Answers to Section Exercises

Chapter 1

Section 1

1. Designing and improving complex products and systems often leads to situations where there is no known theory that can guide decisions. Engineers are then forced to experiment and collect data to find out how a system works, usually under time and monetary constraints. Engineers also collect data in order to monitor the quality of products and services. Statistical principles and methods can be used to find effective and efficient ways to collect and analyze such data.

2. The physical world is filled with variability. It comes from differences in raw materials, machinery, operators, environment, measuring devices, and other uncontrollable variables that change over time. This produces variability in engineering data, at least some of which is impossible to completely eliminate. Statistics must therefore address the reality of variability in data.

3. Descriptive statistics provides a way of summarizing patterns and major features of data. Inferential statistics uses a probability model to describe the process from which the data were obtained; data are then used to draw conclusions about the process by estimating parameters in the model and making predictions based on the model.

Section 2

1. Observational study—you might be interested in assessing the job satisfaction of a large number of manufacturing workers; you could administer a survey to measure various dimensions of job satisfaction. Experimental study—you might want to compare several different job routing schemes to see which one achieves the greatest throughput in a job shop.

2. Qualitative data—rating the quality of batches of ice cream as either poor, fair, good, or exceptional. Quantitative data—measuring the time (in hours) it takes for each of 1,000 integrated circuit chips to fail in a high-stress environment.

3. Any relationships between the variables x and y can only be derived from a bivariate sample.

4. You might want to compare two laboratories in their ability to determine percent impurities in rare metal specimens. Each specimen could be divided in two, with each half going to a different lab. Since each specimen is being measured twice for percent impurity, the data would be paired (according to specimen).

5. Full factorial data structure—tests are performed for all factor-level combinations:

Design	Paper	Loading Condition
delta	construction	with clip
t-wing	construction	with clip
delta	typing	with clip
t-wing	typing	with clip
delta	construction	without clip
t-wing	construction	without clip
delta	typing	without clip
t-wing	typing	without clip

Fractional factorial data structure—tests are performed for only some of the possible factor-level combinations. One possibility is to choose the following "half fraction":

Design	Paper	Loading Condition
delta	construction	without clip
t-wing	construction	with clip
delta	typing	with clip
t-wing	typing	without clip

6. Variables can be manipulated in an experiment. If changes in the response coincide with changes in factor levels, it is usually safe to infer that the changes in the factor caused the changes in the response (as long as other factors have been controlled and there is no source of bias). There is no control or manipulation in an observational study. Changes in the response may coincide with changes in another variable, but there is always the possibility that a *third* variable is causing the correlation. It is therefore risky to infer a cause-and-effect relationship between any variable and the response in an observational study.

Section 3

1. Even if a measurement system is accurate and precise, if it is not truly measuring the desired dimension or characteristic, then the measurements are useless. If a measurement system is valid and accurate, but imprecise, it may be useless because it produces too much variability (and this cannot be corrected by calibration). If a measurement system is valid and precise, but inaccurate, it might be easy to make it accurate (and thus useful) by calibrating it to a standard.

2. If the measurement system is not valid, then taking an average will still produce a measurement that is invalid. If the individual measurements are inaccurate, then the average will be inaccurate. Averaging many measurements only improves precision. Suppose that the long-run average yield of the process is stable over time. Imagine making 5 yield measurements every hour, for 24 hours. This produces 120 individual measurements, and 24 averages. Since the averages are "pulled" to the center, there will be less variability in the 24 averages than in the 120 individual measurements, so averaging improves precision.

3. Unstable measurement systems (e.g., instrument drift, multiple inconsistent devices) can lead to differences or changes in validity, precision, and accuracy. In a statistical engineering study, it is important to obtain valid, precise, and accurate measurements throughout the study. Changes or differences may create excessive variability, making it hard to draw conclusions. Changes or differences can also bias results by causing patterns in data that might incorrectly be attributed to factors in the experiment.

Section 4

1. Mathematical models can help engineers describe (in a relatively simple and concise way) how physical systems behave, or will behave. They are an integral part of designing and improving products and processes.

Chapter 2

Section 1

1. *Flight distance* might be defined as the horizontal distance that a plane travels after being launched from a mechanical slingshot. Specifically, the horizontal distance might be measured from the point on the floor directly below the slingshot to the point on the floor where any part of the plane first touches.

2. If all operators are trained to use measuring equipment in the same consistent way, this will result in better repeatability and reproducibility of measurements. The measurements will be more repeatable because individual operators will use the same technique from measurement to measurement, resulting in small variability among measurements of the same item by the same operator. The measurements will be more reproducible because all operators will be trained to use the same technique, resulting in small variability among measurements made by different operators.

3. This scheme will tend to "over-sample" larger lots and "under-sample" smaller lots, since the amount of information obtained about a large population from a particular sample size does not depend on the size of the population. To obtain the same amount of information from each lot, you should use an absolute (fixed) sample size instead of a relative one.

4. If the response variable is poorly defined, the data collected may not properly describe the characteristic of interest. Even if they do, operators may not be consistent in the way that they measure the response, resulting in more variation.

Section 2

1. Label the 38 runout values consecutively, 1–38, in the order given in Table 1.1 (smallest to largest). First sample labels: {12, 15, 5, 9, 11}; First sample runout values: {11, 11, 9, 10, 11}. Second sample labels: {34, 31, 36, 2, 14}; Second sample runout values: {17, 15, 18, 8, 11}. Third sample labels: {10, 35, 12, 27, 30}; Third sample runout values: {10, 17, 11, 14, 15}. Fourth sample labels: {15, 5, 19, 11, 8}; Fourth sample runout values: {11, 9, 12, 11, 10}. The samples are not identical. *Note:* the population mean is 12.63; the sample means are 10.4, 13.8, 13.4, and 10.6.

3. A simple random sample is not guaranteed to be representative of the population from which it is drawn. It gives every set of n items an equal chance of being selected, so there is always a chance that the n items chosen will be "extreme" members of the population.

Section 3

1. Possible controlled variables: operator, launch angle, launch force, paper clip size, paper manufacturer, plane constructor, distance measurer, and wind. The response is Flight Distance and the experimental variables are Design, Paper Type, and Loading Condition. Concomitant variables might be wind speed and direction (if these cannot be controlled), ambient temperature, humidity, and atmospheric pressure.

2. Advantage: may reduce baseline variation (background noise) in the response, making it easier to see the effects of factors. Disadvantage: the variable may fluctuate in the real world, so controlling it makes the experiment more artificial—it will be harder to generalize conclusions from the experiment to the real world.

3. Treat "distance measurer" as an experimental (blocking) variable with 2 levels. For each level (team member), perform a full factorial experiment using the 3 primary factors. If there are differences in the way team members measure distance, then it will still be possible to unambiguously assess the effects of the primary factors within each "sub-experiment" (block).

4. List the tests for Mary in the same order given for Exercise 5 of Section 1.2. Then list the tests for Tom after Mary, again in the same order. Label the tests consecutively 1–16, in the order listed. Let the digits 01–05 refer to test 1, 06–10 to test 2, . . . , and 76–80 to test 16. Move through Table B.1 choosing two digits at a time. Ignore previously chosen test labels or numbers between 81 and 00. Order the tests in the same order that their corresponding two-digit numbers are chosen from the table. Using this method (and starting from the upper-left of the table), the test labeled 3 (Mary, delta, typing, with clip) would be first, followed by the tests labeled 13, 9, 1, 2, 7, 10, 8, 14, 11, 6, 15, 4, 16, 12, and 5.

5. For the delta/construction/with clip condition (for example), flying the same plane twice would provide information about flight-to-flight variability for that particular plane. This would be useful if you are only interested in making conclusions about that particular plane. If you are interested in generalizing your conclusions to all delta design planes made with construction paper and loaded with a paper clip, then reflying the same airplane does not provide much more information. But making and flying two planes for this condition would give you some idea of variability among different planes of this type, and would therefore validate any general conclusions made from the study. This argument would be true for all 8 conditions, and would also apply to comparisons made among the 8 conditions.

6. Random sampling is used in enumerative studies. Its purpose is to choose a representative sample from some population of items. Randomization is used in analytical/experimental studies. Its purpose is to assign units to experimental conditions in an unbiased way, and to order procedures to prevent bias from unsupervised variables that may change over time.

7. Blocking is a way of controlling an extraneous variable: within each block, there may be less baseline variation (background noise) in the response than there would be if the variable were not controlled. This makes it easier to see the effects of the factors of interest within each block. Any effects of the extraneous variable can be isolated and distinguished from the effects of the factors of interest. Compared to holding the variable constant throughout the experiment, blocking also results in a more realistic experiment.

8. Replication is used to estimate the magnitude of baseline variation (background noise, experimental error) in the response, and thus helps sharpen and validate conclusions drawn from data. It provides verification that results are repeatable and establishes the limits of that repeatability.

9. It is not necessary to know exactly how the entire budget will be spent. Experimentation in engineering is usually sequential, and this requires some decisions to be made in the middle of the study. Although some may think that this is improper from a scientific/statistical point of view, it is only practical to base the design of later stages on the results of earlier stages.

Section 4

1. If you regard student as a blocking variable, then this would be a randomized complete block experiment. Otherwise, it would just be a completely randomized experiment (with a full factorial structure).

2. **(a)** Label the 24 runs as follows:

Labels	Level of A	Level of B	Level of C
1, 2, 3	1	1	1
4, 5, 6	2	1	1
7, 8, 9	1	2	1
10, 11, 12	2	2	1
13, 14, 15	1	1	2
16, 17, 18	2	1	2
19, 20, 21	1	2	2
22, 23, 24	2	2	2

Use the following coding for the test labels: table number 01–04 for test label 1, table number 05–08 for test label 2, . . . , table number 93–96 for test number 24. Move through Table B.1 choosing two digits at a time, ignoring numbers between 97 and 00 and those corresponding to test labels that have already been picked. Order the tests in the same order that their corresponding two-digit numbers are picked from the table. Using this method, and starting from the upper-left corner of the table, the order would be 3, 4, 24, 16, 11, 2, 9, 12, 17, 8, 21, 1, 13, 7, 18, 5, 20, 14, 19, 15, 22, 23, 6, 10. **(b)** Treat day as a blocking variable, and run each of the 8 factor-level combinations once on each day. Blocking allows comparisons among the factor-level combinations to be made within each day. If blocking were not used, differences among days might cause variation in the response which would cloud comparisons among the factor-level combinations. **(c)** List the 8 factor-level combinations separately for each day. For each day, label the runs as follows:

Label	Level of A	Level of B	Level of C
1	1	1	1
2	2	1	1
3	1	2	1
4	2	2	1
5	1	1	2
6	2	1	2
7	1	2	2
8	2	2	2

For each day, move through Table B.1 one digit at a time ignoring the digits 9 and 0 and any that have already been picked. Order the 8 runs in the same order that the numbers were picked from the table. Starting from where I left off in part (a), the order for day 1 is 5, 3, 8, 4, 1, 2, 6 (which implies that run 7 goes last). For day 2, the order is 5, 1, 8, 7, 2, 3, 6 (which implies that run 4 goes last). For day 3, the order is 1, 3, 2, 7, 4, 5, 8, (which implies that run 6 goes last).

The plan is summarized below:

Day	Level of A	Level of B	Level of C
1	1	1	2
1	1	2	1
1	2	2	2
1	2	2	1
1	1	1	1
1	2	1	1
1	2	1	2
1	1	2	2
2	1	1	2
2	1	1	1
2	2	2	2
2	1	2	2
2	2	1	1
2	1	2	1
2	2	1	2
2	2	2	1
3	1	1	1
3	1	2	1
3	2	1	1
3	1	2	2
3	2	2	1
3	1	1	2
3	2	2	2
3	2	1	2

Part (a) randomized all 24 runs together; here, each block of 8 runs is randomized separately.

3. The factor Person is the "block" variable.

Block	Design	Paper
Tom	delta	construction
Tom	t-wing	typing
Juanita	delta	typing
Juanita	t-wing	construction

4. Focusing on Design, you would want each person to test two delta-wing planes and two t-wing planes; this would allow you to clearly compare the two designs. You could separately compare the designs "within" each person. If possible, you would want a plan such that this is true for all three primary factors, simultaneously. This is possible by using the same pattern that is used in Table 2.6:

Person	Design	Paper	Loading Condition
Juanita	delta	construction	with clip
Tom	t-wing	construction	with clip
Tom	delta	typing	with clip
Juanita	t-wing	typing	with clip
Tom	delta	construction	without clip
Juanita	t-wing	construction	without clip
Juanita	delta	typing	without clip
Tom	t-wing	typing	without clip

This design also allows each person to test each Design/Paper combination once, each Design/

Loading combination once, and each Paper/Loading combination once.

5. This is an incomplete block experiment.

Section 5

1. A cause-and-effect diagram may be useful for representing a complex system in a relatively simple and visual way. It enables people to see how the components of the system interact, and may help identify areas which need the most attention/improvement.

Chapter 3

Section 1

1. One choice of intervals for the frequency table and histogram is 65.5–66.4, 66.5–67.4, ..., 73.5–74.4. For this choice, the frequencies are 3, 2, 9, 5, 8, 6, 2, 3, 2; the relative frequencies are .075, .05, .225, .125, .2, .15, .05, .075, .05; the cumulative relative frequencies are .075, .125, .35, .475, .675, .825, .875, .95, 1. The plots reveal a fairly symmetric, bell-shaped distribution.

2. The plots show that the depths for 200 grain bullets are larger and have less variability than those for the 230 grain bullets.

3. **(a)** There are no obvious patterns. **(b)** The differences are -15, 0, -20, 0, -5, 0, -5, 0, -5, 20, -25, -5, -10, -20, and 0. The dot diagram shows that most of the differences are negative and "truncated" at zero. The exception is the tenth piece of equipment, with a difference of 20. This point does not fit in with the shape of the rest of the differences, so it is an outlier. Since most of the differences are negative, the bottom bolt generally required more torque than the top bolt.

Section 2

1. **(a)** For the lengthwise sample: $Median = .895$, $Q(.25) = .870$, $Q(.75) = .930$, $Q(.37) = .880$. For the crosswise sample: $Median = .775$, $Q(.25) = .690$, $Q(.75) = .800$, $Q(.37) = .738$. **(b)** On the whole, the impact strengths are larger and more consistent for lengthwise cuts. Each method also produced an unusual impact strength value (outlier). **(c)** The nonlinearity of the Q–Q plot indicates that the overall shapes of these two data sets are not the same. The lengthwise cuts had an unusually large data point ("long right tail"), whereas the crosswise cuts had an unusually small data point ("long left tail"). Without these two outliers, the data sets would have similar shapes, since the rest of the Q–Q plot is fairly linear.

2. Use the $(i - .5)/n$ quantiles for the smaller data set. The plot coordinates are: (.370, .907), (.520, 1.22), (.650, 1.47), (.920, 1.70), (2.89, 2.45), (3.62, 5.89).

3. The first 3 plot coordinates are: (65.6, -2.33), (65.6, -1.75), (66.2, -1.55). The normal plot is quite linear, indicating that the data are very bell-shaped.

4. Theoretical Q–Q plotting allows you to roughly check to see if a data set has a shape that is similar to some theoretical distribution. This can be useful in identifying a theoretical (probability) model to represent how the process is generating data. Such a model can then be used to make inferences (conclusions) about the process.

Section 3

1. For the lengthwise cuts: $\bar{x} = .919$, $Median = .895$, $R = .310$, $IQR = .060$, $s = .088$. For the crosswise cuts: $\bar{x} = .743$, $Median = .775$, $R = .430$, $IQR = .110$, $s = .120$. The sample means and medians show that the center of the distribution for lengthwise cuts is higher than the center for crosswise cuts. The sample ranges, interquartile ranges, and sample standard deviations show that there is less spread in the lengthwise data than in the crosswise data.

2. These values are statistics. They are summarizations of two samples of data, and do not represent exact summarizations of larger populations or theoretical (long-run) distributions.

4. In the first case, the sample mean and median increase by 1.3, but none of the measures of spread change; in the second case, all of the measures double.

Section 4

1. $\hat{p}$ = the proportion of part orders that are delivered on time to the factory floor. $\hat{u}$ = number of defects per shift produced on an assembly line. A measured value of 65% yield for a run of a chemical process is of neither form.

2. $\hat{p}_{\text{Laid}} = \frac{6}{38} = .158$. $\hat{p}_{\text{Hung}} = \frac{24}{39} = .615$. Most engineering situations call for minimizing variation. The $\hat{p}$ values do not give any indication of how much spread there is in each set of data, and would not be helpful in comparing the two methods with respect to variation.

3. Neither type. These rates represent continuous measurements on each specimen; there is no "counting" involved.

Chapter 4

Section 1

1. **(a)** $\hat{y} = 9.4 - 1.0x$ **(b)** $r = -.945$ **(c)** $r = .945$. This is the negative of the r in part (b), since the $\hat{y}$'s are perfectly negatively correlated with the x's. **(d)** $R^2 = .893 = r^2$ from both (b) and (c). **(e)** $-.4$, $.6$, $-.4$, $.6$, $-.4$. These are the vertical distances from each data point to the least squares line.

3. **(a)** $R^2 = .994$ **(b)** $\hat{y} = -3174.6 + 23.50x$. 23.5 **(c)** Residuals: 105.36, −21.13, −60.11, −97.58, 16.95, 14.48, 42.00, .02. **(d)** There is no replication (multiple experimental runs at a particular pot temperature). **(e)** For $x = 188$ °C, $\hat{y} = 1243.1$. For $x = 200$ °C, $\hat{y} = 1525.1$. It would not be wise to make a similar prediction at $x = 70$ °C because there is no evidence that the fitted relationship is correct for pot temperatures as low as $x = 70$ °C. Some data should be obtained around $x = 70$ °C.

4. **(a)** The scatterplot is not linear, so the given straight-line relationship does not seem appropriate. $R^2 = .723$. **(b)** This scatterplot is much more linear, and a straight-line relationship seems appropriate for the transformed variables. $R^2 = .965$. **(c)** $\widehat{\ln y} = 34.344 - 5.1857 \ln x$. For $x = 550$, $\widehat{\ln y} = 1.6229$ so $\hat{y} = e^{1.6229} = 5.07$ minutes. The implied relationship between x and y is $y = e^{\beta_0} x^{\beta_1}$.

Section 2

1. $\hat{y} = -1315 + 5.6x + .04212x^2$. $R^2 = .996$. For the quadratic model, at $x = 200$ °C, $\hat{y} = 1487.2$, which is relatively close to 1525.1 from part (e) of Exercise 3 of Section 1.

2. **(a)** $\hat{y} = 6.0483 + .14167x_1 - .016944x_2$. $b_1 = .14167$ means that as x_1 increases by 1% (holding x_2 constant), y increases by roughly .142 cm^3/g. $b_2 = -.016944$ means that as x_2 increases by one minute (holding x_1 constant), y decreases by roughly .017 cm^3/g. $R^2 = .807$. **(b)** The residuals are −.015, .143, .492, −.595, −.457, −.188, .695, .143, −.218. **(c)** For $x_2 = 30$, the equation is $\hat{y} = 5.53998 + .14167x_1$. For $x_2 = 60$, the equation is $\hat{y} = 5.03166 + .14167x_1$. For $x_2 = 90$, the equation is $\hat{y} = 4.52334 + .14167x_1$. The fitted responses do not match up well, because the relationship between y and x_1 is not linear for any of the x_2 values. **(d)** At $x_1 = 10\%$ and $x_2 = 70$ minutes, $\hat{y} = 6.279$ cm^3/g. It would not be wise to make a similar prediction at $x_1 = 10\%$ and $x_2 = 120$ minutes because there is no evidence that the fitted relationship is correct under these conditions. Some data should be obtained around $x_1 = 10\%$ and $x_2 = 120$ minutes. **(e)** $\hat{y} = 4.98 + .260x_1 + .00081x_2 - .00197x_1x_2$, and $R^2 = .876$. The increase in R^2 from .807 to .876 is not very large; using the more complicated equation may not be desirable (this is subjective). **(f)** For $x_2 = 30$, the equation is $\hat{y} = 5.0076 + .20084x_1$. For $x_2 = 60$, the equation is $\hat{y} = 5.0319 + .14168x_1$. For $x_2 = 90$, the equation is $\hat{y} = 5.0562 + .08252x_1$. The new model allows there to be a different slope for different values of x_2, so these lines fit the data better than the lines in part (c). But they still do not account for the nonlinearity between x_1 and y. An x_1^2 term should be added to the model. **(g)** There is no replication (multiple experimental runs at a particular NaOH/Time combination). **(h)** These data have a complete (full) factorial structure. The straight-line least squares equation for x_1 is

$\hat{y} = 5.0317 + .14167x_1$ with a corresponding R^2 of .594. The straight-line least squares equation for x_2 is $\hat{y} = 7.3233 - .01694x_2$ with a corresponding R^2 of .212. The slopes in these one-variable linear equations are the same as the corresponding slopes in the two variable equation from (a). The R^2 value in (a) is the sum of the R^2 values from the two one-variable linear equations.

Section 3

1. **(a)** Labeling x_1 as A and x_2 as B, $a_1 = -.643$, $a_2 = -.413$, $a_3 = 1.057$, $b_1 = .537$, $b_2 = -.057$, $b_3 = -.480$, $ab_{11} = -.250$, $ab_{12} = -.007$, $ab_{13} = .257$, $ab_{21} = -.210$, $ab_{22} = .013$, $ab_{23} = .197$, $ab_{31} = .460$, $ab_{32} = -.007$, $ab_{33} = -.453$. The fitted interactions ab_{31} and ab_{33} are large (relative to fitted main effects) indicating that the effect on y of changing NaOH from 9% to 15% depends on the Time (non-parallelism in the plot). It would not be wise to use the fitted main effects alone to summarize the data, since there may be an importantly large interaction. **(b)** $\hat{y}_{11} = 6.20$, $\hat{y}_{12} = 5.61$, $\hat{y}_{13} = 5.18$, $\hat{y}_{21} = 6.43$, $\hat{y}_{22} = 5.84$, $\hat{y}_{23} = 5.41$, $\hat{y}_{31} = 7.90$, $\hat{y}_{32} = 7.31$, $\hat{y}_{33} = 6.88$. Like the plot in part (c) and unlike the plot in (f) of Exercise 2 in Section 4.2, the fitted values for each level of B (x_2) must produce parallel plots; no interactions are allowed. However, unlike parts (c) and (f) of that exercise, the current model allows these fitted values to be nonlinear in x_1 (factorial models are generally more flexible than lines, curves, and surfaces). **(c)** $R^2 = .914$. The plots of residuals versus Time and residuals versus $\hat{y}_i$ both have patterns; these show that the "main effects only" model is not accounting for the apparent interaction between the two factors. Even though R^2 is higher than both of the models in Exercise 2 of Section 4.2, this model does not seem to be adequate.

2. **(a)** $\bar{y}_{...} = 20.792$, $a_2 = .113$, $b_2 = -13.807$, $ab_{22} = -.086$, $c_2 = 7.081$, $ac_{22} = -.090$, $bc_{22} = -6.101$, $abc_{222} = .118$. Other fitted effects can be obtained by appropriately changing the signs of the above. The simplest possible interpretation is that Diameter, Fluid, and their interaction are the only effects on Time. **(b)** $\bar{y}_{...} = 2.699$, $a_2 = .006$, $b_2 = -.766$, $ab_{22} = -.003$, $c_2 = .271$, $ac_{22} = -.003$, $bc_{22} = -.130$, $abc_{222} = .007$. Yes, but the Diameter $\times$ Fluid interaction still seems to be important. **(c)** In standard order, the fitted values are 3.19, 3.19, 1.66, 1.66, 3.74, 3.74, 2.20, 2.20. $R^2 = .974$. For a model with all factorial effects $(\widehat{\ln y_{ijk}} = \overline{\ln y_{ijk}})$, $R^2 = .995$. **(d)** $b_1 - b_2 = 1.532$ ln (sec) decrease; divide the .188 raw drain time by $e^{1.532}$ to get the .314 drain time. This suggests that (.188 drain time/.314 drain time) $= e^{1.532} = 4.63$; the theory predicts this ratio to be 7.78.

3. Interpolation, and possibly some cautious extrapolation, is only possible using surface-fitting methods. In many engineering situations, an "optimal" setting of quantitative factors is sought. This can be facilitated by interpolation (or extrapolation) using a surface-fitting model.

Section 4

1. Transforming data can sometimes make relationships among variables simpler. Sometimes nonlinear relationships can be made linear, or factors and response can be transformed so that there are no interactions among the factors. Transformations can also potentially make the shape of a distribution simpler, allowing the use of statistical models that assume a particular distributional shape (such as the bell-shaped normal distribution).

2. In terms of the raw response, there will be interactions, since x_1 and x_2 are multiplied together in the power law. The suggested plot of raw y versus x_1 will have different slopes for different values of x_2. This means that the effect of changing x_1 *depends* on the setting of x_2, which is one way to define an interaction.

   In terms of the log of y, there will not be interactions, since x_1 and x_2 appear additively in the equation for $\ln y$. Therefore, the suggested plot of $\ln y$ versus x_1 will have the same slope for all values of x_2. This means that the effect of changing x_1 does not depend on the setting of x_2 (there are no interactions).

Section 5

1. A deterministic model is used to describe a situation where the outcome can be almost exactly predicted if certain variables are known. A stochastic/probabilistic model is used in situations where it is not possible to predict the exact outcome. This may happen when important variables are unknown, or when no known deterministic theory can describe the situation. An example of a deterministic model is the classical Economic Order Quantity (EOQ) model for inventory control. Given constant rate of demand R, order quantity X, ordering cost P, and per unit holding cost C, the total cost per time period is $Y = P\left(\frac{R}{X}\right) + C\left(\frac{X}{2}\right)$.

Chapter 5

Section 1

1. (b) 4.1; 1.136.

2. (a) X has a binomial distribution with $n = 10$ and $p = \frac{1}{3}$. Use equation (5.3) with $n = 10$ and $p = \frac{1}{3}$. $f(0)$–$f(10)$ are .0173, .0867, .1951, .2601, .2276, .1366, .0569, .0163, .0030, .0003, .0000. (b) Assuming that they are just guessing, the chance that 7 (or more) out of 10 subjects would be correct is $P(X \geq 7) = .0197$. Under the hypothesis that they are only guessing, this kind of extreme outcome would only happen about 1 in 50 times, so the outcome is strong evidence that they are not just guessing.

3. (a) Using equations (3.4) and (3.5), $\mu = 4$, $\sigma^2 = \frac{5}{3}$, and $\sigma = 1.291$.
(b)

x	2	3	4	5	6
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Since all members of the population are equally likely to be chosen, the probability histogram for X is the same as the population relative frequency distribution. Using equations (5.1) and (5.2), $EX = 4$ and $\mathrm{Var}X = \frac{5}{3}$. **(c)** Label the values 2, 3, 4_1, 4_2, 5, 6.

First Item	Second Item	$\bar{x}$	s^2	Probability
2	3	2.5	.5	$\frac{1}{15}$
2	4_1	3.0	2.0	$\frac{1}{15}$
2	4_2	3.0	2.0	$\frac{1}{15}$
2	5	3.5	4.5	$\frac{1}{15}$
2	6	4.0	8.0	$\frac{1}{15}$
3	4_1	3.5	.5	$\frac{1}{15}$
3	4_2	3.5	.5	$\frac{1}{15}$
3	5	4.0	2.0	$\frac{1}{15}$
3	6	4.5	4.5	$\frac{1}{15}$
4_1	4_2	4.0	0	$\frac{1}{15}$
4_1	5	4.5	.5	$\frac{1}{15}$
4_1	6	5.0	2.0	$\frac{1}{15}$
4_2	5	4.5	.5	$\frac{1}{15}$
4_2	6	5.0	2.0	$\frac{1}{15}$
5	6	5.5	.5	$\frac{1}{15}$

Using the above table, the probability distribution for $\overline{X}$ is:

$\bar{x}$	2.5	3	3.5	4	4.5	5	5.5
$P(\overline{X} = \bar{x})$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{3}{15}$	$\frac{3}{15}$	$\frac{2}{15}$	$\frac{1}{15}$

Using equations (5.1) and (5.2), $E\overline{X} = 4$ and $\mathrm{Var}\overline{X} = \frac{2}{3}$. As might be expected, the mean of $\overline{X}$ is the same as the mean of X, and the variance is smaller. The probability distribution for S^2 is

s^2	0	.5	2	4.5	8
$P(S^2 = s^2)$	$\frac{1}{15}$	$\frac{6}{15}$	$\frac{5}{15}$	$\frac{2}{15}$	$\frac{1}{15}$

4. For $p = .1$, $f(0)$–$f(5)$ are .59, .33, .07, .01, .00, .00; $\mu = np = .5$; $\sigma = \sqrt{np(1-p)} = .67$. For $p = .3$, $f(0)$–$f(5)$ are .17, .36, .31, .13, .03, .00; $\mu = 1.5$; $\sigma = 1.02$. For $p = .5$, $f(0)$–$f(5)$ are .03, .16, .31, .31, .16, .03; $\mu = 2.5$; $\sigma = 1.12$. For

$p = .7$, $f(0)$–$f(5)$ are .00, .03, .13, .31, .36, .17; $\mu = 3.5$; $\sigma = 1.02$. For $p = .9$, $f(0)$–$f(5)$ are .00, .00, .01, .07, .33, .59; $\mu = 4.5$; $\sigma = .67$.

5. Binomial distribution: $n = 8$, $p = .20$. **(a)** .147 **(b)** .797 **(c)** $np = 1.6$ **(d)** $np(1 - p) = 1.28$ **(e)** 1.13

6. Geometric distribution: $p = .20$. **(a)** .08 **(b)** .59 **(c)** $1/p = 5$ **(d)** $(1 - p)/p^2 = 20$ **(e)** 4.47

7. For $\lambda = .5$, $f(0), f(1), \ldots$ are .61, .30, .08, .01, .00, .00, ...; $\mu = \lambda = .5$; $\sigma = \sqrt{\lambda} = .71$. For $\lambda = 1.0$, $f(0), f(1), \ldots$ are .37, .37, .18, .06, .02, .00, .00, ...; $\mu = 1.0$; $\sigma = 1.0$. For $\lambda = 2.0$, $f(0), f(1), \ldots$ are .14, .27, .27, .18, .09, .04, .01, .00, .00, ...; $\mu = 2.0$; $\sigma = 1.41$. For $\lambda = 4.0$, $f(0), f(1), \ldots$ are .02, .07, .15, .20, .20, .16, .10, .06, .03, .01, .00, .00, ...; $\mu = 4.0$; $\sigma = 2.0$.

8. **(a)** .323 **(b)** .368

9. **(a)** .0067 **(b)** $Y \sim \text{Binomial}(n = 4, p = .0067)$; .00027

10. Probability is a mathematical system used to describe random phenomena. It is based on a set of axioms, and all the theory is deduced from the axioms. Once a model is specified, probability provides a deductive process that enables predictions to be made based on the theoretical model.

Statistics uses probability theory to describe the source of variation seen in data. Statistics tries to create realistic probability models that have (unknown) parameters with meaningful interpretations. Then, based on observed data, statistical methods try to estimate the unknown parameters as accurately and precisely as possible. This means that statistics is inductive, using data to draw conclusions about the process or population from which the data came.

Neither is a subfield of the other. Just as engineering uses calculus and differential equations to model physical systems, statistics uses probability to model variation in data. In each case the mathematics can stand alone as theory, so calculus is not a subfield of engineering and probability is not a subfield of statistics. Conversely, statistics is not a subfield of probability just as engineering is not a subfield of calculus; many simple statistical methods do not require the use of probability, and many engineering techniques do not require calculus.

11. A relative frequency distribution is based on *data*. A probability distribution is based on a theoretical model for probabilities. Since probability can be interpreted as long-run relative frequency, a relative frequency distribution approximates the underlying probability distribution, with the approximation getting better as the amount of data increases.

Section 2

1. **(a)** 2/9 **(c)** .5

(d) $$F(x) = \begin{cases} 0 & \text{for } x \le 0 \\ \frac{10x - x^2}{9} & \text{for } 0 < x < 1 \\ 1 & \text{for } x \ge 1 \end{cases}$$

(e) 13/27; .288

2. **(a)** .2676 **(b)** .1446 **(c)** .3393 **(d)** .3616 **(e)** .3524 **(f)** .9974 **(g)** 1.28 **(h)** 1.645 **(i)** 2.17

3. **(a)** .7291 **(b)** .3594 **(c)** .2794 **(d)** .4246 **(e)** .6384 **(f)** 48.922 **(g)** 44.872. **(h)** 7.056

4. **(a)** .4938 **(b)** Set μ to the midpoint of the specifications: $\mu = 2.0000$; .7888 **(c)** .0002551

5. **(a)** $P(X < 500) = .3934$; $P(X > 2000) = .1353$ **(b)** $Q(.05) = 51.29$; $Q(.90) = 2{,}302.58$

6. **(b)** *Median* $= 68.21 \times 10^6$ **(c)** $Q(.05) = 21.99 \times 10^6$; $Q(.95) = 128.9 \times 10^6$

Section 3

1. Data that are being generated from a particular distribution will have roughly the same shape as the density of the distribution, and this is more true for larger samples. Probability plotting provides a sensitive graphical way of deciding if the data have the same shape as a theoretical probability distribution. If a distribution can be found that accurately describes the data generating process, one can then estimate probabilities and quantiles and make predictions about future process behavior based on the model.

2. Fit a line (by eye or some other method) through the points on the plot. The x-intercept is an approximate mean, and an approximate standard deviation is $\sigma \approx \frac{1}{\text{slope}} = \frac{\Delta x}{\Delta y} = \frac{\Delta \text{ data quantiles}}{\Delta \text{ std. normal quantiles}}$.

3. (b) $\mu \approx 69.5$; $\sigma \approx 1/\text{slope} = 2.1$

4. (a) First 3 coordinates of the normal plot of the raw data: (17.88, −2.05), (28.92, −1.48), (33.00, −1.23). The normal plot is not linear, so a Gaussian (normal) distribution does not seem to fit these data. First 3 coordinates of the normal plot of the natural log of the data: (2.884, −2.05), (3.365, −1.48), (3.497, −1.23). This normal plot is fairly linear, indicating that a lognormal distribution fits the data well. $\mu \approx 4.1504$, $\sigma \approx .5334$. 3.273; 26.391. **(b)** The first 3 coordinates of the Weibull plot are (2.88, −3.82), (3.36, −2.70), (3.50, −2.16). The Weibull plot is fairly linear, indicating that a Weibull distribution might be used to describe bearing load life. $\alpha \approx 81.12$, $\beta \approx 2.3$; 22.31.

5. (b) The exponential plot is fairly linear, indicating that an exponential distribution fits the data well. Since a line on the plot indicates that $Q(0) \approx 0$, no need for a threshold parameter greater than zero is indicated.

Section 4

1. If X and Y are independent, then observing the actual value of X does not in any way change probability assessments about the yet-to-be-observed Y, or vice-versa. Independence provides great mathematical simplicity in the description of the behavior of X and Y.

2. (a) For $x = 0, 1, 2$, $f_X(x) = .5, .4, .1$. For $y = 0, 1, 2, 3, 4$, $f_Y(y) = .21, .19, .26, .21, .13$. **(b)** No, since $f(x, y) \neq f_X(x)\, f_Y(y)$. **(c)** .6; .44 **(d)** 1.86; 1.74 **(e)** For $y = 0, 1, 2, 3, 4$, $f_{Y|X}(y \mid 0) = .3, .2, .2, .2, .1$; 1.6.

3. (a) For $y = 1, 2, 3, 4$, $f_{Y|X}(y \mid 0) = 0, 0, 0, 1$ and $f_{Y|X}(y \mid 1) = .25, .25, .25, .25$. $f(0, 1) = f(0, 2) = f(0, 3) = 0$, $f(0, 4) = p$, $f(1, 1) = f(1, 2) = f(1, 3) = f(1, 4) = .25(1 - p)$.
(b) $2.5 + 1.5p$ **(c)** $p > .143$

4. (a) Since X and Y are independent, $f(x, y) = f_X(x)\, f_Y(y)$ (Definition 27),

$$f(x, y) = \begin{cases} \frac{1}{.05}\frac{1}{.06} & \text{for } x \in (1.97, 2.02) \text{ and } \\ & y \in (2.00, 2.06) \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} 333.33 & \text{for } x \in (1.97, 2.02) \text{ and } \\ & y \in (2.00, 2.06) \\ 0 & \text{otherwise} \end{cases}$$

(b) Find the volume below this density over the region in which $2.00 < y < x$ and $1.97 < x < 2.02$. This is .0667. (Using calculus, this is $\int_{2.00}^{2.02} \int_{2.00}^{x} 333.33\, dy\, dx$.)

5. (a)

$$f_X(x) = \begin{cases} 2x & \text{for } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases};$$

$$f_Y(y) = \begin{cases} 2(1 - y) & \text{for } 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases};$$

$\mu = EX = 2/3$.
(b) Yes, since $f(x, y) = f_X(x)\, f_Y(y)$. **(c)** .7083
(d) $E(X|Y = .5) = 2/3$

6. (a)

$$f(x, y) = f_X(x)\, f_Y(y) = \begin{cases} e^{-x}e^{-y} & \text{if } x \geq 0 \text{ and } y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

(b) e^{-2t} **(c)** $f_T(t) = \begin{cases} 2e^{-2t} & \text{for } t \geq 0 \\ 0 & \text{otherwise} \end{cases}$.

This is an exponential distribution with mean .5.
(d) $(1 - e^{-t})^2$.

(e) $f_T(t) = \begin{cases} 2e^{-t}(1-e^{-t}) & \text{for } t \geq 0; \\ 0 & \text{otherwise} \end{cases}$.

$E(T) = 1.5$

Section 5

1. mean = .75 in.; standard deviation = .0037.

2. **(a)** Propagation of error formula gives 1.4159×10^{-6}. **(b)** The lengths.

3. **(a)** 13/27; .0576 **(b)** $\overline{X} \sim$ Normal with mean 13/27 and standard deviation .0576. **(c)** .3745 **(d)** .2736 **(e)** 13/27, .0288; $\overline{X} \sim$ Normal with mean 13/27 and standard deviation .0288; .2611; .5098.

4. .7888, .9876, 1.0000

5. Rearrange the relationship in terms of g to get $g = \frac{4\pi^2 L}{\tau^2}$. Take the given length and period to be approximately equal to the means of these input random variables. To use the propagation of error formula, the partial derivatives need to be evaluated at the means of the input random variables and $\frac{\partial g}{\partial L} = \frac{4\pi^2}{\tau^2} = 6.418837$ and $\frac{\partial g}{\partial \tau} = \frac{-8\pi^2 L}{\tau^3} = -25.8824089$. Then applying equation (5.59), $\text{Var}(g) \approx (6.418837)^2(.0208)^2 + (-25.8824089)^2 \times (.1)^2 = 6.7168$ ft^2/sec^4 so the approximate standard deviation of g is $\sqrt{6.7168} = 2.592$ ft/sec^2. The precision in the period measurement is the principal limitation on the precision of the derived g because its term (variance $\times$ squared partial derivative) contributes much more to the propagation of error formula than the length's term.

Chapter 6

Section 1

1. [6.3, 7.9] ppm is a set of plausible values for the mean. The method used to construct this interval correctly contains the true mean in 95% of repeated applications. This particular interval either contains the mean or it doesn't (there is no probability involved). However, because the *method* is correct 95% of the time, we might say that we have 95% confidence that it was correct this time.

2. **(a)** [111.0, 174.4] **(b)** [105.0, 180.4] **(c)** 167.4 **(d)** 174.4 **(e)** [111.0, 174.4] ppm is a set of plausible values for the mean aluminum content of samples of recycled PET plastic from the recycling pilot plant at Rutgers University. The method used to construct this interval correctly contains means in 90% of repeated applications. This particular interval either contains the mean or it doesn't (there is no probability involved). However, because the *method* is correct 90% of the time, we might say that we have 90% confidence that it was correct this time.

3. $n = 66$

4. **(a)** $\bar{x} = 4.6858$ and $s = .02900317$ **(b)** = [4.676, 4.695] mm **(c)** [4.675, 4.696] mm. This interval is wider than the one in (b). To increase the confidence that μ is in the interval, you need to make the interval wider. **(d)** The lower bound is 4.677 mm. This is larger than the lower endpoint of the interval in (b). Since the upper endpoint here is set to ∞, the lower endpoint must be increased to keep the confidence level the same. **(e)** To make a 99% one-sided interval, construct a 98% two-sided interval and use the lower endpoint. This was done in part (a), and the resulting lower bound is 4.676. This is smaller than the value in (d); to increase the confidence, the interval must be made "wider." **(f)** [4.676, 4.695] ppm is a set of plausible values for the mean diameter of this type of screw as measured by this student with these calipers. The method used to construct this interval correctly contains means in 98% of repeated applications. This particular interval either contains the mean or it doesn't (there is no probability involved). However, because the *method* is correct 98% of the time, we might say that we have 98% confidence that it was correct this time.

Section 2

1. $H_0: \mu = 200$; $H_a: \mu > 200$; $z = -2.98$; p-value $\doteq .9986$. There is no evidence that the mean aluminum content for samples of recycled plastic is greater than 200 ppm.

2. **(a)** $H_0: \mu = .500$; $H_a: \mu \neq .500$; $z = 1.55$; p-value $\doteq .1212$. There is some (weak) evidence that the mean punch height is not .500 in. (The rounded $\bar{x}$ and s given produce a z that is quite a bit different from what the exact values produce. $\bar{x} = .005002395$ and $s = .002604151$, computed from the raw data, produce $z = 1.85$, and a p-value of $2(.0322) = .0644$.) **(b)** [.49990, .50050] **(c)** If uniformity of stamps on the same piece of material is important, then the standard deviation (spread) of the distribution of punch heights will be important (in addition to the mean).

3. The mean of the punch heights is almost certainly not exactly equal to .50000000 inches. Given enough data, a hypothesis test would detect this as a "statistically significant" difference (and produce a small p-value). What is practically important is whether the mean is "close enough" to .500 inches. The confidence interval in part (b) answers this more practical question.

4. $H_0: \mu = 4.70$; $H_a: \mu \neq 4.70$; $z = -3.46$; p-value $\doteq .0006$. There is very strong evidence that the mean measured diameter differs from nominal.

5. Although there is evidence that the mean is not equal to nominal, the test does not say anything about how *far* the mean is from nominal. It may be "significantly" different from nominal, but the difference may be practically unimportant. A confidence interval is what is needed for determining how far the mean is from nominal.

Section 3

1. The normal distribution is bell-shaped and symmetric, with fairly "short" tails. The confidence interval methods depend on this regularity. If the distribution is skewed or prone to outliers/extreme observations, the normal-theory methods will not properly take this into account. The result is an interval whose real confidence level is different from the nominal value (and often lower than the nominal value).

2. **(a)** Independence among assemblies; normal distribution for top-bolt torques. **(b)** H_0: $\mu = 100$; H_a: $\mu \neq 100$; $t = 4.4$; p-value $\doteq .001$. There is strong evidence that the mean torque is not 100 ft lb. **(c)** [104.45, 117.55] **(d)** Independence among assemblies; normal distribution for differences. **(e)** H_0: $\mu_d = 0$; H_a: $\mu_d < 0$ (where differences are Top – Bottom); $t = -2.10$ on 14 df; $.025 < p\text{-value} < .05$. **(f)** [−13.49, 1.49]

3. **(a)** [−0.0023, .0031] mm **(b)** H_0: $\mu_d = 0$; H_a: $\mu_d \neq 0$; $z = .24$; p-value $= .8104$. There is no evidence of a systematic difference between calipers. **(c)** The confidence interval in part (a) contains zero; in fact, zero is near the middle of the interval. This means that zero is a very plausible value for the mean difference—there is no evidence that the mean is not equal to zero. This is reflected by the large p-value in part (b).

4. **(a)** The data within each sample must be iid normal, and the two distributions must have the same variance σ^2. One way to check these assumptions is to normal plot both data sets on the same axes. For such small sample sizes, it is difficult to definitively verify the assumptions. But the plots are roughly linear with no outliers, indicating that the normal part of the assumption may be reasonable. The slopes are similar, indicating that the common variance assumption may be reasonable. **(b)** Label the Treaded data Sample 1 and the Smooth data Sample 2. $H_0: \mu_1 - \mu_2 = 0$; $H_a: \mu_1 - \mu_2 \neq 0$; $t = 2.49$; p-value is between .02 and .05. This is strong evidence of a difference in mean skid lengths. **(c)** [2.65, 47.35] **(d)** [2.3, 47.7]

Section 4

1. **(a)** [9.60, 37.73] **(b)** 57.58 **(c)** $H_0: \frac{\sigma_T^2}{\sigma_S^2} = 1$ $H_a: \frac{\sigma_T^2}{\sigma_S^2} \neq 1$; $f = .64$ on 5,5 df; p-value $> .50$ **(d)** [.36, 1.80]

2. **(a)** $[7.437, \infty)$ **(b)** $[44.662, \infty)$ **(c)** Top and bottom bolt torques for a given piece are probably not sensibly modeled as independent.

Section 5

1. **(a)** Conservative method: [.562, .758]; .578. Other method: [.567, .753]; .582. **(b)** $H_0: p = .55$; H_a:

$p > .55$; $z = 2.21$; p-value $= .0136$. **(c)** Conservative method: $[-.009, .269]$. Other method: $[-.005, .265]$. **(d)** $\mathrm{H}_0: p_\mathrm{S} - p_\mathrm{L} = 0$; $\mathrm{H}_\mathrm{a}: p_\mathrm{S} - p_\mathrm{L} \neq 0$; $z = 1.87$; p-value $= .0614$.

2. 9604

3. Conservative method: [.22, .35]. Other method: [.23, .34].

4. $\mathrm{H}_0: p_1 - p_2 = 0$; $\mathrm{H}_\mathrm{a}: p_1 - p_2 \neq 0$; $z = -.97$; p-value $= .3320$.

Section 6

1. A consumer about to purchase a single auto would be most interested in a prediction bound, because the single auto that the consumer will purchase is likely to have mileage above the bound. This is not true for a confidence bound for the mean. That may be more useful for the EPA official, since this person wants to be sure that the manufacturer is producing cars that exceed some minimum average mileage. The design engineer would be most interested in a lower tolerance bound for most mileages, to be sure that a high percentage of the cars produced are able to cruise for at least 350 miles. A confidence for the mean or prediction bound does not answer this question.

2. **(a)** [132.543, 297.656] **(b)** [92.455, 337.745] **(c)** The tolerance interval is much wider than the prediction interval. The interval in (b) is meant to bracket 90% of all observations, while the the one from (a) is meant only to bracket a single additional observation. **(d)** The confidence interval for mean lifetime is smaller than both the prediction interval and the tolerance interval. It is meant only to bracket the mean/center of the population, not additional observation(s). **(e)** $[152.811, \infty)$ **(f)** $[113.969, \infty)$

3. **(a)** [3.42, 6.38]; [30.6, 589.1] **(b)** [3.87, 5.93]; [48.1, 375.0] **(c)** The intervals in (a) are wider than those in (b). This is usually true when applying tolerance intervals and prediction intervals in the same situation.

4. 92.6%; 74.9%

Chapter 7

Section 1

1. **(a)** The plot reveals two outliers. The assumptions of the one-way normal model appear to be less than perfectly met in this problem. (Both of the outliers come from the 8,000 psi condition. This is an indication that the common σ part of the one-way normal model may be less than perfect.) **(b)** .02057. This measures the magnitude of baseline variation in any of the five treatments, assuming it is the same for all five treatments; [.01521, .03277].

2. **(a)** The plot reveals one outlier/unusual residual (the 1.010 value from Van #1 produces the residual $-.0094$). One should proceed under the one-way model assumptions only with caution. **(b)** The standardized residuals tell the same story told in part (a). **(c)** $s_\mathrm{p} = .0036$ measures the (supposedly common) variation in tilt angle for repeated measurement of a particular van; [.0026, .0058].

Section 2

1. **(a)** .02646; 75% **(b)** .03742 **(c)** $[-.0724, .0572]$ provides no convincing evidence of nonlinearity over the range from 2,000 to 6,000, as it includes 0.

2. **(a)** The intervals in numerical order of the four vans are: [1.0875, 1.0984], [.9608, 9716], [1.0145, 1.0242], [.9968, 1.0076]; at least 96% simultaneous confidence. **(b)** $\Delta = .0077$; $\Delta = .0073$ **(c)** [.013516, .02408]

3. Before the data are collected,the probability is .05 that an individual 95% confidence interval will be in error—that it will not contain the quantity that it is supposed to contain. If several of these individual intervals are made, then the probability that *at least* one of the intervals is in error is greater than .05. (If each interval has a .05 chance of failing, then the overall chance of at least one failure is greater than .05.) When making several intervals, most people would like the overall or simultaneous error probability to be small. In order to make sure, for example, that the overall error probability is .05, the error probability associated with the individual intervals must be made smaller than .05. This is equivalent to increasing the individual confidences (above 95%), which makes the intervals wider.

Section 3

1. **(a)** .03682; it is larger. **(b)** .05522; it is larger.

2. **(a)** $k_2^* = 2.88$ so the intervals in numerical order of the four vans are: [1.0878, 1.0982], [.9610, 9714], [1.0147, 1.0240], [.9970, 1.0074]. **(b)** $\Delta =$ .0097; $\Delta = .0092$. These are larger than the earlier Δ's. The confidence level here is a simultaneous one while the earlier level was an individual one. The intervals here are doing a more ambitious job and must therefore be wider.

Section 4

1. **(a)** Small, since some means differ by more than the Δ there. **(b)** $SSTr = .285135$, $MSTr = .071284$, df $= 4$; $SSE = .00423$, $MSE = .000423$, df $= 10$; $SSTot = .289365$, df $= 14$; $f = 168.52$ on 4,10 df; p-value $< .001$. $R^2 = .985$.

2. **(a)** Small, since some sample means differ by more than the Δ's there. **(b)** $SSTr = .034134$, $MSTr = .011378$, df $= 3$; $SSE = .000175$, $MSE = .000013$, df $= 13$; $SSTot = .034308$, df $= 16$; $f = 847$ on 3,13 df; p-value $< .001$.

3. **(a)** To check that the μ_i's are normal, make a normal plot of the $\bar{y}_i$'s. To check that the ϵ_i's are normal, make a normal plot of the residuals. (Normal plotting each sample individually will not be very helpful because the sample sizes are so small.) Both plots are roughly linear, giving no evidence that the one-way random effects model assumptions are unreasonable. **(b)** $SSTr = 9310.5$, $MSTr = 1862.1$, df $= 5$; $SSE = 194.0$, $MSE = 16.2$, df $= 12$; $SSTot = 9504.5$, df $= 17$; $f = 115.18$ on 5,12 df; p-value $< .001$. $\hat{\sigma} = 4.025$ measures variation in y from repeated measurements of the same rail; $\hat{\sigma}_\tau = 24.805$ measures the variation in y from differences among rails. **(c)** [3.46, 13.38]

4. **(a)** Unstructured multisample data could also be thought of as data from one factor with r levels. In many situations, the specific levels of the factor included in the study are the levels of interest. For example, in comparing three drugs, the factor might be called"Treatment." It might have four levels: Drug 1, Drug 2, Drug 3, and Control. The experimenter is interested in comparing the specific drugs used in the study to each other and to the control. Sometimes the specific levels of the factor are not of interest in and of themselves, but only because they may represent (perhaps they are a random sample of) many different possible levels that could have been used in the study. A random effects analysis is appropriate in this situation. For an example, see part (b). **(b)** If there are many technicians, and five of these were randomly chosen to be in the study, then interest is in the variation among all technicians, not just the five chosen for the study. **(c)** $\hat{\sigma} = .00155$ in.; $\hat{\sigma}_\tau = .00071$ in.

Section 5

1. **(a)** Center line$_{\bar{x}} = 21.0$, $UCL_{\bar{x}} = 22.73$, $LCL_{\bar{x}} = 19.27$. Center line$_R = 1.693$, $UCL_R = 4.358$, no LCL_R. **(b)** Center line$_s = .8862$, $UCL_s = 2.276$, no LCL_s. **(c)** 1.3585; 1.3654; $s_p = 1.32$. **(d)** Center line$_{\bar{x}} = 21.26$, $UCL_{\bar{x}} = 23.61$, $LCL_{\bar{x}} = 18.91$. Center line$_R = 2.3$, $UCL_R = 5.9202$, no LCL_R. **(e)** Center line$_{\bar{x}} = 21.26$, $UCL_{\bar{x}} = 23.62$, $LCL_{\bar{x}} = 18.90$. Center line$_s = 1.21$, $UCL_s = 3.10728$, no LCL_s.

2. **(a)** $\frac{\bar{R}}{d_2} = \frac{4.052632}{2.326} = 1.742318 \times .001$ in.; $\frac{\bar{s}}{c_4} = \frac{1.732632}{.9400} = 1.843226 \times .001$ in. **(b)** For the R chart Center Line$_R$ $= 2.326(1.843226) = 4.287344 \times .001$ in., $UCL_R = 4.918(1.843226) = 9.064985 \times .001$ in. and there is no lower control limit. For the s chart Center Line$_s = 1.732632 \times .001$ in. $UCL_s = 2.089(1.732632) = 3.619468 \times .001$ in. and there is again no lower control limit. Neither chart indicates that the short-term variability of the process (as measured by σ) was unstable. **(c)** Use Center Line$_{\bar{x}} = 11.17895 \times .001$ in. above nominal, $LCL_{\bar{x}} = 11.17895 - 3\frac{1.843226}{\sqrt{5}} = 8.706 \times .001$

in. above nominal and $UCL_{\bar{x}} = 11.17895 + 3\frac{1.843226}{\sqrt{5}} = 13.65189 \times .001$ in. above nominal. $\bar{x}$ from sample 16 comes close to the upper control limit, but overall the process mean seems to have been stable over the time period. **(d)** The $\bar{x}$'s from samples 9 and 16 seem to have "jumped" from the previous $\bar{x}$. The coil change may be causing this jump, but it could also be explained by common cause variation. It may be something worth investigating. **(e)** Assuming that the mean could be adjusted (down), you need to look at one of the estimates of σ to answer this question about individual thread lengths. (You should not use control limits to answer this question!) If μ could be made equal to zero, then (assuming normally distributed thread lengths), almost all of the thread lengths would fall in the interval $\pm 3\sigma$. Using the estimate of σ based on $\bar{s}$ from part (a), this can be approximated by $3(1.843226) = 5.53 \times .001$ in. It does seem that the equipment is capable of producing thread lengths within .01 in. of nominal. If the equipment were not capable of meeting the given requirements, the company could invest in better equipment. This would "permanently" solve the problem, but it might not be feasible from a financial standpoint. A second option is to inspect the bolts and remove the ones that are not within .01 in. of nominal. This might be cheaper than investing in new equipment, but it will do nothing to improve the quality of the process in the long run. A third option is to study the process (through experimentation) to see if there might be some way of reducing the variability without making a large capital investment.

3. Control charting is used to monitor a process and detect changes (lack of stability) in a process. The focus is on detecting changes in a meaningful parameter such as μ, σ, p, or λ. Points that plot out of control are a signal that the process is not stable at the standard parameter value (for a standards given chart) or was not stable at any parameter value (for a retrospective chart). The overall goal is to reduce process variability by identifying assignable causes and taking action to eliminate them. Reducing variability increases the quality of the process output.

4. Shewhart control charts do not physically control a process in the sense of guiding or adjusting it. They only monitor the process, trying to detect process instability. There is an entirely different field dedicated to "engineering control"; this field uses feedback techniques that manipulate process variables to guide some response. Shewhart control charts simply monitor a response, and are not intended to be used to make "real time" adjustments.

5. Out-of-control points should be investigated. If the causes of such points can be determined and eliminated, this will reduce long-term variation from the process. There must be an active effort among those involved with the process to improve the quality; otherwise, control charts will do nothing to improve the process.

6. Control limits for an $\bar{x}$ chart are set so that, under the assumption that the process is stable, it would be very unusual for an $\bar{x}$ to plot outside the control limits. The chart recognizes that there will be *some* variation in the $\bar{x}$'s even if the process is stable, and prevents overadjustment by allowing the $\bar{x}$'s to vary "randomly" within the control limits. If the process mean or standard deviation changes, $\bar{x}$'s will be more likely to plot outside of the control limits, and sooner or later the alarm will sound. This provides an opportunity to investigate the cause of the change, and hopefully take steps to prevent it from happening again. In the long run, such troubleshooting may improve the process by making it less variable.

Section 6

1. **(a)** Center line$_{\hat{p}} = .02$, $UCL_{\hat{p}} = .0438$, no $LCL_{\hat{p}}$. **(b)** Center line$_{\hat{p}} = .0234$, $UCL_{\hat{p}} = .0491$, no $LCL_{\hat{p}}$.

2. Center line$_{\hat{u}_i} = .138$ for all i, $UCL_{\hat{u}_i} = .138 + 3\sqrt{\frac{.138}{k_i}}$, no $LCL_{\hat{u}_i}$ for all i.

3. **(a)** Center line$_{\hat{u}_i}$ $= .714$ for all i, $UCL_{\hat{u}_i} = .714 + 3\sqrt{\frac{.714}{k_i}}$, no $LCL_{\hat{u}_i}$ for all i. The process seems to be stable. **(b)** (i) if $k_i = 1$, .0078; if $k_i = 2$, .0033. (ii) if $k_i = 1$, .0959; if $k_i = 2$, .1133.

4. $\hat{p} = \frac{18}{250} = .072$, so Center Line$_{\hat{p}_i}$ $= .072$. The control limits depend on the sample size n_i. For $n_i = 20$, $.072 - 3\sqrt{\frac{.072(1-.072)}{20}} = -.101399 < 0$, so there is no lower control limit, while $UCL_{\hat{p}_i} = .072 + 3\sqrt{\frac{.072(1-.072)}{20}} = .245399$. For $n_i = 30$, $.072 - 3\sqrt{\frac{.072(1-.072)}{30}} = -.06957966 < 0$, so there is no lower control limit, while $UCL_{\hat{p}_i} = .072 + 3\sqrt{\frac{.072(1-.072)}{30}} = .2135797$. For $n_i = 40$, $.072 - 3\sqrt{\frac{.072(1-.072)}{40}} = -.05061158 < 0$, so there is no lower control limit, while $UCL_{\hat{p}_i} = .072 + 3\sqrt{\frac{.072(1-.072)}{40}} = .1946116$. There is no evidence that the process fraction nonconforming was unstable (changing) over the time period studied.

5. If different data collectors have different ideas of exactly what a “nonconformance” is, then the data collected will not be consistent. A stable process may look unstable (according to the c chart) because of these inconsistencies.

6. It may indicate that the chart was not applied properly. For example, if hourly samples of size $m = 4$ are collected, it may or may not be reasonable to use a retrospective $\bar{x}$ chart with $m = 4$. If the 4 items sampled are from 4 different machines, 3 of which are stable at some mean and the 4th stable at a different mean, then the sample ranges and standard deviations will be inflated. This will make the control limits on the $\bar{x}$ chart too wide. Also, the $\bar{x}$'s will show very little variation about a center line somewhere between the two means. This is all a result of the fact that each sample is really coming from four different processes. Four different control charts should be used.

Chapter 8

Section 1

1. **(a)** Error bars: $\bar{y}_{ij} \pm 23.54$. **(b)** $a_1 = 21.78$, $a_2 = -21.78$, $b_1 = -41.61$, $b_2 = 16.06$, $b_3 = 25.56$, $ab_{11} = -1.94$, $ab_{12} = 1.39$, $ab_{13} = .56$, $ab_{21} = 1.94$, $ab_{22} = -1.39$, $ab_{23} = -.56$. Interactions: $ab_{ij} \pm 9.52$. A main effects: $a_i \pm 6.73$. B main effects: $b_j \pm 9.52$. Interactions are not detectable, but main effects for both A and B are. **(c)** $\bar{y}_{.j} - \bar{y}_{.j'} \pm 20.18$

2. **(a)** $s_p = 33.25$ measures baseline variation in y for each factor-level combination, assuming it is the same for all factor-level combinations. **(b)** Error bars: $\bar{y}_{ij} \pm 27.36$. **(d)** $a_1 = -2.77$, $a_2 = -17.4$, $a_3 = 20.17$, $b_1 = -13.33$, $b_2 = -1.20$, $b_3 = 14.53$, $ab_{11} = .033$, $ab_{12} = -5.40$, $ab_{13} = 5.37$, $ab_{21} = -2.13$, $ab_{22} = -.567$, $ab_{23} = 2.70$, $ab_{31} = 2.104$, $ab_{32} = 5.97$, $ab_{33} = -8.07$. **(e)** 18.24. No. **(f)** Use $(a_i - a_i') \pm 22.35$. **(g)** Use $(a_i - a_i') \pm 26.88$.

Section 2

1. **(a)** $\hat{E} \pm .014$. B and C main effects, BC interaction. **(b)** $s_{FE} = .0314$ with 20 df; close to $s_p = .0329$. **(c)** Using few effects model: [3.037, 3.091]. Using general method: [3.005, 3.085].

2. **(a)** Only the main effect for A plots “off the line.” **(b)** Since the D main effect is almost as big (in absolute value) as the main effect for A, you might choose to include it. For this model, the fitted values are (in standard order): 16.375, 39.375, 16.375, 39.375, 16.375, 39.375, 16.375, 39.375, −4.125, 18.875, −4.125, 18.875, −4.125, 18.875, −4.125, 18.875. **(c)** Set A low (unglazed) and D high (no clean). [0, 9.09].

3. **(a)** $\bar{y}_{....} = 3.594$, $a_2 = -.806$, $b_2 = .156$, $ab_{22} = -.219$, $c_2 = -.056$, $ac_{22} = -.031$, $bc_{22} = .081$, $abc_{222} = .031$, $d_2 = -.056$, $ad_{22} = -.156$, $bd_{22} = .006$, $abd_{222} = -.119$, $cd_{22} = -.031$, $acd_{222} = -.056$, $bcd_{222} = -.044$, $abcd_{2222} = .006$. **(b)** It appears that only the main effect for A is detectably larger than the rest of the effects, since the point for a_2 is far away from the rest of the fitted effects. **(c)** To minimize y, use A(+) (monks cloth) and B(+) (treatment Y).

Section 3

1. Since A $\leftrightarrow$ BCDE, if both are large but opposite in sign, their estimated sum will be small.

2. **(a)** 8.23, .369, .256, $-.056$, .344, $-.069$, $-.081$, $-.093$, $-.406$, .181, .269, $-.344$, $-.094$, $-.156$, $-.069$, .019. **(b)** .312. The sums $\alpha_2 + \beta\gamma\delta\epsilon_{2222}$, $\gamma_2 + \alpha\beta\delta\epsilon_{2222}$, $\delta_2 + \alpha\beta\gamma\epsilon_{2222}$, and $\alpha\beta\delta_{222} + \gamma\epsilon_{22}$ are detectable. Simplest explanation: A, C, D main effects and CE interaction are responsible for these large sums. **(c)** A (+), C (+), D ($-$), and E ($-$). The abc combination, which did have the largest observed bond strength.

3. **(b)** (1), ad, bd, ab, cd, ac, bc, abcd. Estimated sums of effects: 3.600, $-.850$, .100, $-.250$, $-.175$, $-.025$, $-$, 075, $-.025$. **(c)** The estimate of $\alpha_2 + \beta\gamma\delta_{222}$ plots off the line. Still, one might conclude that this is due to the main effect for A, but the conclusion here would be a little more tentative.

Section 4

1. The advantage of fractional factorial experiments is that the same number of factors can be studied using less experimental runs. This is important when there are a large number of factors, and/or experimental runs are expensive. The disadvantage is that there will be ambiguity in the results; only sums of effects can be estimated. The advantage of using a complete factorial experiment is that all means can be estimated, so all effects can be estimated.

2. It will be impossible to separate main effects from two-factor interactions. You would hope that any interactions are small compared to main effects; the results of the experiment can then be (tentatively) summarized in terms of main effects. (If all interactions are really zero, then it is possible to estimate all of the main effects.) Looking at Table 8.35, the best possible resolution is 3 (at most).

3. Those effects (or sums of effects) that are nearly zero will have corresponding estimates that are "randomly" scattered about zero. If all of the effects are nearly zero, then one might expect the estimates from the Yates algorithm (excluding the one that includes the grand mean) to be bell-shaped around zero. A normal plot of these estimates would then be roughly linear. However, if there are effects (or sums of effects) that are relatively far from zero, the corresponding estimates will plot away from the rest (off the line), and may be considered more than just random noise. The principle of "sparsity of effects" says that in most situations, only a few of the many effects in a factorial experiment are dominant, and their estimates will then plot off the line on a normal plot.

4. **(a)** I $\leftrightarrow$ ABCDF $\leftrightarrow$ ABCEG $\leftrightarrow$ DEFG **(b)** ABDF, ABEG, CDEFG **(c)** +, +; $-$, $-$ **(d)** That only A, F, and their interaction are important in describing y.

5. 3.264

6. **(a)** I $\leftrightarrow$ ABCE $\leftrightarrow$ BCDE $\leftrightarrow$ ADEF **(b)** $-$, $-$; +, $-$ **(c)** .489

Chapter 9

Section 1

1. **(a)** $s_{LF} = 67.01$ measures the baseline variation in Average Molecular Weight for any particular Pot Temperature, assuming this variation is the same for all Pot Temperatures. **(b)** Standardized residuals: 2.0131, $-.3719$, $-.9998$, -1.562, .2715, .2394, .7450, .0004 **(c)** [22.08, 24.91] **(d)** [1761, 1853], [2630, 2770] **(e)** [1745, 1869], [2605, 2795] **(f)** 1705; 2590 **(g)** 1627; 2503 **(h)** $SSR = 4{,}676{,}798$, $MSR = 4{,}676{,}798$, df = 1; $SSE = 26{,}941$, $MSE = 4490$, df = 6; $SSTot = 4{,}703{,}739$, df = 7; $f = 1041.58$ on 1,6 df; p-value $< .001$

2. **(a)** $b_0 = 4345.9$, $b_1 = -3160.0$, $s_{LF} = 26.76$ (close to $s_p = 26.89$) **(b)** Standardized residuals: 1.32, $-.48$, $-.04$, $-.91$, .52, -1.07, 1.94, $-.04$, -1.09. **(c)** [-357.4, -274.64] **(d)** $t = -14.47$ on 7 df, p-value $< .001$; or $f = 209.24$ on 1,7 df, p-value $< .001$. **(e)** [2744.8, 2787.0] **(f)** [2699.2, 2832.6] **(g)** 2698.5

Section 2

1. **(a)** $s_{SF} = .04677$ measures variation in Elapsed Time for any particular Jetting Size, assuming this variation is the same for all Jetting Sizes. **(b)** Standardized residuals: $-.181$, $.649$, $-.794$, $-.747$, 1.55, -1.26. **(c)** $[81.32, 126.66]$; $[-3.17, -1.89]$; $[.01344, .02245]$ **(d)** $[14.462, 14.596]$; $[14.945, 15.145]$ **(e)** $[14.415, 14.644]$; $[14.875, 15.215]$ **(f)** 14.440; 14.942 **(g)** 14.323; 14.816 **(h)** $SSR = .20639$, $MSR = .01319$, df $= 2$; $SSE = .00656$, $MSE = .00219$, df $= 3$; $SSTot = .21295$, df $= 5$; $f = 42.17$ on 2,3 df; p-value $= .005$. H_0 means that Elapsed Time is not related to Jetting Size. **(i)** $t = 9.38$; p-value $= .003$. $H_0: y \approx \beta_0 + \beta_1 x + 0$; i.e., Elapsed Time is related to Jetting Size only linearly (no curvature).

2. **(a)** $s_{SF} = .4851$ measures baseline variation in y for any (x_1, x_2) combination, assuming this variation is the same for all (x_1, x_2) combinations. **(b)** Standardized residuals: $-.041$, $.348$, 1.36, -1.44, -1.00, $-.457$, 1.92, $.348$, $-.604$. **(c)** $[5.036, 7.060]$; $[.0775, .2058]$; $[-.0298, -.0041]$ **(d)** $[5.992, 6.622]$; $[5.933, 6.625]$ **(e)** $[5.798, 6.816]$; $[5.720, 6.838]$ **(f)** 5.571; 5.535 **(g)** 5.017; 4.970 **(h)** $SSR = 5.8854$, $MSR = 2.9427$, df $= 2$; $SSE = 1.4118$, $MSE = .2353$, df $= 6$; $SSTot = 7.2972$, df $= 8$; $f = 12.51$ on 2,6 df; p-value $= .007$.

Section 3

1. **(a)** $\hat{y} = 31.40 + 7.430 \ln x_1 - .08101 x_2 - .2760 (\ln x_1)^2 + .00004792 x_2^2 - .006596 x_2 \ln x_1$. $R^2 = .724$. $s_{SF} = 1.947$. $s_p = 2.136$, which is greater than s_{SF}, so there is no indication that the model is inappropriate. **(b)** Factor-level combinations have fitted values that differ by as much as .77. **(d)** (i) $[.128, 2.781]$. (ii) $[-2.693, 5.601]$. (iii) -2.332.

2. **(a)** Estimate of $\mu_{...} = .67407$; estimate of $\alpha_2 = .12407$; estimate of $\beta_2 = -.30926$. **(b)** There is some hint of a pattern in the plot of Standardized Residuals versus levels of C, indicating that the amount of additive may be having a small effect that the model is not accounting for. Otherwise, the residuals do not provide any evidence that the model is inadequate. **(c)** $s_{FE} = .09623$. $s_p = .12247$. No; $s_{FE} < s_p$.

Appendix A (selected answers only)

Section 1

1. **(a)** .1865 **(b)** .6083

2. **(a)** .54 **(b)** .78

3. **(a)** .505 **(b)** .998

4. **(a)** .76 **(b)** .78 **(c)** .974

5. **(a)** .75 **(b)** .80 **(c)** .75 **(d)** Yes, since the answers to parts (a) and (c) are the same. **(e)** One such pair is “ring meets spec.s on first grind” and “ring is ground twice.”

Section 2

1. $r = .99979$

2. $k = 2$

Section 3

1. **(a)** 1.7310×10^{13} **(b)** 2.2777×10^{12} **(c)** .1316

2. **(a)** .0000081 **(b)** .03402

3. **(a)** 1,757,600 **(b)** .00167 **(c)** .0167

4. **(a)** .5 **(b)** .167

Section 4

1. **(a)** 20; 15.81 **(b)** $1 - \frac{3}{2}\exp\left(-\frac{t}{15}\right) + \frac{1}{2}\exp\left(-\frac{t}{5}\right)$ **(c)** $f_T(t) = \frac{1}{10}\left(\exp\left(-\frac{t}{15}\right) - \exp\left(-\frac{t}{5}\right)\right)$ **(d)** $S_T(t) = \frac{3}{2}\exp\left(\frac{t}{15}\right) - \frac{1}{2}\exp\left(-\frac{t}{5}\right)$, $h_T(t) = \frac{1}{5}\left(\frac{\exp\left(-\frac{t}{15}\right) - \exp\left(-\frac{t}{5}\right)}{3\exp\left(-\frac{t}{15}\right) - \exp\left(-\frac{t}{5}\right)}\right)$. $h_T(t)$ is not constant. It starts at 0, and increases to an asymptote of $1/15$.