Problem Set #1

Issued: Friday, Feb. 7, 2003
Due: Wednesday, Feb. 12, 2003

Please write and hand in solutions to the following problems or scenarios. Write your solutions from the perspective of an engineer sending an explanation to a coworker (or a customer, or a supplier) of your solution. That is to say, you must write supporting text explaining equations and formulas that you use; a few mathematical expressions and a number or plot at the end will not earn full credit.

Problem 1

A random sample of 20 die is non-destructively examined from each wafer lot at the end of the process. Given that the die yield of the process is 97%, calculate the probability that we will find less than three defective die in this sample.

Problem 2

Part a. Use a statistics or mathematics package (e.g. JMP, Splus, Excel, Matlab) to generate 20 samples x_i, $i = 1$ to 20, where each x_i is independent and identically distributed as a normal distribution with mean 11 and standard deviation 2, i.e. $x_i \sim N(11, 4)$. We will take these as measurements of the dielectric constant of the polysilicon layer on each wafer in a 20 wafer lot. Prepare a table with your sample measurements, and calculate $\bar{x}$, the sample mean, and s_x, the sample standard deviation. Plot (using your stats/ math package) a histogram of your 20 values. Does the distribution appear normal?

Part b. Generate a sample y_i, $i = 1$ to 100 from the same assumed distribution, $y_i \sim N(11, 4)$. You do not need to show the specific 100 values, but plot a histogram of the observed values. Are the sample mean $\bar{y}$ and sample standard deviation s_y closer to the "true" population values of 11 and 2 than with your earlier sample of size 20? Does this surprise you?

Part c. Treat the measurements from part a as samples from lot 1. Create nine more 20 sample data sets; all together, these ten sets of 20 measurements represent the measurements from each of lots $j = 1$ to 10. Prepare a table showing your ten sample means $\bar{x}_j$ and sample standard deviations s_{x_j}, as well as the sample variances $s_{x_j}^2$.

Part c-1: Plot a histogram of your ten sample means. What is the mean and standard deviation across the observed values $\bar{x}_j$? What would you ideally expect for the form (as well as the mean and standard devi-

ation) of the distribution of your histogram, given the "true" population distribution from which the original measurements came?

Part c-2: Plot a histogram of your ten sample variances. What would you ideally expect for the form of the distribution of your histogram of variances, given the "true" population distribution from which the original measurements came?

Problem 3

This problem is almost the same as problem 2, except in this case the underlying "true" distribution is a Poisson distribution, representative of the number of particles detected on each wafer.

Part a. Use a statistics or mathematics package (e.g. JMP, Splus, Excel, Matlab) to generate 20 samples x_i, $i = 1$ to 20, where each x_i is independent and identically distributed as a Poisson distribution with mean 4, i.e. $x_i \sim P(4)$. We will take these as measurements of the number of particles greater than $1\mu m$ in size detected on each wafer in a 20 wafer lot. Prepare a table with your sample measurements, and calculate $\bar{x}$, the sample mean, and s_x, the sample standard deviation. Plot (using your stats/math package) a histogram of your 20 values. Does the distribution appear normal? Poisson? Does the sample mean and standard deviation you calculated for your set of 20 values support the idea that the underlying distribution is Poisson distributed?

Part b. Generate a sample y_i, $i = 1$ to 100 from the same assumed distribution, $y_i \sim P(4)$. You do not need to show the specific 100 values, but plot a histogram of the observed values. Are the sample mean $\bar{y}$ and sample standard deviation s_y closer to the "true" population values (4 and 2 for the mean and standard deviation, respectively) than with your earlier sample of size 20? Does this surprise you?

Part c. Treat the measurements from part a as samples from lot 1. Create nine more 20 sample data sets; all together, these ten sets of 20 measurements represent the measurements from each of lots $j = 1$ to 10. Prepare a table showing your ten sample means $\bar{x}_j$ and sample standard deviations s_{x_j}.

Part c-1: Plot a histogram of your ten sample means. What is the mean and standard deviation across the observed values $\bar{x}_j$? What would you expect for the approximate form (as well as the mean and standard deviation) of the distribution of your histogram, given the "true" population distribution from which the original measurements came?