Basic Probability and Statistics Problems

1) A computer system uses passwords that consist of five letters followed by a single digit.
   a) How many passwords are possible?
   b) How many possible passwords end in an even digit?
   c) How many possible passwords do not contain the letter O?
2) A production lot has 100 units of which 20 are known to be defective. A random sample of 4 units is selected. What is the probability that the sample will contain no more than 2 defective units?
3) The investigation of consumer product complaints by the Federal Trade Commission (FTC) has generated much interest by manufacturers in the quality of their products. A manufacturer of food processors conducted an analysis of a large number of consumer complaints and found that they fell into the categories shown below. If a consumer complaint is received, what is the probability that the cause of the complaint was product appearance given that the complaint originated during the guarantee period?

	Reason for Complaint		
	Electrical	Mechanical	Appearance
During Guarantee Period	18%	13%	32%
After Guarantee Period	12%	22%	3%

4) Assume that the engine component of a spacecraft consists of two engines in parallel. If the main engine is 95% reliable, the backup is 80% reliable, and the engine component as a whole is 99% reliable,
   a) what is the probability that both engines will be operable?
   b) what is the probability that the main engine will fail but the backup will be operable?
   c) what is the probability that the backup engine will fail but the main engine will be operable?
   d) what is the probability that the engine component will fail?
   e) what is the probability that, in an engine system such as that described, the backup engine will function given that the main engine fails?
5) An engineer is in charge of a particular process at an oil refinery. Past experience indicates that 10% of all shutdowns are due to equipment failure alone, 5% are due to a combination of equipment failure and operator error, and 40% involve operator error. A shutdown occurs. Find the probability that
   a) equipment failure or operator error is involved
   b) operator error alone is involved
   c) neither operator error nor equipment failure is involved
   d) operator error is involved given that equipment failure occurs
   e) operator error is involved given that equipment failure does not occur
6) The completion time of a construction project depends on whether the carpenters or the plumbers working on the project go on strike. The probabilities of delay are 100%, 80%, 40%, and 5% if both go on strike, carpenters alone go on strike, plumbers alone go on strike, and neither of them strikes respectively. Also, there is 60% chance that plumbers will strike if carpenters strike, and if plumbers go on strike there is a 30% chance that carpenters would follow. It is known that the chance for the plumbers' strike is 10%.
   a) Determine the probability that both plumbers and carpenters strike.
   b) Determine the probability that carpenters strike.
   c) Determine the probability that carpenters alone go on strike.
   d) Determine the probability that plumbers alone go on strike.
   e) Determine the probability that neither of them strikes.
   f) Determine the probability of delay in completion.
   g) If there is a delay in completion, determine the following:
      i) Probability that both carpenters and plumbers are on strike.
      ii) Probability that carpenters strike and plumbers are not on strike.
      iii) Probability that carpenters are on strike.
7) Let X and Y be independent random variables with $E[X] = 2$, $E[X^2] = 29$, $E[Y] = 4$, and $E[Y^2] = 52$. Define $W = X + Y$ and $Z = 2X$. The random variables W and Z are not independent. What is their correlation coefficient?

8) The occurrence of flood may be modelled by a Poisson Process. If the mean occurrence rate of floods for a certain region is once every 8 years, determine the probability
   a) of no floods in a 10-year period
   b) of more than three floods in a 15-year period.
9) The daily water consumption of a city may be assumed to be normally distributed with a mean of 500,000 gpd and a standard deviation of 150,000 gpd.
   a) For a given day, what is the probability that the water supply must exceed 750,000 gallons to avoid a shortage?
   b) If the city engineer wants the probability of shortage to be no more than 1% in any given day, how much water supply is required?
10) The students Losen, Cahoy, and Lewis measured the lengths of some spanner bushings of a particular type purchased from a local machine supply shop. The lengths (in inches) obtained by one of the students were as follows:

    1.1375, 1.1390, 1.1420, 1.1430, 1.1410, 1.1360, 1.1395, 1.1380
    1.1350, 1.1370, 1.1345, 1.1340, 1.1405, 1.1340, 1.1380, 1.1355

   a) If one is to build a confidence interval for the population mean measured length for these bushings, what assumptions must be made? Use a histogram to assess the reasonableness of this assumption.
   b) No matter what you concluded in (a), build a 90% confidence interval for the mean measured length for bushings of this type.
   c) The supply shop claims that the length of a bushing is 1.14in. Based on the confidence interval found in (b), what do you conclude?
   d) The variability in the length of the bushings is obviously a concern. Build a 98% confidence interval for the standard deviation of measured length for these bushings.

   The above study also included measurement of the outside diameters of the 16 bushings. Two of the students measured each of the bushings with the following results;

Bushing	1	2	3	4	5	6	7	8
Student A	.3690	.3690	.3690	.3700	.3695	.3700	.3695	.3690
Student B	.3690	.3695	.3695	.3695	.3695	.3700	.3700	.3690

Bushing	9	10	11	12	13	14	15	16
Student A	.3690	.3695	.3690	.3690	.3695	.3700	.3690	.3690
Student B	.3700	.3690	.3695	.3695	.3690	.3695	.3690	.3690

   e) If one wants to compare the two students' average measurements, how must one proceed? Why?
   f) What assumptions must be made in order to compute a confidence interval as described in (e)? Do these look valid? How can you check?
   g) Build a 95% confidence interval for the mean difference in outside measurements for the two students.
   h) Are the students coming up with the same mean measurement? Conclude based on the confidence interval found in (g)?
   i) Explain why one cannot, in general, compare the mean in the measurements of Student A to the mean in the measurements of Student B and still arrive at the same conclusion as in part h (in this case, you would arrive at the same conclusion, but what happens to the CI in general?).
11) Eastman, Frye, and Schnept worked with a company that mass-produces plastic bags. Their work at the company eventually focused on startup problems of a particular machine that could be operated at either a high speed or a low speed. One part of the data they collected consisted of counts of faulty bags produced in the first 250 manufactured after changing a roll of plastic feedstock. The counts they obtained for both low- and high-speed operation of the machine were 147 faulty under high-speed operation and 12 faulty under low-speed operation. In what follows, suppose that it is sensible to think of the machine as operating in a physically stable fashion during the production of the first 250 bags after changing a roll of plastic, with a constant probability (possibly different under different speeds) of any particular bag produced being faulty.
   a) Find a point estimate and a 95% confidence interval estimate for the probability of obtaining a faulty bag under high-speed operation.
   b) Find a point estimate and a 95% confidence interval estimate for the probability of obtaining a faulty bag under low-speed operation.
   c) Find a point estimate and a 95% confidence interval estimate for the difference in the two probabilities.
   d) Do the two speeds of operation yield the same proportion of faulty bags? Comment based on results found in (c).

12) Every hour a sample of ten fittings to be used in an aircraft hydraulic system is drawn from the production line and the pitch diameter of the threads is measured. The measurements (in inches) for two different hours were

    hour1: 0.4036 0.4035 0.4034 0.4033 0.4032 0.4033 0.4035 0.4035 0.4039 0.4036,
    hour2: 0.4031 0.4031 0.4034 0.4032 0.4030 0.4035 0.4036 0.4029 0.4027 0.4032

    a) Build a 90% confidence interval for a quantity that will allow you to compare the variability in the pitch diameters for the two different hours.
    b) Suppose you had to build a confidence interval for the difference in the two mean pitch diameters (the one from the first hour and the one from the second), would you feel comfortable with making the necessary assumptions? Comment based on results found in (a) and any plots you feel are relevant.
    c) Build a 96% confidence interval for the difference in the two mean pitch diameters.
    d) Is there an hourly difference in the mean pitch diameter? Comment based on (c).

13) It is thought that the application of a plasma coating that contains submicron particles of tungsten carbide will reduce wear to rotary valves used in the pulp and paper industry. Tests are conducted to evaluate the wear in coated valves. Data are gathered on the wear of the part in millimeters over the test period and are

    0.077, 0.085, 0.078, 0.096, 0.092, 0.106, 0.078, 0.116,
    0.099, 0.074, 0.082, 0.149, 0.088, 0.081, 0.072, 0.099.

    a) Find a 98% confidence interval for the standard deviation of the wear (You can use Minitab to make the necessary mean and variance computations).
    b) What assumptions are required for the result in (a) to hold? Do they appear valid? Perform a check.

14) The marketing director of a large battery manufacturer is thinking about spending a lot of money on a promotional campaign. To decide the issue, she wishes to find the manufacturer's current share of the battery market. If their market share is less than 25%, she will proceed with the expensive campaign. She calls upon her Research Division to come up with more information on the subject. The latter decide to choose 20 distributors at random and see if they distribute the manufacturer's batteries.

    a) What hypotheses should be tested, and why?
    b) In this context, what is the type I error?
    c) In this context, what is the type II error?
    d) For the marketing director, which is the more serious error? Why?

15) Recall #10.

    a) The supply shop claims that the length of a bushing is 1.14in. Test this claim using a formal procedure at $\alpha = .05$. What do you conclude? Note all required assumptions and comment on their validity. Also report a p-value.
    b) Further recall that the study also included measurement of the outside diameters of the 16 bushings. Two of the students measured each of the bushings. Are the two students' average measurements the same? Test using a formal procedure at $\alpha = .05$. Note all required assumptions and comment on their validity. Also report a p-value.
    c) Could you have reached the conclusion drawn in (b) above based on the 95% confidence interval that was built in #10(g)? Comment.

16) Recall #11.

    a) The company claims that the probability of obtaining a faulty bag under high-speed operation is no more that 50%. Test this claim using a formal procedure at $\alpha = .05$. What do you conclude? Also report a p-value.
    b) The company claims that the probability of obtaining a faulty bag under low-speed operation is no more than 4%. Test this claim using a formal procedure at $\alpha = .05$. What do you conclude? Also report a p-value.
    c) Is there sufficient evidence to show that high-speed operation yields a higher proportion of faulty bags than low-speed operation? Test using a formal procedure at $\alpha = .05$. Also report a p-value.

17) Recall #12. Is there an hourly difference in the mean pitch diameter? Test using a formal procedure at $\alpha = .05$. Also report a p-value.