## Chapter 1 Section 1

**1.** If quality of conformance has to do with small variation and one wishes to assure it, it will be necessary to measure, monitor, find sources of and seek ways to reduce variation. All of these require data (information on what is happening in a system producing a product) and therefore the tool of statistics. Hence, quality and statistical methods are directly related.

**2.** Mechanical devices whose features of interest vary substantially tend to be noisy, prone to breakdown, difficult to service and inefficient. In the service sector, variation from what is promised/expected is the principle source of customer dissatisfaction. Customer dissatisfaction is undesirable because if a customer is not satisfied, he will seek another vendor or source to meet his need.

**3.** If a good or service is designed properly that does not guarantee quality. Quality of conformance may be an issue, i.e., variation of important features as described in question 2 above can lead to serious customer dissatisfaction therefore poor quality.

**4.** If a good or service conforms to design specifications, that does not guarantee quality because the design may not produce a good or service that is fit for use when no variation occurs, i.e., a poor design for the proposed performance.

## Chapter 1 Section 2

**1.** If processes can be made to work effectively, resulting products or services will be good is the rationale behind a process orientation. Further, root causes of problems will more likely be identified and removed. Material and time will be saved as well as producing goods or services that are considered quality.

**2.** A customer focus relates to quality in two ways. Studying customer behavior and desires can drive creation of a designed product that is fit for use. Receiving feedback from customers and collecting data concerning a current product or service gives insight as to variation in important features of a good or service. High variation is directly connected with customer dissatisfaction and must be addressed immediately, i.e., poor quality of conformance. Low variation with an appropriately designed product or service results in positive customer feedback.

**3.** Motivations for a corporate continuous quality improvement emphasis are survival and growth. Competitiveness in the marketplace will force companies who aren't continually improving quality of design and conformance from the marketplace.

**4.** Effective measurement is a prerequisite to success in process improvement because if one cannot reliably measure important characteristics of what is being done to produce a good or service, there is no way to tell whether design requirements are being met and customer needs are genuinely being met.

**5.** Control charts are the basic tools used for monitoring processes and issuing warnings of apparent process instability.

**6.** If a process is stable or consistent, it is not necessarily producing high quality goods. The feature(s) of interest could be taking values that are consistent but far from the desired or designed value(s). Or, the feature(s) of interest could be consistently of high variation, directly related to poor quality.

## Chapter 1 Section 3

**1.** The top-to-bottom direction of a flowchart usually corresponds to a time dimension.

**2.** Extra "columns" could be constructed that correspond to, say, different locations or plants or different department spheres of responsibility, still maintaining the top-to-bottom time dimension for each column.

**3.** A "cause and effect" or "fishbone" diagram are other names of the Ishikawa chart.

**4.** One purpose of the Ishikawa chart is to provide a tool that organizes ideas from a brainstorming session concerning some matter of interest, either a problem or quality issue. Further, the Ishikawa chart is constructed in such a way that gives clear direction for future action.

## Chapter 1 Section 4

**1.** It is more desirable to have data that provide a true picture of process behavior than to obtain "good numbers" or "favorable results" because effective decision-making can be made only when the true picture of process behavior is understood.

**2.** People who have seen data collected by themselves or others that were used to harm them or their colleagues will most likely not be cooperative in a data collection event. Further those who have made an honest and sincere effort at data collection in the past only to see their efforts ignored will almost surely guarantee that future data collection efforts produce nothing useful.

**3.** If operational definitions are not clear before the data collection effort begins, the collected data may very well represent values for multiple unknown variables, i.e., nothing useful can be obtained from an analysis of the collected data.

**4.** A knowledge of who, how and when the data were collected is most likely not known. Thus, an accurate understanding of how to proceed with an appropriate analysis cannot be reliably made.

**5.** Through documentation of who, how and when the data were collected, ambiguities can be eliminated that prohibit an appropriate analysis of the data collected.

**6.** The *x*, *y*, symbol, color, time, symbol size could represent six variables.

**7.** A checksheet can be easily and quickly constructed with a simple interpretation.

**8.** A large sample is not necessarily optimal. Instead, one should think in terms of (1) the size of variation that must be accounted for and (2) the size of an effect that is of practical importance. If no variation, a sample of size $n = 1$ is sufficient. If some variation exists and a small effect is of practical interest, more data may be needed.

**<u>Chapter 1 Section 5</u>**

**1.** A simple histogram can portray “spread” of a process and “location”. Also, for a stable process, shape or distribution of process data can be inferred.

**2.** Time trends of the process data cannot be portrayed by a simple histogram.

**3.** The run chart can depict trends in process data and where in time outliers occur. This gives insight into possible special or assignable causes.

**4.** Beginning at time 1, data slowly trend upward to the mid time point where data occur randomly around the center and then begin to trend up again after the ¾ time point. Or vice versa, at time 1, data slowly trend down from above the center and after the ¾ time point, continue to trend down.

**5.** The Pareto chart is particularly useful for getting people to prioritize their efforts and focus first on the biggest quality problems an organization faces.

**6.** The rationale behind the Pareto chart is the most often occurring situation should perhaps receive first attention and likewise the second most often occurring situation the second attention. Most of anything is traceable to a few causes is the underlying theory of the Pareto Chart.

**Chapter 2 Section 1**

**1.** Comparing a measurement method or device to a standard one and, if necessary, working out conversions that will allow the method to produce "correct" (converted) values on average is calibration. Outputs from the measurement device to a "known" or "standard" value permits the analyst to compare what is being recorded to what really is, i.e., an assessment can be made and using calibration, a correction made to the measured value so the result is on average correct.

**2.** Measurand (x) is the true density (g/cc) of a selected pellet after firing at 1400°C for a selected length of time. The symbol y is the recorded density (g/cc) of a selected pellet after firing at 1400°C for a selected length of time. The term $\varepsilon$ is the error from a recorded *y* value and the measurand *x* for a selected pellet after firing at 1400°C for a selected length of time. The term δ is the bias or difference in average recorded value from repeat observations of a single pellet fired for a selected length of time at 1400°C.

**3.** Assuming constant bias, independent of original density and different length of firing times, implies 5 *x*'s, 5 $\varepsilon$'s, 5 *y*'s and one δ. If the constant bias is only for a selected firing time with possibly different original densities, then 5 δ's, one for each of the different firing times.

**4.** No, should have recorded original density. Without original density, cannot get the difference " after minus before" which reflects firing effect.

**5.** 1 measurand, 5 y's, 5 $\varepsilon$'s and one δ.

**6.** $\sqrt{3.4}$

**7.** **a.** .797

**b**. $\sigma_y = \sqrt{\sigma_x^2 + \sigma_{measurement}^2}$

**8.** **a.** 5.7

**b**. $\mu_x + \delta$

**Chapter 2 Section 2**

**1.** **a**. $x + \mu_\delta$

**b**. $\sigma_\delta^2 + \sigma_{device}^2$

**c.** Part (b) is more important because the square root of (b) is the $\sigma_{R\&R}$.

**2.** **a.** $x_{1i} - x_{2i} = d_i$, -.1, .9, -1, -1.1, -.9; $\bar{d} = -.44$; estimates $\delta_1 - \delta_2$

**b.** $s_d = .847;\ estimates\ \sqrt{\sigma_{device1}^2 + \sigma_{device2}^2}$

**c.** $\bar{d} \pm t_{4,.05}\, s_d/\sqrt{5}$ ; ( -1.248, .368)

**3**. **a.** (.733, 7.042). Since 95% C.I. for $\frac{\sigma_1}{\sigma_2}$ includes 1, implies no difference in consistency.

**b.** (-1.441, .561) using df = 5 because Sattherwaite approx. df is 5.49, rounding down gives 5. Could have used df = min ( (5-1), (5-1)) = 4. No difference in bias, $\delta_1 - \delta_2$ doesn't depart from 0 since the confidence interval includes 0.

**4**. **a.** $s_1$ *estimates* $\sqrt{\sigma_x^2 + \sigma_{device1}^2}$

**b.** $s_2$ *estimates* $\sqrt{\sigma_x^2 + \sigma_{device2}^2}$

**c.** $\delta_1 - \delta_2$ (equipment 1 minus equipment 2).

**5.** The method in problem 3 is better because the variation in the estimate of $\mu_d = \delta_1 - \delta_2$ is smaller.

**6.** $\bar{y}_1$ estimates $\delta + x_1$ and $\bar{y}_2$ estimates $\delta + x_2$, so $\bar{y}_1 - \bar{y}_2$ estimates $x_1 - x_2$.

**7**. **a.** The same as in problem 3(b), but this interval now estimates $\mu_{x1} - \mu_{x2}$.

The 95% confidence interval using Satterthwaite df approximation of df = 5 becomes (-1.441, .561). The df truncated from 5.49. Could have used a more conservative df = min ( (5-1), (5-1)) = 4.

**b.** The average density after firing using method 1 for a selected length of time minus that for method 2, i.e., $\mu_{x1} - \mu_{x2}$.

**c.** No, only one device is considered, device 1, and its bias $\delta$ cannot be split out, i.e., $\bar{y}_1$ *estimates* $\mu_{x1} + \delta$ *and* $\bar{y}_2$ *estimates* $\mu_{x2} + \delta$.

**8.** $\bar{y} \pm t_{4;.025}\, s/\sqrt{n}$ becomes (4.711, 6.689) for $\mu_x + \delta$.

**9.** **a.** 95% C.I. for $\sigma_{measurement}$ (.477, 2.290)

**b.** $\bar{y} \pm t_{4;.025}\, s/\sqrt{n}$ becomes (4.711, 6.689) for $x + \delta$.

## Chapter 2 Section 3

**1.** **a.** $s_y^2 = 16.3333, n_y = 3;\ s^2 = 3, n = 4;\ \ \hat{\sigma}_x = \sqrt{\max(0, 16.3333 - 3)}$ = 3.6514

**b.** approximate df = 177.7769/(133.3883 + 3) = 1.3 so let approx. df = 1. 95% confidence interval for $\sigma_x$ becomes ( 3.6514 $\sqrt{1/5.024}$, 3.6514 $\sqrt{1/.001}$ ) or (1.629, 115.469).

**2**. **a.** $\sqrt{6.66}$ = 2.5807 = $\hat{\sigma}_{repeatability}$, the confidence interval is
$(\sqrt{3.15},\ \sqrt{22.22})\ or\ (1.7748,\ 4.7138)$.

**b.** $\sqrt{1.92}$ = 1.38564 = $\hat{\sigma}_{reproducibility}$, the confidence interval is (0, 3.018)

**c**. Instrument quality should be addressed, variation operator to operator is less than repeated measurements on same item.

**3.** **a.** $\sqrt{3.75} = 1.9365 = \hat{\sigma}_{device}$; $(1.3304, 3.5355)$ is the 95% C.I. for $\sigma_{device}$.

**b.** $\sqrt{1.96} = 1.4 = \hat{\sigma}_x$; $(1.9442,\ 2.773)$ is the 95% C.I. for $\sigma_x$ .

**c.** No, $\hat{\sigma}_x < \hat{\sigma}_{device}$

**<u>Chapter 2 Section 4</u>**

**1.** **a.** m = 4, I = 3, J = 1.

**b.** $R_{11} = 4$, $R_{21} = 5$, $R_{31} = 3$, $\bar{R} = 4$, $\frac{\bar{R}}{d_2(4)} = \frac{4}{2.059} = 1.9426 = \hat{\sigma}_{device}$

**c.** $\bar{y}_i = \mu + \alpha_i + \bar{\epsilon}_i$, I = 3, J = 1, m = 4. So, Var $\bar{y}_i$ is $\sigma_\alpha^2 + \frac{\sigma^2}{4}$.

In the context, $\sigma_x^2\ is\ \sigma_\alpha^2$.

$\bar{y}_i{'s}$ are 20.5, 18, 21.25. $\frac{\bar{\Delta}}{d_2(3)} = \frac{3.25}{1.693} = 3.68512.$

$\hat{\sigma}_x = \sqrt{\max(0, [\left(\frac{3.25}{1.693}\right)^2 - \left(\frac{1}{4}\right)(1.9426)^2])} = 1.6558.$

**2.** No, only one operator.

**3.** **a.** I = 1, J = 3, m = 4.

**b.** $\bar{R} = \frac{17}{3}, d_2(4) = 2.059. So, \frac{\bar{R}}{d_2(4)} = 2.752 = \hat{\sigma}_{repeatability}$.

**c.** $\bar{y}_i{'s}$ are 20.5, 21, 17.5, $\Delta = 3.5, d_2(3) = 1.693,$

$\hat{\sigma}_{reproducibility} = \sqrt{\max(0, [\left(\frac{3.5}{1.693}\right)^2 - \left(\frac{1}{4}\right)(2.752)^2])} = \sqrt{2.3805} = 1.5428.$

**d.** No, we only have data from one “x”.

**4.** **a.** I = 3, J =2, m = 4.

**b.** $\hat{\sigma}_{repeatability} = \sqrt{.000005} = .0022361.$

$$\hat{\sigma}_{reproducibility} = \sqrt{\max\left(0, \quad \left[\left(\frac{1}{12}\right)(.0000007) + \left(\frac{2}{12}\right)(.0000053) - \left(\frac{1}{4}\right)(.000005)\right]\right)}$$
$$= \sqrt{0} = 0.$$

$$\hat{\sigma}_{R\&R} = \sqrt{\max\left(0, \quad \left[\left(\frac{1}{12}\right)(.0000007) + \left(\frac{2}{12}\right)(.0000053) + \left(\frac{3}{4}\right)(.000005)\right]\right)}$$
$$= .002181$$

**c.** (3)(2)(3) = 18 df for 95% C.I. of

$\sigma_{repeatability}$ $becomes \left[ .0022361\sqrt{\frac{18}{31.526}}, .0022316\sqrt{\frac{18}{8.231}} \right]$ $or$ $[.001689, .0033067]$

No confidence interval for $\sigma_{reproducibility}$.

df for confidence interval for $\sigma_{R\&R}$ is

$$\frac{(.002166)^4}{\frac{1}{16}\left[\frac{(.0000007)^2}{9}+\frac{2(.0000053)^2}{9}+\frac{3(.000005)^2}{6}\right]} = 18.737 \ or\ truncating\ gives\ 18\ as\ the\ df.$$

The 95% confidence interval for $\sigma_{R\&R}$ becomes

$$[.002181\sqrt{\frac{18}{31.526}}, \quad .002181\sqrt{\frac{18}{8.231}}] \ or\ [.001648, .0032253].$$

**d.** $\widehat{GCR} = \frac{6(.002181)}{.2} = .06543$, 95% $confidence\ limits\ for\ GCR\ become$

$\left[\frac{6(.001648)}{.2}, \frac{6(.0032253)}{.2}\right]$ $or$ $[.0494\ .09676]$.

## Chapter 2 Section 5

**1.** **a.** $\hat{y}(x) = .00801 + 1.00104x$

**b.** $\sqrt{MSE} = \sqrt{.00267054} = .051677$

**c.** $[.051677\sqrt{\frac{12}{23.337}}, .051677\sqrt{\frac{12}{4.404}}]$ $or$ $[.037056, .085303]$ is the 95% CI for $\sigma_{repeatability}$.

**2.** **a.** $\hat{x} = \frac{6.11 - .00801}{1.00104} = 6.09565$

**b.** $1.00104 \pm t_{12;.025}$ (.00311) becomes $1.00104 \pm (2.179)(.00311)$ or 95% C.I. for the slope becomes [ .99426, 1.00781 ] , yes, it includes 1.

**3.** $\hat{y}(8) = 8.0163$, $the$ 95% $prediction\ interval\ is$ $(7.8997, 8.1329)$.

**4.** No, $y_{new}$ is outside y's in the data used to model the relationship.

## Chapter 2 Section 6

**1.** **a.**

| $\bar{\bar{p}}$ | $\bar{\bar{p}}\,(1-\bar{\bar{p}})$ | $\overline{\hat{p}(1-\hat{p})}$ |
|---|---|---|
| 0.656250 | 0.225586 | 0.224609 |
| 0.656250 | 0.225586 | 0.220703 |
| 0.500000 | 0.250000 | 0.24807 |
| 0.921875 | 0.072021 | 0.069336 |
| 0.765625 | 0.179443 | 0.174805 |
| 0.953125 | 0.044678 | 0.043945 |
| 0.796875 | 0.161865 | 0.155274 |
| 0.968750 | 0.030273 | 0.029297 |
| 0.890625 | 0.097412 | 0.094727 |
| 0.984375 | 0.015381 | 0.014649 |

**b.** $\frac{\hat{\sigma}^2_{\text{reproducibility}}}{\hat{\sigma}^2_{R\&R}} = \frac{.0027245}{.130225} = .0209 \; or \; 2.09\%$; $Note$: .130225 is the average of the $\bar{\bar{p}}\left(1-\bar{\bar{p}}\right)$ column and thus equals $\hat{\sigma}^2_{R\&R}$. Also, $\hat{\sigma}^2_{\text{repeatability}} = .1275 =$ the average of the $\overline{\hat{p}(1-\hat{p})}$ column.
So, $\hat{\sigma}^2_{\text{reproducibility}} = \hat{\sigma}^2_{R\&R} - \hat{\sigma}^2_{\text{repeatability}} = .130225 - .1275 = .0027245$.

**c.** $\hat{\sigma}_{reproducibility} = \sqrt{.0027245} = .052196$

**d.** $\hat{\sigma}_{repeatability} = \sqrt{.1275} = .35707$

**e.** $1 - 1 \pm 1.645\sqrt{\frac{(2)(.9)(.1)}{16}} \; or \; 0 \pm .1744$ No, the $\hat{p}'s$ are very close for each part.

**f.** $\hat{p}_{1i} - \hat{p}_{3i} = d_i$; $\bar{d} \pm t_{9;.95}\frac{s_d}{\sqrt{n}}$; $s_d = .0574$; $90\%\; C.I.\; (-.0458, .0209)$.

## Chapter 3, Section 1

**1.** Identifying when the effect of a special cause may have entered the process and evaluate whether a process is stable with respect to aim or variation are the purposes of control charting.

**2.** Standards given implies $\mu \; and \; \sigma$ are known. Retrospective implies $\mu \; and \; \sigma$ are not known but must be estimated from process data.

**3.** Control charts are designed to detect special cause variation. Common cause variation is variation caused by factors built-into or assumed to be part of the process or system. Special cause variation is variation caused by factors not built-in or assumed part of the process or system.

**4.** Control limits get closer to the center line as subgroup size increases.

**5.** If the multiple changes from 3 to 2, more frequent false alarms will result. But if a change in the process occurs, the change will be detected sooner.

**6.** If the multiple increases above 3, less frequent false alarms will result. If a change in the process occurs, the change will not be detected as soon.

**7.** No, a retrospective chart could be stable and centered around, say, 10, but the target is 4. It is likely most items will be outside specs. Another scenario could be the retrospective or standards given chart determines a stable process with large variation. Thus, the specs may be tight enough that much of the stable process occurs outside specs.

**8.** If a process is judged unstable or out-of-control, estimating or predicting the % of items inside specs doesn't make sense, in other words, the prediction is not reliable. Moment to moment will produce different percentages of items inside specs. If $\sigma$ is very small, it is possible a large % of items are inside specs, i.e., cycling or trending, thus unstable, yet all inside control limits, with a possible small $\sigma$ if the R or s chart is stable but the chart for aim is not stable.

**9.** **a.** 10
**b.** 5
**c.** Analyst 2, eliminates known line to line variation.

**Chapter 3 Section 2**

**1.** **a.** A set of 9 containers sampled in a selected hour is the subgroup. Subgroup size is 9. 40 subgroups are selected.
**b.** $4 \pm 3\frac{.1}{\sqrt{9}}$ ; $4 \pm .1$; $(3.9,\ 4.1)$; $LCL_{\bar{x}} = 3.9$; $UCL_{\bar{x}} = 4.1$.
**c.** Subgroup ranges; $D_1\sigma = .546(.1) = .0546$; $D_2\sigma = 5.394(.1) = .5394$.
**d.** In (b), it is Standards given, $\mu$ and $\sigma$ known. In (c), it is also standards given because $\mu$ and $\sigma$ are known.
**e.** One decision rule implying stability is "all points within control limits". However, there is a small probability a subgroup statistic falls outside the control limits even when the process is stable.

**2.** **a.** $\bar{\bar{x}} = 3.9, \bar{R} = .56, \bar{s} = .48$, $LCL_R = .184(.56) = .10304$, $UCL_R = 1.816(.56) = 1.0169$.
**b.** $3.9 \pm .337\,(.56)$ ; $3.9 \pm .18872$; $LCL_{\bar{x}} = 3.7112$, $UCL_{\bar{x}} = 4.0887$.
**c.** $LCL_s = .239(.48) = .1147$; $UCL_s = 1.7611(.48) = .8453$.

**d.** $LCL_{\bar{x}} = 3.9 - 1.032(.48) = 3.4046;$

$UCL_{\bar{x}} = 3.9 + 1.032(.48) = 4.3953.$

**e.** $\frac{\bar{R}}{d_2} = \frac{.56}{2.97} = .1885;$ $\frac{\bar{s}}{c_4} = \frac{.48}{.9693} = .4952.$ $\bar{\bar{x}} = 3.9$ is an estimate of the avg. distance from bottom to handle.

**3.** **a.** $LCL_{\bar{x}} = 6 - 3\frac{1.5}{\sqrt{4}} = 3.75;$

$UCL_{\bar{x}} = 6 + 3\frac{1.5}{\sqrt{4}} = 8.25.$ No out of control points.

**b.** $LCL_s = 0$; $UCL_s = 2.088(1.5) = 3.132;$ pt. 3 is outside limits.

**c.** $LCL_{\bar{x}} = 6.58 - 1.628(1.72) = 3.779;$

$UCL_{\bar{x}} = 6.58 + 1.628(1.72) = 9.38.$

$LCL_s = 0$; $UCL_s = 2.266(1.72) = 3.897;$ No instability.

**d.** $\frac{\bar{s}}{c_4} = \frac{1.72}{.9213} = 1.867;$ $E(R) = d_2 1.867 = 2.534(1.867) = 4.7307.$

**4.** **a.** The subgroup is a single Series XX transmission housing.

The subgroup size is 1.

**b.** 34

**c.** .0001, .0003

**d.** Individuals retrospective. Subgroup size is 1.

$\overline{MR} = \frac{.02472}{34} = .000727;$ $\frac{\overline{MR}}{d_2} = \frac{.000727}{1.128} = .000645.$

$LCL_x = 3.7805 - 3(.000645) = 3.77856;$

$UCL_x = 3.7805 + 3(.000645) = 3.782435.$ All points inside limits.

**Chapter 3 Section 3**

**1.** **a.** Attributes because the variable values are counts.
**b.** Poisson ($\lambda = .05$)
**c.** 30(.05) = 1.5, Poisson ($\lambda = 1.5$)

**d.** $\sqrt{1.5} = 1.2247$

**e.** E($X$/30) = 1.5/30 = .05; Var($X$/30) = 1.5/900 = .0016667;

$\sigma\left(\frac{x}{30}\right) = .0408233.$

$LCL_{x/30} = 0;\ UCL_{x/30} = .05 + 3(.0408233) = .17247.$

No points outside limits**.**

**f.** u-chart

**g.** $\hat{\lambda}_{pooled} = \frac{8}{210} = .03809.$

$$LCL_{\frac{x}{30}} = .03809 - 3\sqrt{\frac{\hat{\lambda}_{pooled}}{30}} = 0;$$

$$UCL_{\frac{x}{30}} = .03809 + 3\sqrt{\frac{\hat{\lambda}_{pooled}}{30}} = .14498$$

**2.** **a.** Attribute because the variable values are counts.

**b.** Binomial, $n$ = 30; $p$ = .05.

**c.** $\mu_x = 30(.05) = 1.5.$

**d.** $\sigma_x = \sqrt{30(.05).95} = 1.19373.$

**e.** $LCL_{\hat{p}} = .05 - 3\sqrt{\frac{(.05).95}{30}} = .05 - .11937 = 0.$

$UCL_{\hat{p}} = .05 + 3\sqrt{\frac{(.05).95}{30}} = .05 + .11937 = .16937.$

**f.** $LCL_x = np - 3\sqrt{np(1-p)} = 1.5 - 3.5812 = 0.$

$UCL_x = np + 3\sqrt{np(1-p)} = 1.5 + 3.5812 = 5.0812.$

**g**. $LCL_x = n\hat{p} - 3\sqrt{n\hat{p}(1-\hat{p})} = 1.142857 - 3.14545 = 0.$

$UCL_x = n\hat{p} + 3\sqrt{n\hat{p}(1-\hat{p})} = 1.142857 + 3.14545 = 4.288.$

## <u>Chapter 3 Section 4</u>

**1.** No patterns and all Q's are within control limits.

**2.** If only "outside control limits" rule is applied, it is possible for Q's to trend up or down or cycle and all be inside control limits. Using only the rule "outside limits" would incorrectly interpret these scenarios as stable.

**3.** The frequency of false alarms will increase if the extra alarm rules are used in addition to the "3 sigma" limits.

**Chapter 3 Section 5**

**1.** ARL means average run length until the first “believed”out-of-control point is identified. All OK ARL is the average run length until the first “believed” out-of-control point is identified, when, in fact, the process is stable or “All OK”.

**2.** A long or large ARL is desired when a process is stable. Under non-stable situations, a short or small ARL is desired.

**3.** **a.** 370

**b.** $UCL_{\bar{x}} = 20 + 3\,\frac{4}{\sqrt{4}} = 26.$

$LCL_{\bar{x}} = 20 - 3\,\frac{4}{\sqrt{4}} = 14.$

$p = Prob\left( Z > \frac{26-21}{2} \right) + Prob\left( Z < \frac{14-21}{2} \right) = .0064$

$\frac{1}{p} = 156.25.$ When $\mu = 21,$ the ARL is about 156 or 157.

**4.** **a.** $UCL = 4 + 3\sqrt{4} = 10,\; LCL = 0.\; Letting\ \lambda = 4,\;\; p = 1 - P(X \le 10).$

$p = 1 - .99716 = .00284.\; \frac{1}{p} = 352.1\,.\, ARL = 352\ or\ 353.$

**b.** $p = 1 - P(X \le 10).\, \lambda_{new} = 8. \quad p = 1 - .815886 = .184114.\; \frac{1}{p} =$ $5.43\,.\; ARL = 5\ or\ 6.$

**c.** $Let\ X = \#\ items\ per\ 2\ units. So, X \backsim Poisson(8).\; UCL_x = 8 + 3\sqrt{8} = 16.485,\;\; LCL_x = 0.\; Letting\ \lambda\ per\ 2\ items = 8,\;\; All\ OK\ ARL,\;\; p = 1 - P(X \le 16).$

$p = 1 - .996282 = .003718.\; \frac{1}{p} = 268.96\,.\, ALL\ OK\ ARL = 268\ or\ 269.$

$Using\ UCL_x = 16.485\ and\ LCL_x = 0, and$

$Letting\ X = \#\ items\ per\ 2\ units\ and\ a\ new\ \lambda\ per\ item = 8\;\; X \backsim Poisson(16).\; p = 1 - P(X \le 16).\; p = 1 - .56596 = .4340.\; \frac{1}{p} =$ $2.30\,.\; ARL = 2\ or\ 3.$

**5.** (b)

**6.** **a**. Control limits are a function of *n*.

**b.** Because the control limits are the same number of $\frac{\sigma}{\sqrt{n}}$ from $\mu$.

## Chapter 3 Section 6

**1.** **a.** T($t$) = 4. When the process is hitting T($t$) = 4, the process is considered optimal**.**

**b.** E($t$ = 1) = 2; E($t$ = 2) = 3; E($t$ = 3) = 4. $\Delta E(2) = 3 - 2 = 1$. $\Delta E(3) = 4 - 3 = 1. \Delta^2 E(3) = 1 - 1 = 0$. E($t$) is departure of the process from a targeted value. $\Delta E(t)$ is an indicator of how the departure from the process is changing. $\Delta^2 E(t)$ indicates how the changing departure from the process is changing.

**c.** $\Delta \mathrm{X}(3) = .8\Delta E(3) + 1.6E(3) + 1.9\,\Delta^2 E(3) = .8(1) + 1.6(4) + 1.9(0)$ = 7.2.

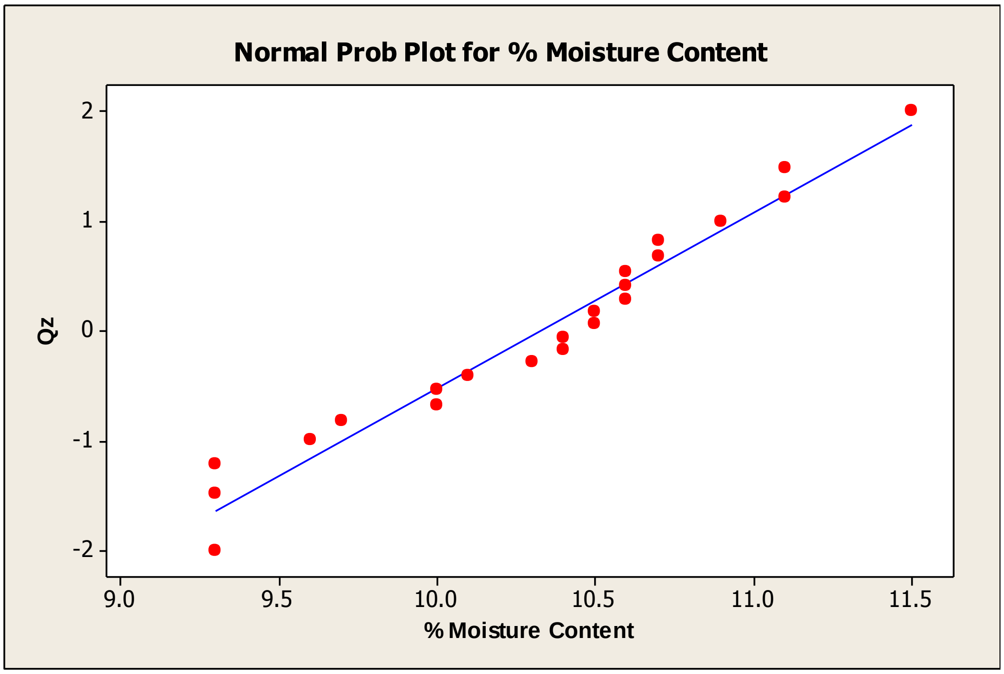
**2.** **a.** If no relationship exists between Y and X, most likely nothing will happen. The process will be similar to a random walk until deterioration of mechanics occur, then trending strongly one way or the other or wildly oscillating. If Y is inversely related to X and since the "first" $\Delta \mathrm{X}(3)$ is positive with all positive K's, the Y's will most likely be pushed down away from T(t), at least initially.

**b.** Do an experiment. Choose, say, 3 sets of K's, each for *n* time periods**.** Calculate $S_1$, $S_2$, and $S_3$ where $S_i = \frac{1}{n}\sum_1^n E(t)^2$. The smallest $S_i$ suggests where to start with a potentially effective set of K's.

**c.** Control charts monitor a process and do not provide a real time adjustment to the process**.** Control charting can point to a point in time (or earlier) where efforts need to be made to find a cause for any unusual patterns on the chart.

## Chapter 4 Section 1

**1.** **a.**

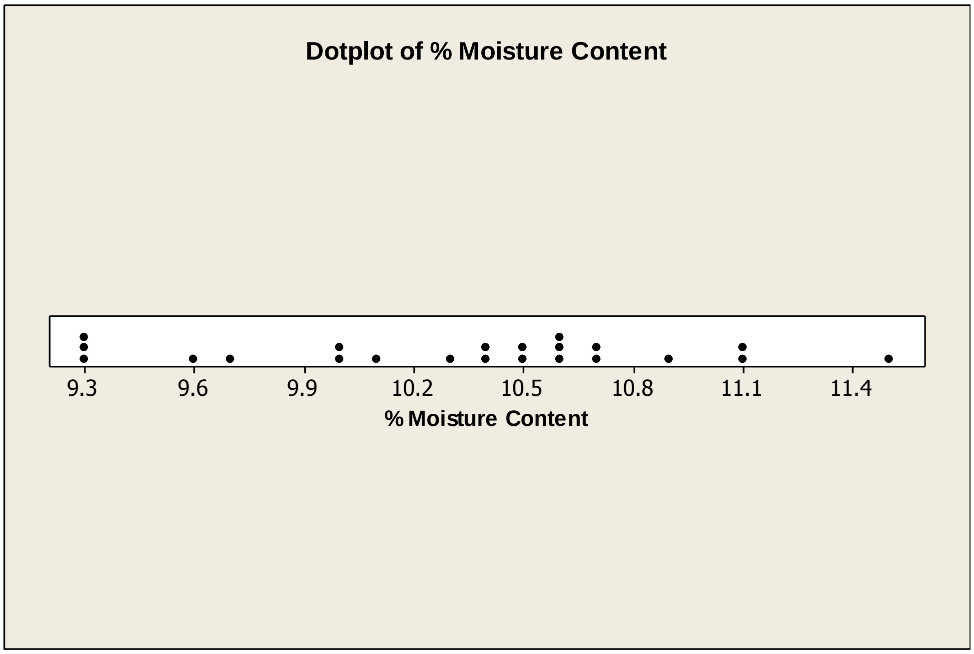



It appears normal distribution is a reasonable model for % moisture content because the plot is approximately linear.

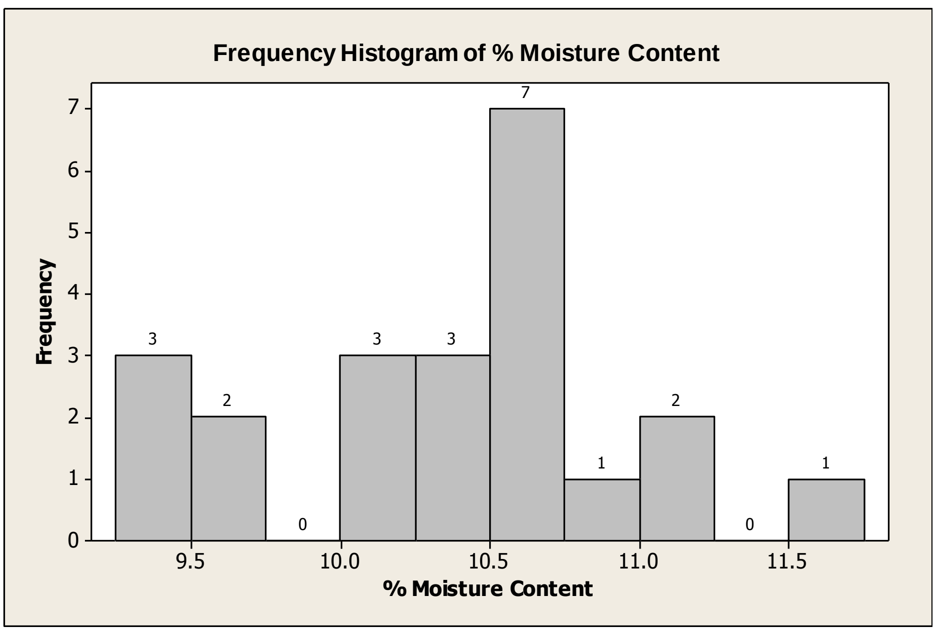
**b.** The above plot is Qz vs % Moisture content. 1.59614 = slope $\approx \frac{1}{\sigma}$, or $\sigma \approx \frac{1}{1.59614} = .6265$. $\frac{-\mu}{\sigma} \approx y - intercept$ = -16.4838. So, $\mu \approx 16{,}4737 * .6265 = 10.327$.

Regressing %Moisture Content (y) vs. Qz (x) gives the estimated slope to be .601 which estimates $\sigma$ and vertical intercept of 10.327 which estimates $\mu$.

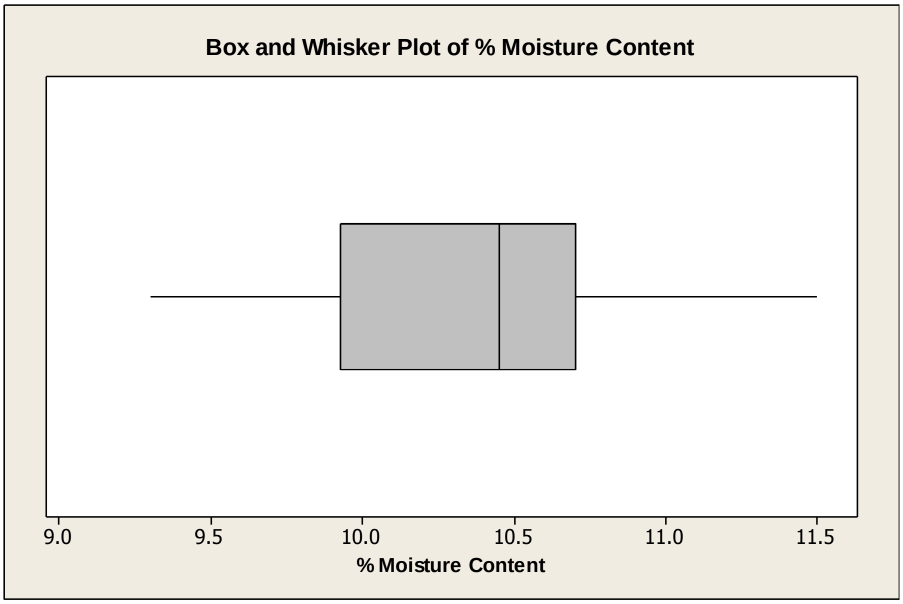
c.


Dotplot of % Moisture Content
9.3
9.6
9.9
10.2
10.5
10.8
11.1
11.4
% Moisture Content

d.


Frequency Histogram of % Moisture Content
Frequency
% Moisture Content
0
1
2
3
4
5
6
7
9.5
10.0
10.5
11.0
11.5
3
2
0
3
3
7
1
2
0
1

e.



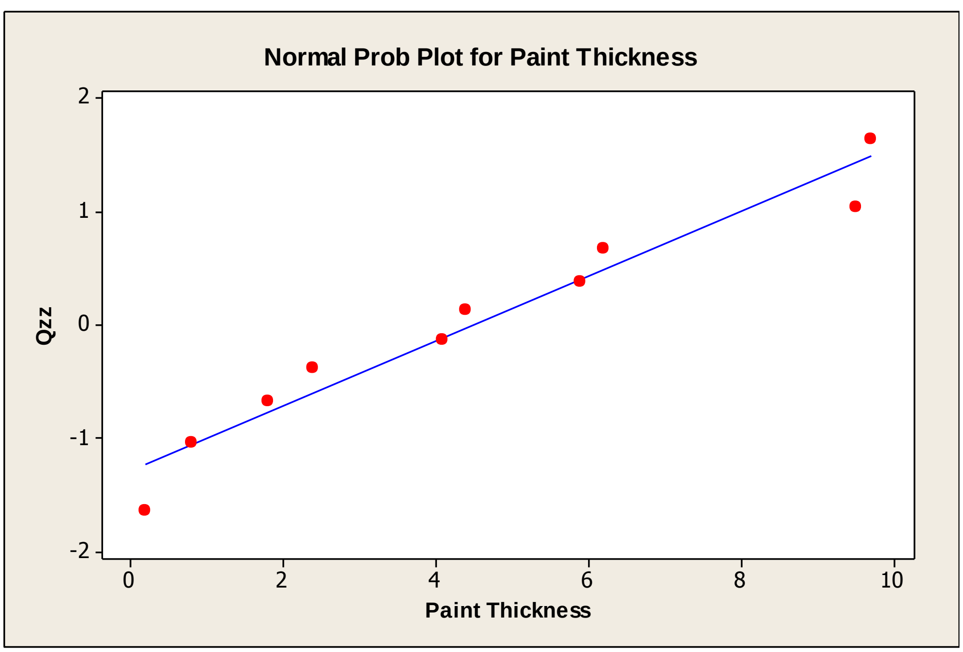

f. Looking at column of (i - .5)/22, 10 is the 25th quantile, 10.45 is the 50th quantile and 10.7 is the 75th quantile.

g. The IQR is 10.7 – 10 = .7

2. a. The time order of measured lots must be recorded.

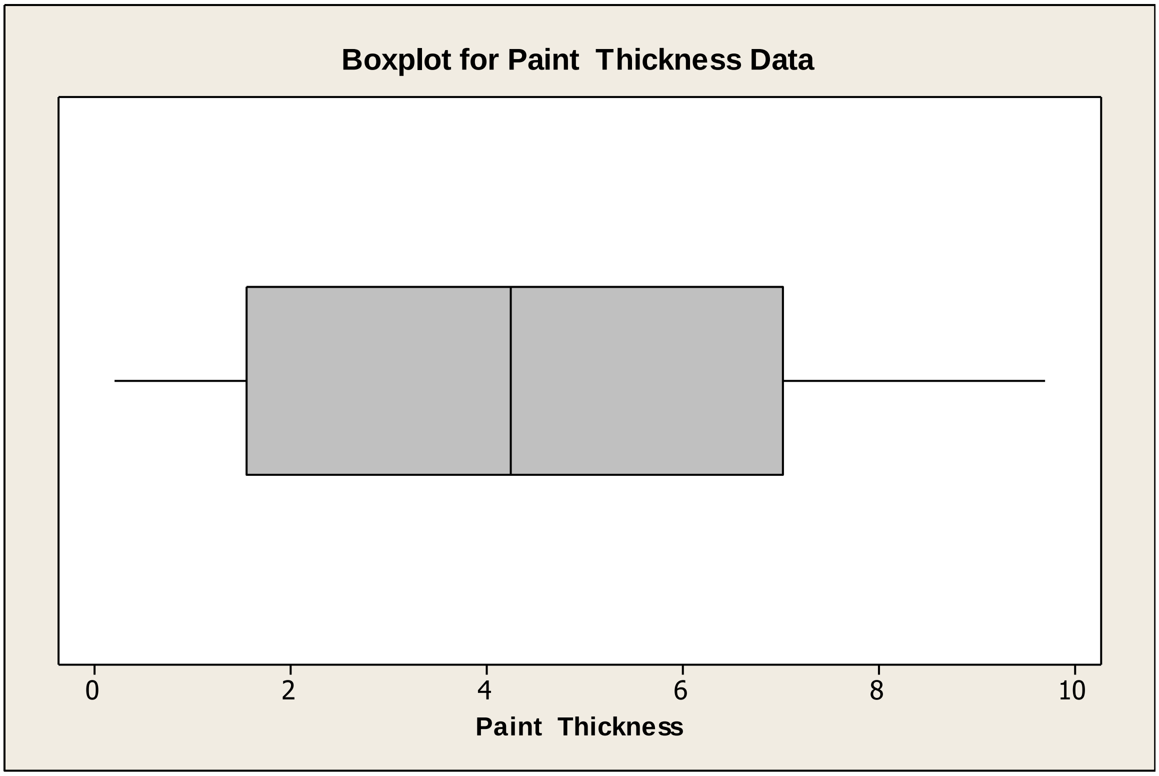
b. The % moisture content must be stable or consistent over time, no trends or cycling over time.

3. a.



A straight line appears to match the plot, implying a normal distribution of paint thickness is a reasonable assumption.

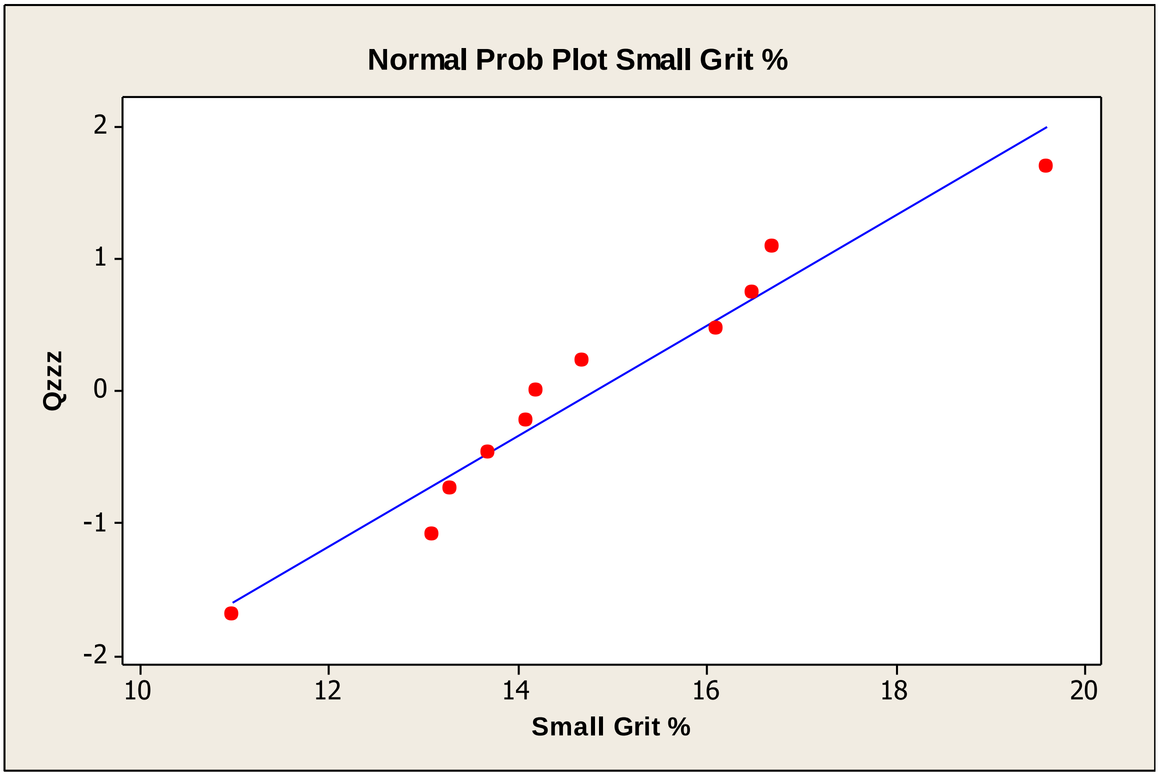
**b.**




No outliers are detected.

**c.** 25th quantile is 1.8, $50^{th}$ quantile is 4.25 and $75^{th}$ quantile is 6.2.

**4.** **a.**




No strong departures from linearity, so a normal distribution is a reasonable assumption.

**b.** $\bar{x} = 14.818$ and $s = 2.299$

**c.** 17.86 is approx. 90th quantile. .9 is 40% of the distance between .863636 and .954545. So, 17.86 is 40% of the distance from 16.7 to 19.6.

**<u>Chapter 4 Section 2</u>**

**1.** **a.** Lower 95% confidence limit = (6)(.61)$\sqrt{21/\chi^2_{21;.975}}$ = 2.82

Upper 95% confidence limit = (6)(.61) $\sqrt{21/\chi^2_{21;.025}}$ = 5.23

**b.** No, (2.82, 5.232) doesn't set within (0, 3).

**c.** Estimated $C_{pk}$ = min { (10.327 – 9)/1.83 , (12-10.327)/1.83 }
= min { .7251, .9142 } = .7251. Since this is less than

1.0, we have more than 1% outside specs.

**d.** $\frac{U-L}{6\hat{\sigma}} = \frac{3}{6(.610)} = .82$. Even if we could center the process, the "best" quality would be a $C_{pk}$ estimate of .82. Need to reduce variation, even if we center the process.

**e.** 95% C.I. for $C_{pk}$; $.7251 \pm 1.96\sqrt{\frac{1}{198} + \frac{(.7251)^2}{44-2}}$ or $.7251 \pm .25979$; (.4653, .9849).

**f.** $\left\{ \frac{U-L}{6s}\sqrt{\frac{\chi^2_{lower}}{n-1}},\ \frac{U-L}{6s}\sqrt{\frac{\chi^2_{higher}}{n-1}} \right\}$ becomes $\left\{\frac{3}{6(.61)}\sqrt{\frac{10.283}{21}},\ \frac{3}{6(.61)}\sqrt{\frac{35.479}{21}}\right\}$

(.57357, 1.0654). 95% C.I. for $C_p$. Not centered. $\bar{\bar{x}} = 10.327 \neq 10.5$.

**2.** **a.** Widen spec limits.

**b.** Same as (a).

**3.** **a.** $\left\{ \frac{U-L}{6s}\sqrt{\frac{\chi^2_{lower}}{n-1}},\ \frac{U-L}{6s}\sqrt{\frac{\chi^2_{higher}}{n-1}} \right\}$ becomes $\left\{\frac{40}{6(6.18)}\sqrt{\frac{3.325}{9}},\ \frac{40}{6(6.18)}\sqrt{\frac{16.919}{9}}\right\}$

( .6556, 1.479) 90% C.I. for $C_p$

**b.** Estimated $C_{pk}$ =

= min{ (2291.3 – 2280)/(3)(6.18), (2320 – 2291.3)/(3)(6.18) }

= min { .609, 1.548 }

= .609.

90% C.I. for $C_{pk}$

$.609 \pm 1.645\sqrt{\frac{1}{90} + \frac{(.609)^2}{18}}$ or $.609 \pm .2929$;

(.3161, .9019).

**c.** If the process can be centered, the quality will improve but still not high enough, the 90% C.I. for $C_p$ in a**.** is not completely above 1.

**4.** $\frac{\sqrt{\chi^2_{higher}}}{\sqrt{\chi^2_{lower}}} = \sqrt{45.722/16.047} = 1.688$

**5.** **a.** Lower 90% confidence limit for $6\sigma$ = (6)(3.35)$\sqrt{9/\chi^2_{9;.95}}$ = 14.659

Upper 90% confidence limit for $6\sigma$ = (6)(3.35) $\sqrt{9/\chi^2_{9;.05}}$ = 33.069

**b.** $\bar{x} = 4.5$, $s = 3.35$; No, $\bar{x} \pm s$ falls completely outside specs.

**6.** **a.** Estimated $C_p$ = 3.6/6s = .261.

**b.** Estimated $C_{pk}$ =

= min { (14.818 – 13)/3(2.299), (16.6 – 14.818)/3(2.299) }

= min { .2636, .2584 } = .2584

**c.** 95% C.I. for $C_p$;

$\{\frac{U-L}{6s}\sqrt{\frac{\chi^2_{lower}}{n-1}}, \frac{U-L}{6s}\sqrt{\frac{\chi^2_{higher}}{n-1}}\}$ becomes $\{\frac{3.6}{6(2.299)}\sqrt{\frac{3.247}{10}}, \frac{3.6}{6(2.299)}\sqrt{\frac{20.483}{10}}\}$

( .1487 , .3735) 95 % C.I. for $C_p$

95% C.I. for $C_{pk}$;

$.2584 \pm 1.96\sqrt{\frac{1}{99} + \frac{(.2584)^2}{20}}$ or $.2584 \pm .2272$;

(.0312, .4856).

**d.** Potential and present not good. Need both centering and reduce “s”.

**<u>Chapter 4 Section 3</u>**

**1.** **a.** $10.327 \pm t_{21;.975}(.61)\sqrt{1+\frac{1}{22}}$; $t_{21;.975}$ = 2.080;

$10.327 \pm 1.2973$; $(9.0297, 11.6243)$

**b.** $p$ = .95; n = 22; $1 - p^n - n(1-p)p^{n-1}$ = 1 - .323533 - .37461 = .30.
30% confident that 95% of additional lots have between 9.3% and 11.5% moisture content.

**c.** 95% sure the interval contains 95% of the moisture contents
$10.327 \pm 2.7(.61)$; $(8.68, 11.974)$

**2.** **a.** n = 10; n/(n + 1) = 10/11

**b.** 95% sure the interval contains 99% of the moisture contents
$2291.3 \pm 4.437(6.18)$; $(2263.879,\ 2318.72)$

**c.** $2291.3\ \pm t_{9;.975}(6.18)\sqrt{1+\frac{1}{10}}$; $t_{9;.975}$ = 2.262;

$2291.3\ \pm\ 14.6614$; $(2276.638\ ,\ 2305.96)$ contains length of one item with 95% probability.

**3.** **a.** 95% confidence to contain 99% of paint thicknesses.
4.5 $\pm 4.437(3.35)$; $(0,\ 19.3639)$

**b.** 9/11 or 81.81%

**c.** $p$ = .9; n = 10; $1 - p^n - n(1-p)p^{n-1}$ = 1 - .34867 - .38742 = .2639.
26.39% confident that 90% of additional paint thicknesses are between .2 and 9.7.

**4.** **a.** 14.818 $\pm t_{10;.995}(2.299)\sqrt{1+\frac{1}{11}}$; $t_{10;.995}$ = 3.169;

14.818 $\pm\ 7.609$; $(7.209\ ,\ 22.4275)$ contains lot % of small grit particles from the next lot with 99% prob.

**b.** 99% confident the interval contains % of small grit particles for 90% of lots.
14.818 $\pm 3.429(2.299)$; $(6.935,\ 22.701)$

**c.** (a) is a prediction interval; (b) is a tolerance interval

d. 95% confident 95% of all lots have % small grit particles that exceed $L$. 14.818– $2.815(2.299)$; $L =\ 8.436$

<u>**Chapter 4 Section 4**</u>

**1.** **a.** $F = kW$. $\mu_F\ \approx\ \mu_k\,\mu_W =\ .3(10) = 3.$

**b.** $\sigma_F^2 \approx\ 10^2(.01)^2 + (.3^2)(.04) =\ .01 +\ .0036 =\ .0136.$
$\sigma_F \approx\ .1166.$

**c**. $3 \pm\ 2(.1166)$; $(2.7668,\ 3.2332)$

**2.** **a.** $x_1 + x_2 + \ldots + x_{200}$; The thickness of each page is a random variable.

**b.** $\mu_1 + \mu_2 + \cdots + \mu_{200} = 2.$

**c.** $\sigma_T = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_{200}^2} = \sqrt{(.0001)^2(200)} = .001414.$

**d.** $1.96\sigma_T = 1.96(.001414) = .00277.$ $2 \pm .00277\text{inches}.$
95% of all books have thicknesses within .00277 inches of 2 inches.

**3.** **a.** $\sigma_u = \sqrt{\sigma_{x1}^2 + \sigma_{x2}^2} = \sqrt{2(.01)^2} = .01414.$

**b.** D = $\sqrt{(5+u)^2 + v^2}$ ; $\frac{\partial D}{\partial u} = \frac{1}{2}(5^2 + 10u + u^2 + v^2)^{-\frac{1}{2}}(10 + 2u)$
$\frac{\partial D}{\partial v} = \frac{1}{2}(5^2 + 10u + u^2 + v^2)^{-\frac{1}{2}}(2v).$

$(\frac{\partial D}{\partial u})^2\sigma_u^2 = .25(25+0)^{-1}10^2(.0004)$= .0004.

$(\frac{\partial D}{\partial v})^2\sigma_v^2 = .25(25+0)^{-1}(.0004)(0) = 0.$
$\sigma_D = \sqrt{.0004 + 0} = .02.$

**c**. 5.0017 $\pm\ t_{19;.95}(.0437)\sqrt{1 + \frac{1}{20}}$ ; $t_{19;.95}$ = 1.729;

$5.0017 \pm .07742;$ $(4.9243\ ,\ 5.079)$ contains next distance with 90% probability.

## Chapter 5 Section 1

**1.** $y_{ij} \sim N(\mu_i\,,\ \sigma^2)$

**2.** $s_p$ estimates $\sigma$ .

**3.** $s^2_p$ = 16.444. $s_p$ = 4.055.

**4.** $3 = 2\frac{s_p}{\sqrt{n_1}}$; $4 = 2\,\frac{s_p}{\sqrt{n_2}}$; $\frac{9}{4} + 4 = s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$; $\frac{5}{2} = \sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ ;

$\bar{y}_1 - \bar{y}_2 \pm 2(\frac{5}{2})$ or $\bar{y}_1 - \bar{y}_2 \pm 5$

**5.** **a.** 12 = 4(3) = df., $s_p$ = 3, $\bar{y}_i \pm t_{12;.975}\,\frac{3}{\sqrt{4}}$; $t_{12;.975}$ = 2.179;

$\triangle = (2.179)\frac{3}{2} = 3.2685$

**b.** $1 - 4(1\text{-}.95) \leq \gamma$

*where $\gamma$ is the group or simultaneous confidence level. So,* $\gamma \geq .80$

because of Bonferroni. So, no, confidence is not 95% all include the parameters of interest. Individually they are 95% confident but simultaneous inclusion reduces the “family- wise” confidence because the confidence must take into account the joint multiple confidences all occurring at once.

**6.** **a.** $s^2_p$ = 45.585; $s_p$ = 6.7516.

**b.** 35 $\pm\ t_{8;.975}(6.7516/\sqrt{3})$ ; $t_{8;.975}$ = 2.306;

$35 \pm 8.988$; $(26.011\ ,\ 43.988)$; 95% confidence

**c.** $\bar{y}_4 - \bar{y}_1 \pm t_{8;.975}\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ becomes $-15 \pm 12.712$ or (-27.71, -2.288)

**d.** $\frac{1}{2}\,(\mu_2 + \mu_4) - \frac{1}{2}(\mu_3 + \mu_5)$ corresponds to Design X minus Design Y.

**e**. Individual measurements for each design come from different prototypes. This permits legitimate inference to performance of new or old prototypes for the given designs.

**7.** **a**. -22.5 $\pm$ 9.063 or (-31.563, -13.437); Design X minus Design Y.

**b.** $\binom{5}{2} = 10$.
**c.** $1 - 10(1 - \alpha) = 1 - 10(1 - .99) \geq .90;\ \gamma \geq .90$. So, each must be 99%.

## Chapter 5 Section 2

**1.** Two-way factorial implies there are two factors, each having perhaps a different number of levels. Perhaps one factor is Pressure and the other is Moisture. Pressure could be at, say, hi, med or low and Moisture could be at 2%, 4%, 6%, 8%, 10%. So a total of 15 = 3 X 5 treatment combinations.

**2.** **a.** $\bar{y} + a_3 + b_3 + ab_{33} = 10 + -5 + -1 + -1 = 3$

**b.** $\Delta = t_{9;.975}(2)\sqrt{\frac{4}{(3)(3)(2)}} = 2.1326$; $t_{9;\,.975} = 2.262$.
**c.** Yes, some departure from parallelism because some $\widehat{\alpha\beta}_{ij}$ *exceed* 2.1326 *in* absolute value, i.e., significant interaction exists.

**3.** **a.** No, only A effect at a selected level of B. The simple effects of A are different from level to level of B. It is possible both "simple" effects of A are of the same "sign", meaning one could average both simple effects and make a general inference about an A effect independent of level of B.
**b.** Yes, we have an interaction effect so if there is an A effect it changes for different levels of B.
**c.** No, we have interaction implying simple effects of A change for different levels of B.

**4.** **a.** No, the .0008 inch std. dev. is understandably smaller than the $s_p$ = .0017 because the .0008 value came from repeat measurements on the same item, whereas the $s_p$ value came from measurements on different copies of the same CAD drawing pooled across different machine/enlargements.

**b.** -.0018, .0042, .0012, -.0008, -.0028

**c.** Yes, both fit and assumed common variance can be evaluated.

**d.** $ab_{13}$ = .00438, $ab_{23}$ = -.00302, $ab_{33}$ = -.00136, $ab_{32}$ = .00104, $ab_{31}$ = .00032

**e.** $\Delta = t_{36;.975}(.0017)\sqrt{\frac{4}{(3)(3)(5)}}$ = .001029; $t_{36;\,.975}$ = 2.03. Yes, the absolute value of most estimated interaction effects exceed .001029**.**

**f.** Estimated $\alpha_1 - \alpha_2 = .00799$. Estimated std. dev. Is (.0017)(6/45)$^{.5}$ = **.**0006208. So, $\Delta = 2.03(.0006208) = .00126$. .00799 $\pm$ $.00126$ or (.00673, .00925); Not credible to use for every enlargement level because important interaction exists.

**g.** Plot not given here.

**<u>Chapter 5 Section 3</u>**

**1.** $2^5$ = 32 treatments

**2.** **a.** $\hat{\mu} = 118.125$; $\hat{\alpha}_2 = 11.375$; $\hat{\beta}_2 = -14.625$; $\widehat{\alpha\beta}_{22} = -6.875$; $\hat{\gamma}_2 = 55.125$; $\widehat{\alpha\gamma}_{22} = -3.125$; $\widehat{\beta\gamma}_{22} = 1.375$; $\widehat{\alpha\beta}\gamma_{222} = 4.625$.

**b.** $\Delta = t_{56;.975}\,(7.2)\frac{1}{2^3}\sqrt{\frac{8}{8}} = \frac{(2.005)(7.2)}{8} = 1.8045$. All fitted effects are significant except $\beta\gamma_{22}$.

**c.** Grand Avg**. +** Hi A (11.375) + Lo B (14.625) + Hi C (55.125) + HiA/LoB(6. 875) = 118.125 + 88 = 206.125.

**3.** $\hat{\alpha}_2 = 3$; $\hat{\beta}_2 = 1$;

**4.** 5 cycles, divided by 32.

**5.** $\Delta = t_{24;.975}\left(s_p\right)\frac{1}{2^3}\sqrt{\frac{8}{4}} = 2.064\left(\frac{1}{8}\right)s_p\sqrt{2} = .3648\,s_p$

**6.** (a)$(b)$

**7.** (a)

**8.** (a) (b) (c)

**Chapter 6 Section 1**

**1.** **a.** $2^3 = 8$

**b.** $\Delta = t_{3;.975}\, s_p\, \frac{1}{2^3}\sqrt{\frac{3}{2}+5}$ = (3.182)(.31868)$s_p$ = 1.014$s_p$

**c.** 1 generator

**d.** $\mathrm{I} \leftrightarrow \mathrm{ABCD}$ is the defining relation and the generator is

$\mathrm{D} \leftrightarrow \mathrm{ABC}$. So, from the problem, the following are judged detectable:

$\alpha_2 + \beta\gamma\delta_{222}$
$\delta_2 + \alpha\beta\gamma_{222}$
$\beta\gamma_{22} + \alpha\delta_{22}$

Further, assuming all two-factor and higher interactions are negligible, the A effect and D effect are what is driving differences in the responses.

**2.** Hi A, Hi B, Hi C and Low D are recommended. This combination is not represented in the fractional factorial, $2^{4-1}$ where the generator is $\mathrm{D} \leftrightarrow \mathrm{ABC}$. All hi for A, B, C and D occurs in the experimental setup.

**3.** **a.** 9 factors

**b.** $2^9 = 512$ combinations

**c.** $2^{9-1} = 256$ combinations

**d.** 16

**e.** 1/32

**f.** 5 generators

**g.** 31 = $2^q - 1 = 2^5 - 1$

**4.** **a.**

| E | F | G | H | J |
|---|---|---|---|---|
| - | - | + | - | + |
| + | - | - | + | - |

**b**. CDG, DH, BJ

**c.** Assuming all two-factor and higher interactions are not important implies the following, Hi D and Hi H important influences on y.

$-6.38$ estimates $\delta_2 +$ all higher order aliases, and

$-10.13$ estimates $\alpha\delta_{22} + h_2 + a$ll other two factor and higher aliases.

Since the two-factor and higher order interactions are all assumed unimportant, $-6.38$ estimates $\delta_2$ and $-10.13$ estimates $h_2$.

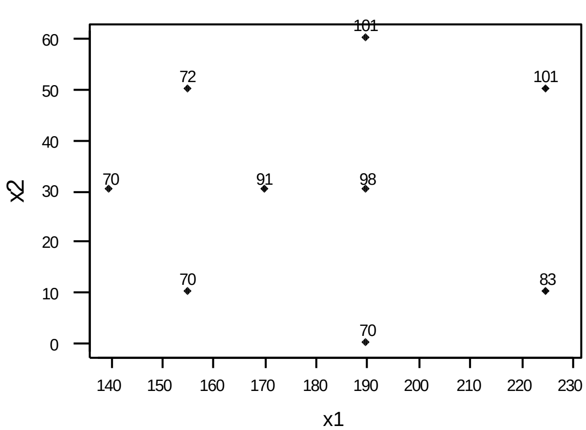
**d.** From **c**. we select Hi D, Hi H. As for the others, since all two-factor and higher interactions are judged not important, -1.25 estimates $j_2$ so select Hi level of J, -.75 estimates $e_2$ so select Hi level of E, .13 estimates $f_2$ so select Lo F, .13 estimates $g_2$ so select Lo G, 3.75 estimates $\alpha_2$ , select Lo A and 1.25 estimates $\beta_2$, select Lo B. Finally select either Hi or Lo C.

$$\hat{y} = 92\ - 3.75 + \ -1.25 + 0 + \ -6.38 - \ .75 - \ .13 - \ .13 - 10.38 - 1.25$$

$$\hat{y} = 92 - 24.02 = 67.98$$

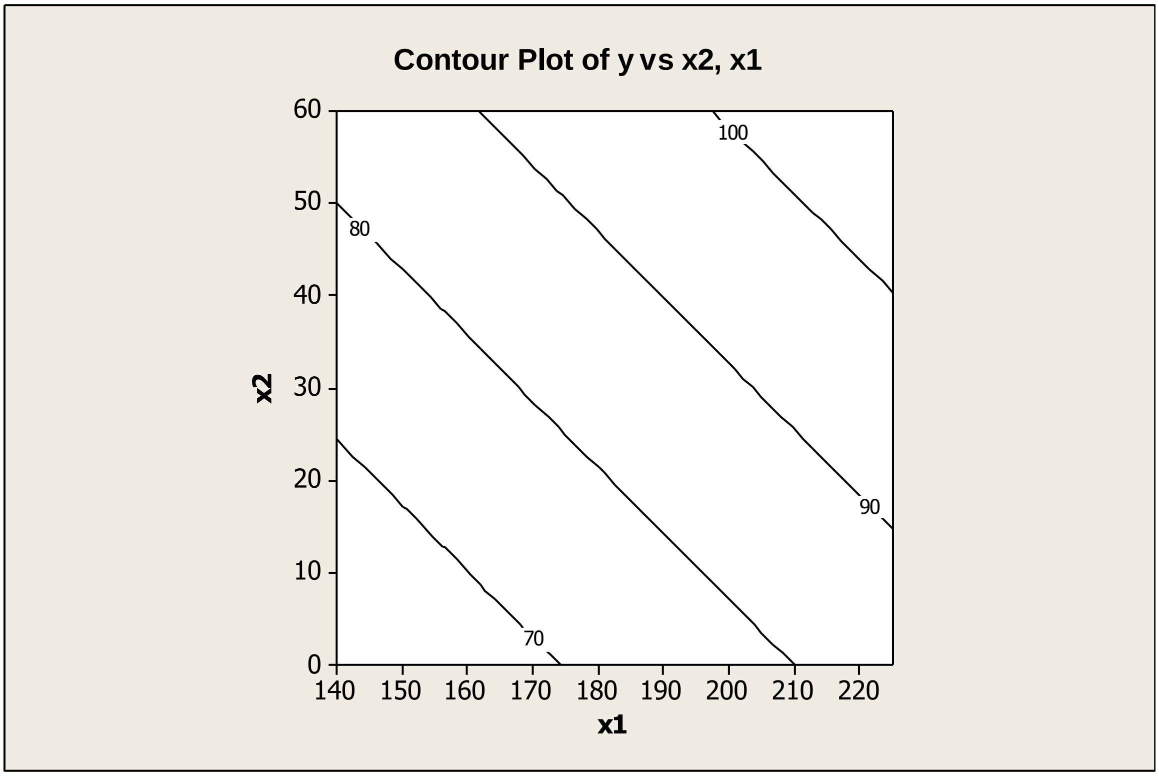
**Chapter 6 Section 2**

**1.** **a.**



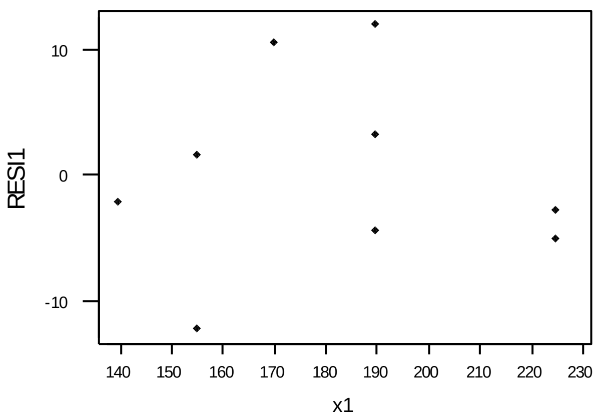
**b.** $\hat{y}(X_1, X_2) = 21.2816 + .2798X_1 + .3912X_2$

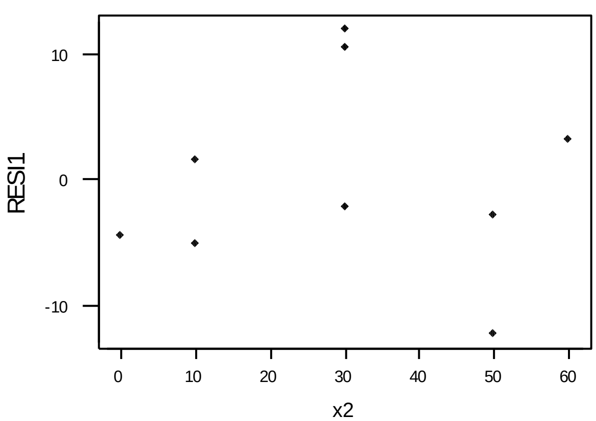
**c.**

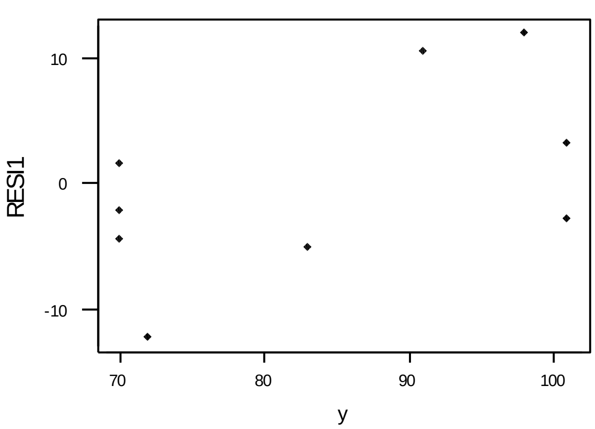



It appears the smallest predicted density occurs for ($X_1$, $X_2$) near their simultaneous minimum values over the experimental region, i.e., $X_1$ = 155 and $X_2$ = 10. The largest predicted density appears to be where ($X_1$, $X_2$) are simultaneously large within the experimental region, say, $X_1$ = 225, $X_2$ = 50.

**d.**


RESI1
10
0
-10
140
150
160
170
180
190
200
210
220
230
x1

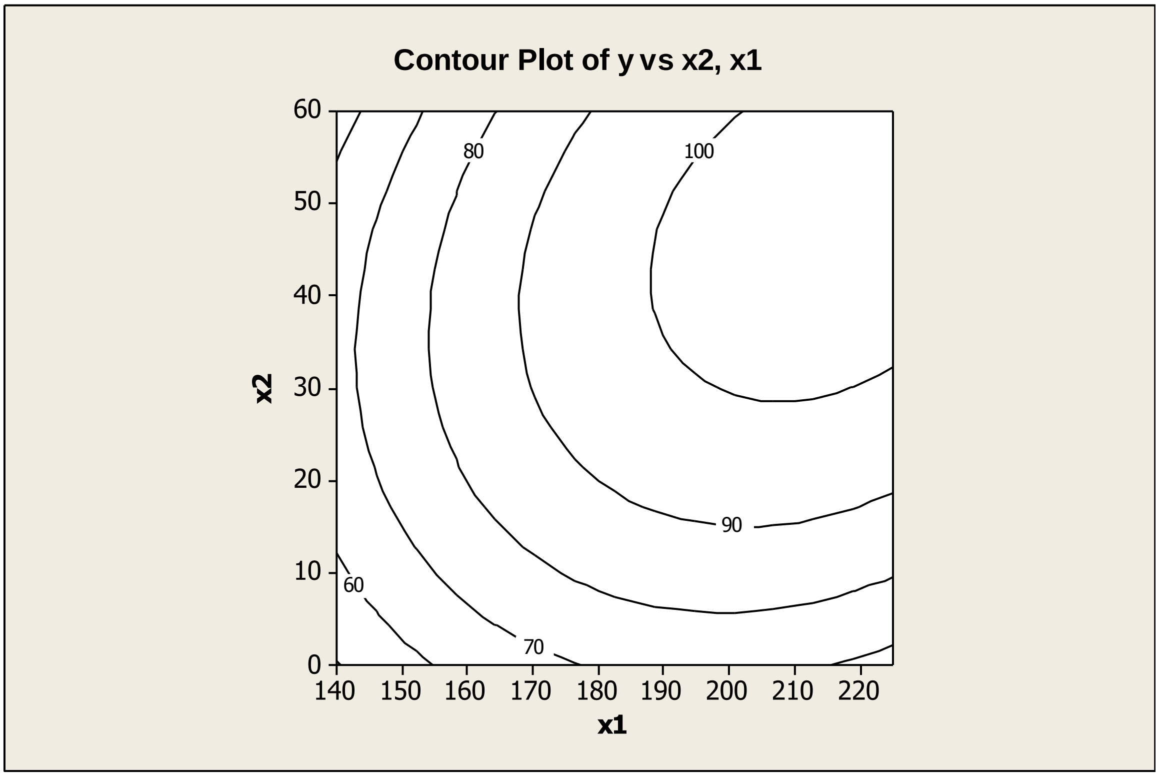

RESI1
10
0
-10
0
10
20
30
40
50
60
x2


RESI1
10
0
-10
70
80
90
100
y

Only a slight curvature is suggested. Negative, positive, negative trends of residuals vs $X_1$ or vs $X_2$ are seen.

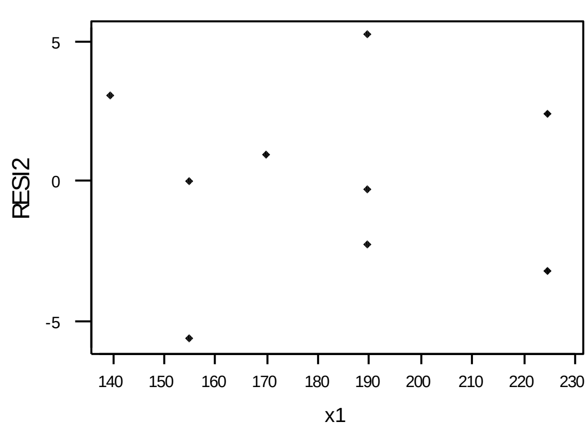
**e.** $\hat{y}(X_1, X_2) = -206.63 + 2.8424X_1 + .256X_2 + .005714X_1X_2 - .007233X_1^2 - .015842X_2^2$

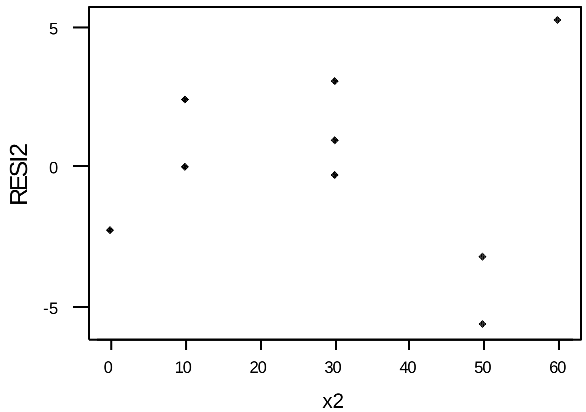
**f.**

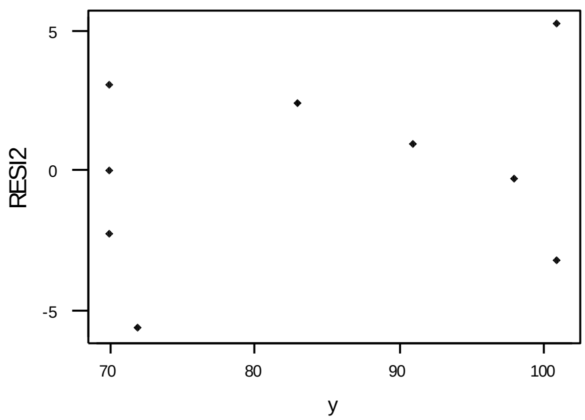



The largest predicted density is for $X_1$ close to 225 and $X_2$ close to 50.

**g.**

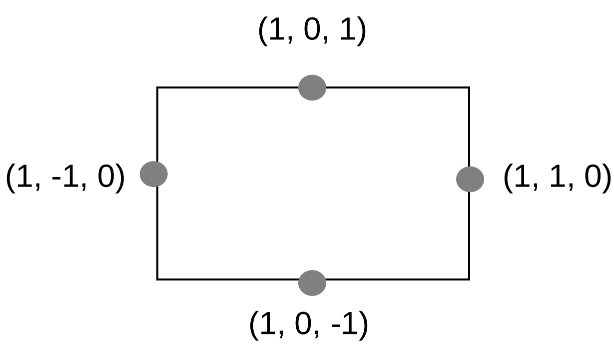





These residual plots do not show much improvement over those in (d). However, the $R^2$ for the fitted model in (e) is much larger (94.2%) than the $R^2$ for the fitted model in (b) (69.9%).

**2.** **a.** The "front" side of the cube looks like



The first ordinate is 1 = “out of the page”, 0 = “ on page” and -1 = “behind the page”. The 2nd ordinate is “left to right”, i.e. -1, 0, 1. The 3rd ordinate is “top to bottom”, i.e. 1, 0, -1.

Each of the 6 sides of the cube looks like the above sketch. The design points are located in the same relative positions. The 13th design point is the center of the cube at (0, 0, 0). The experimental region is the cube with each corner “sawed” off.

**b.** The 13 design points do not constitute a central composite design. A $2^3$ central composite requires design points at the eight $(X_1,X_2,X_3)$ distinct points such that each $X_1$, $X_2$, and $X_3$ must be 1 or -1. Further, the center point (0, 0, 0) must be included and

(2)(3) = six "star" points

$(\alpha, 0, 0)$, $(-\alpha, 0, 0)$, $(0, \alpha, 0)$, $(0, -\alpha, 0)$, $(0, 0, \alpha)$, $(0, 0, -\alpha)$.

**c.** Yes, there was replication at the center point (0, 0, 0). Three runs were taken at this point.

**d.** $\hat{y}(X_1, X_2, X_3) = 25.10 - 8.10X_1 - 15.08X_2 + 2.01X_3$

$R^2 = 80.7\%$, residual plots suggest a model that contains squared terms and cross product terms.

**e.** $$\hat{y}(X_1, X_2, X_3) = 20.1233 - 8.095X_1 - 15.0763X_2 + 2.0062X_3 + 8.275X_1X_2 + .15X_1X_3 - 1.5675X_2X_3 - 1.0679X_1^2 + 8.2696X_2^2 + 2.1196X_3^2$$

$R^2 = 99.6\%$, residual plots affirm this fit.

**f.** Confidence intervals (90% level) for the coefficients of $X_2^2$, $X_3^2$, $X_2X_3$ and $X_1X_2$ all contain values exclusive of zero. Thus, these terms were helpful to add to the model fitted in (d). Further, the $R^2$ has increased significantly

**g.** $s = 1.456$ using the full quadratic model in (e).