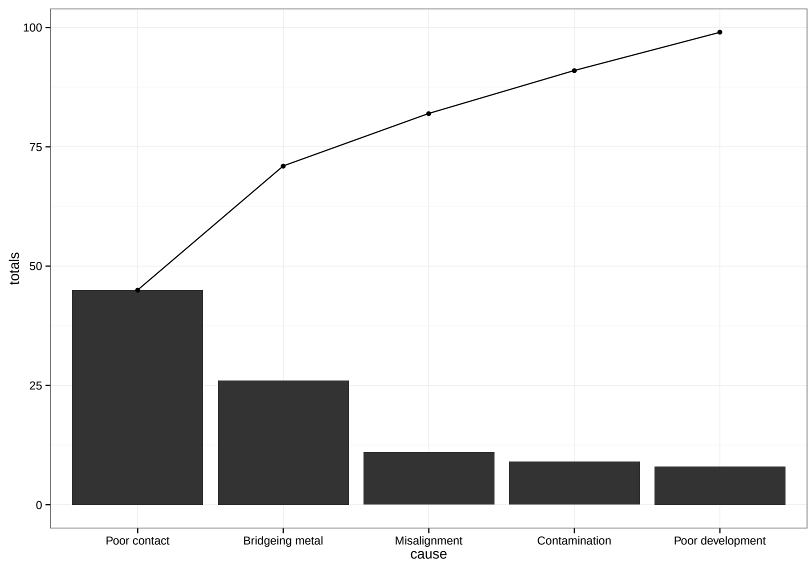
1. (a) $X \sim \text{Bin}(500, 0.002)$

   (b)
$$(P(X < 2))^{10} = (0.736)^{10} = 0.736^{10}$$

2. (a) $X \sim \text{HyGeom}(N = 1000, K = 5, n = 25)$

   (b)
$$P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{\binom{5}{0}\binom{1000-5}{25-0}}{\binom{1000}{25}} + \frac{\binom{5}{1}\binom{1000-5}{25-1}}{\binom{1000}{25}} = 0.994$$

3. Pareto chart for a check sheet:



4. (a)
$$C_P = \frac{USL - LSL}{6\sigma} = \frac{26.2 - 23.4}{6 \cdot .5} = 0.933$$

   (b) Based on Table 8.3 in the text, we have that this process capability ratio does not meet the minimum recommended level of 1.33. While the process may be capable of meeting specifications, it doesn't seem likely that it would happen as consistently as desired.

   (c) The proportion of units currently produced not meeting specifications is given by
$$\begin{aligned} P(x < 23.4) + P(x > 26.2) &= P\Big(z < \frac{23.4 - 25}{1.25}\Big) + P\Big(z > \frac{26.2 - 25}{1.25}\Big) \\ &= \Phi(-1.28) + (1 - \Phi(0.96)) \\ &= 0.269 \end{aligned}$$

   (d) I would say absolutely not. If over a quarter of units produced are not acceptable, then there is clearly a problem and the process is not meeting specificiations.

(e) If the manufacturer seeks to increase the C_p, they must first, make sure the center is lowered from its current value of $\mu = 25$ to $\mu_1 = 23.5 + \frac{26.1-23.5}{2} = 24.8$ so that the process is centered. Also, the characteristic standard deviation $\sigma = 1.25$ is far too large for the process to be considered capable of meeting specifications.

5. (a) We have that $\hat{\sigma} = \frac{\bar{R}}{d_2} = 1.684$.

(b) We have that

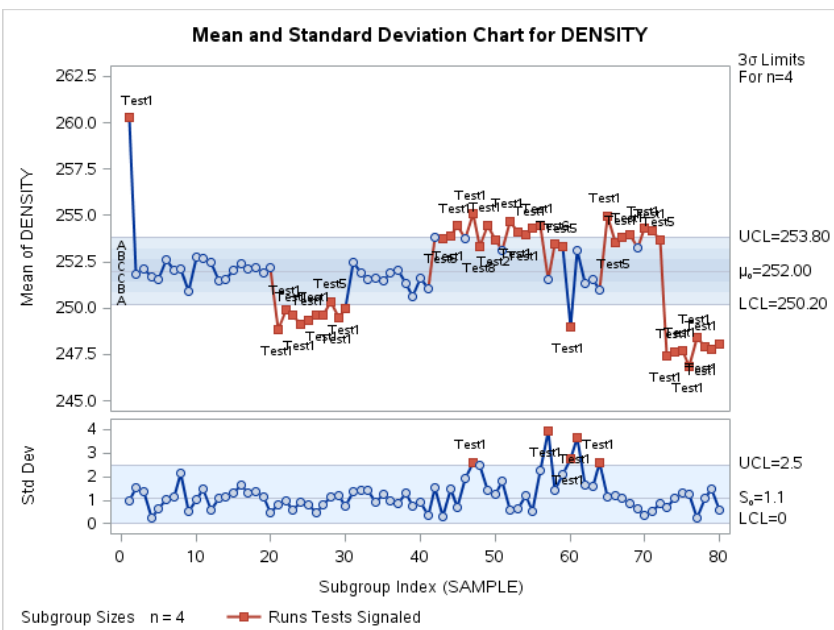
$$UCL = \bar{x} + 3\hat{\sigma} = 150.37 + 3 \cdot 1.684 = 152.054$$
$$LCL = \bar{x} - 3\hat{\sigma} = 150.37 - 3 \cdot 1.684 = 148.686$$

(c) I would not suggest adopting these trial control limits if the machine is considered to be meeting some desired specification as there is one out of control signal in the I-chart labeled "Test". I would also want to investigate into possible explinations for such a high observation.

(d) The difference between the last and second to last observation is greater than the average absolute difference between observations ($\bar{R} = 1.9$) so removal of the last observation would slightly decrease the CL for the MR chart. The LCL would remain 0 and the UCL would decrease slightly as it is a linear function of $\bar{R}$ (namely, $UCL = D_4\bar{R}$).

(e) Because the last observation is slightly above the sample average of $\bar{x} = 150.37$, we have that the centerline would decrease for the I chart. The UCL and LCL are positive and negative linear functions of $\bar{R}$, respectively. Removing the last observation causes a decrease in $\bar{R}$ and thus a decrease and increase in the UCL and LCL for the I chart.

6. (a) $\bar{X} - S$ chart:



(b) It appears that for samples 47, 57, 60, 61, and 64 that the process variability is out of control. All other samples do not signal an out of control process in terms of process variability.

(c) It appears that samples 1, 44-45, 47, 49, 52-56, 65, 67-68, 70-71, are above aim and samples 21-27, 29-30, 60, and 73-80 are all below aim. All other samples appear to be on target (within the control limits).

(d) I would not suggest changing the control limits as it appears that there is reason to suspect that out-of-control signals are being caused by external factors. For example, why did the first sample have such a large mean? There were also *several* runs that were consistently below or above aim. Causes for these phenomena should be determined, and removed, before further analysis of the control limits.

7. (a) Centerlines and control limits for $\overline{x}/R$ charts:

	$\overline{X}$ chart	R chart
UCL	$\overline{\overline{x}} + A_2\overline{R} = 3.245$	$D_4\overline{R} = 0.308$
LCL	$\overline{\overline{x}} - A_2\overline{R} = 3$	$D_3\overline{R} = 0$
CL	$\overline{\overline{x}} = 3.1227$	$\overline{R} = .1197$

(b) There are far more points outside of the control limits than would be expected. A total of 10 data points are outside the control limits in the $\overline{X}$ chart out of 74. We would expect only 0.3% to be outside the control limits.

(c) Centerlines and control limits for I/MR charts:

	I chart	MR chart
UCL	$\overline{\overline{x}} + \frac{3}{d_2}\overline{MR} = 3.267$	$D_4\overline{MR} = 0.265$
LCL	$\overline{\overline{x}} - \frac{3}{d_2}\overline{MR} = 2.979$	$D_3\overline{MR} = 0$
CL	$\overline{\overline{x}} = 3.1227$	$\overline{MR} = .0812$

(d) I would say that in this instance, the proportion of data points outside of the control limits is consistent with what would be expected as there are no data points outside of the UCL and LCL for either the I chart of the MR chart.

(e) For the data used in plot 1, we have

$$\hat{C}_p = \frac{USL - LSL}{6\hat{\sigma}} = \frac{3.23 - 3.01}{6\overline{R}/d_2} = 0.519$$

For the data used in plot 2, we have

$$\hat{C}_p = \frac{USL - LSL}{6\hat{\sigma}} = \frac{3.23 - 3.01}{6\overline{MR}/d_2} = 0.509$$

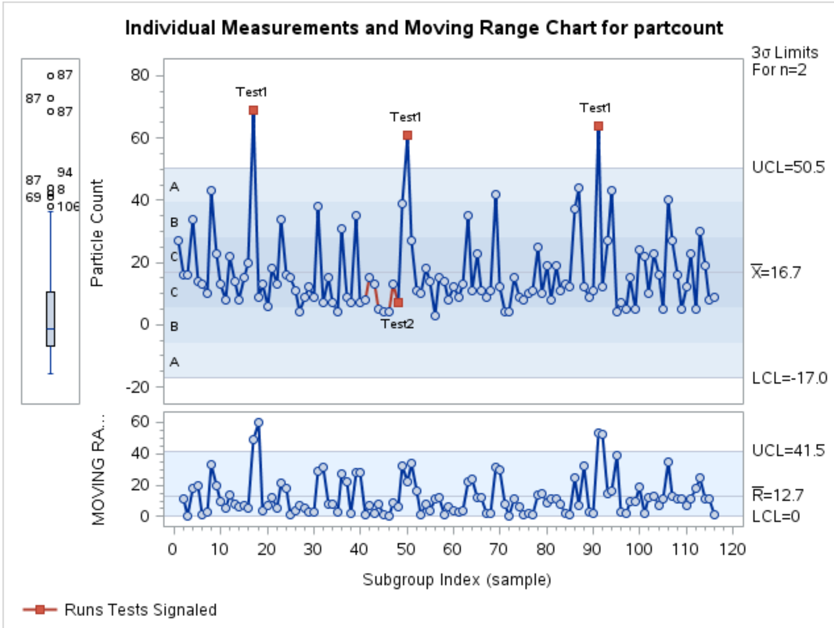
These are both small $\hat{C}_p$ values and do not suggest that the process can consistently product wafers within specifications.

(f) For plot 1 data, $\hat{\sigma}^2 = (\frac{\overline{R}}{d_2})^2 = (\frac{.1197}{1.693})^2 = 0.005$. For plot 2 data, $\hat{\sigma}^2 = (\frac{\overline{MR}}{d_2})^2 = (\frac{.0812}{1.128})^2 = 0.005$. The first represents the estimated variability of all the seperate measurements and the second represents the estimated variability of between wafer measurements.

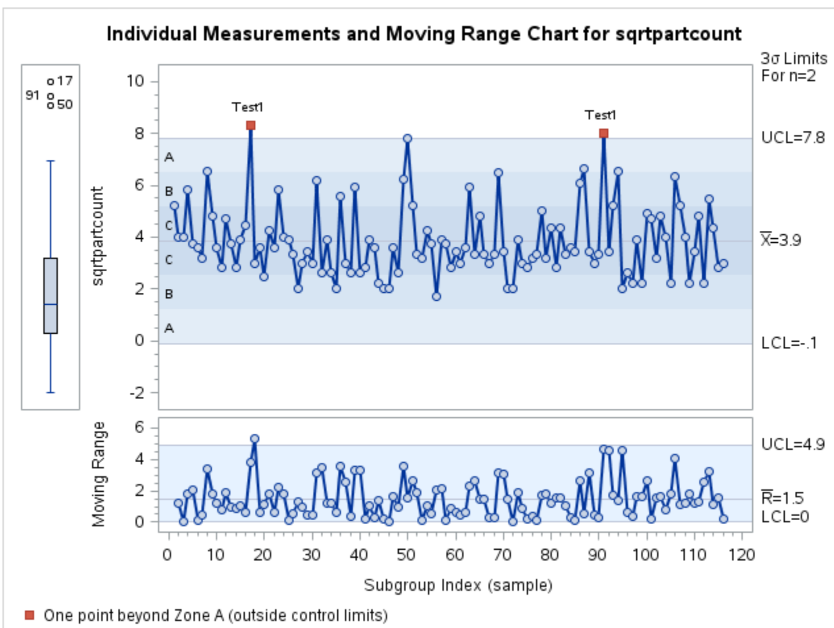
(g) The $\overline{x}$ chart and I chart are the exact same. It is only the definition of out-of-control that changes. Thus, the charts are not consistent in that they draw different conclusions; if considered single measurements, the process is considered statistically in control, if not, then it is out of control.

(h) I would recommend using a control chart process that takes into consideration the correlation between measurements in a single sample.

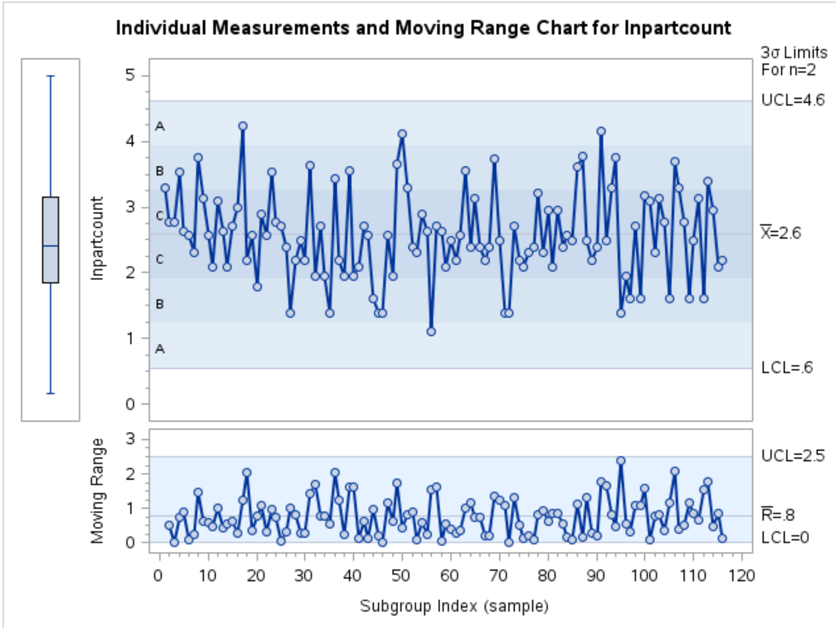
8. (a) The process does not appear to be in control. It's pretty obvious in the boxplot provided that the data is not symmetric. This is not ideal for process monitoring via control charts.



(b) The problems noted just above appear to be remediated a little by taking the square root of the response but there is still a right skew to the data and the process could still be considered out of control.



(c) Now, there appears to be symmetry in the data and the process is not yielding any out of control signals.



(d) I would recommend that they use the log transformed data as that appears more symmetric and is better modeled using a normal distribution. This is an important assumption made in several of the quality control analysis techniques.