Name: ______________________________

NANO 114 – Probability and Statistical Methods for Engineers

Professor Shyue Ping Ong

Final

Winter 2015, March 16 2015

Instructions:

1. Write your name on the top right of this page.
2. Please answer **ALL** questions as far as possible.
3. You have a total of **3 hours** exactly.
4. Write your answers in the space provided following the questions. If you need additional space, you may request for additional blank sheets of paper during the exam. If you use additional blank sheets, you should mark the top left of each sheet with your name and the question number and today's date, e.g., "John Smith, Q1, Mar 16 2015".
5. Indicative maximum points for each sub-question are provided only as a reference. Actual grading points may vary.
6. Write neatly and legibly. Illegible answers will be considered as wrong.

Question 1 (16 points): ____________

Question 2 (16 points): ____________

Question 3 (18 points): ____________

Question 4 (16 points): ____________

Question 5 (16 points): ____________

Question 6 (18 points): ____________

This exam has a total of 18 pages, including the cover page.

Qn 1. A sample preparation for a chemical measurement can be completed by either an experienced technician, or an inexperienced technician. The experienced technician is able to prepare the sample correctly 80% of the time, with minor errors 15% of the time and with major errors 5% of the time. The inexperienced technician, on the other hand, can only prepare the sample correctly 20% of the time, with minor errors 30% of the time, and with major errors 50% of the time.

In a research facility, there are one experienced technician and five inexperienced technicians. A technician is randomly selected to prepare a sample.

i. What is the probability of the sample having minor errors? (4 points)
ii. If the sample has minor errors, what is the probability that it was prepared by the experienced technician? (8 points)
iii. It has been determined that the first technician was inexperienced and major errors were made in the preparation of the first sample. Another technician is now randomly selected from the remaining technicians, and asked to do a second sample preparation. What is the probability that the second preparation contains no errors? (4 points)

Qn 2. Flaws occur at random along the length of a thin copper wire according to a Poisson process at an average rate of 2 flaws per mm.

i. What is the probability that there are three or fewer flaws in 4 mm of wire? (2 points)
ii. What is the probability that there are more than two flaws in 3 mm of wire? (2 points)
iii. You are given five meters of the same copper wire. Approximate the probability that there are between 9900 and 10250 flaws in that five meters of wire. (4 points)
iv. The manufacturer also sells a higher quality copper wire made with a more stringent process. Based on their testing, there is a 90% probability that any random stretch of 5mm of this higher quality wire has no flaws. Determine the mean number of flaws per mm of this higher quality wire. (4 points)
v. A random 10 mm length of the higher quality wire in part iv was cut into exactly 10 pieces of wire of exactly 1 mm each. What is the probability that exactly 3 of those 10 pieces have no flaws (the rest of the pieces having at least one flaw)? (4 points)

Qn 3. A battery manufacturer develops lithium-ion batteries for mobile phones. The batteries that come off the assembly line have a normally distributed mean capacity of 3000 mAh with a standard deviation of 200 mAh.

i. What is the probability that a randomly selected battery has a capacity of less than 2600 mAh? (2 points)

ii. The batteries are packed in lots of 10. What is the probability that the average capacity in a lot is less than 2900 mAh? (2 points)

iii. To ensure that its batteries meet stringent quality control, the manufacturer conducts tests on random lots of 10 batteries. If the sample average is significantly less from the population mean capacity, it could indicate that a particular production line has a problem. In one test, the measured capacities are as follows:

Specimen #	1	2	3	4	5	6	7	8	9	10
Capacity (mAh)	3129	2994	2824	2791	3053	3041	2642	2803	2731	2813

Perform a hypothesis test at the 5% level to determine if the mean capacity of the sample is significantly less from the population mean capacity. State your null and alternative hypotheses and outline all steps in the test clearly. (4 points)

iv. Using the same data as in part iii, construct a 99% confidence interval for the population mean capacity. (4 points)

v. Upon inspecting a faulty production line, the manufacturer found that the line in fact had a mean capacity of 2800 mAh and a standard deviation of 100 mAh. Sketch a graph with two normal PDFs indicating the probability that a hypothesis test at the 5% level would erroneously conclude that a lot of 10 batteries from the faulty line is fine, i.e., a Type II error as a shaded region. The graph just needs to be a simple sketch where you indicate the mean and variance of two normal PDFs and shading the region that corresponds to a Type II error. (2 points) Calculate the probability of a type II error.

Q4. An experiment was conducted to determine the effect of C_2F_6 flow rate on the uniformity of the etch on a silicon wafer used in integrated circuit manufacturing. Three flow rates are used in the experiment, and for each flow rate, four observations (in percent) were measured. The data is in the table below. Additional blank columns and rows are provided to make it easier for you to work out the solutions. You can choose whether you want to use them. If you use them, partial credit will be given for the correct useful numbers.

C_2F_6 Flow (SCCM)	Observations				
	1	2	3	4	
125	2.7	4.6	2.6	3	
160	4.9	4.6	5	4.2	
200	4.6	3.4	2.9	3.5	

Let us denote the mean uniformity at each flow rate as $\mu_{flowrate}$, e.g., μ_{125}. We wish to use an ANOVA analysis to determine if there are any substantive differences in uniformity for different flow rates.

i. Construct the null hypothesis and alternative hypothesis for the test. (4 points)
ii. Calculate the mean square variability within and between samples. (4 points)
iii. Perform a hypothesis test at the 0.05 level and determine any significant differences in uniformity with flow rate. In particular, identify the specific pairs of flow rates with different uniformity using Tukey's HSD test. (8 points)

Q5. Six individuals are randomly selected for a medical trial of a new drug to lower the blood pressure. Their weight and systolic blood pressures (BP) were measured at the start of the trial. After the end of the trial, their systolic BPs were measured again. The data is summarized in the table below. Additional blank columns and rows are provided to make it easier for you to work out the solutions. You can choose whether you want to use them. If you use them, partial credit will be given for the correct useful numbers. You may assume that weight and blood pressure are jointly normally distributed.

Subject Number	**Weight (lbs), X**	**Starting Systolic BP (mm Hg), Y**	**Ending Systolic BP (mm Hg), Y'**					
1	165	130	127					
2	167	133	110					
3	180	150	151					
4	155	128	120					
5	212	151	140					
6	175	146	150					

i. Determine the mean and median weight and systolic BP before the experiment. (4 points)

ii. It is hypothesized that there is a linear relationship between the starting systolic BP, denoted by the variable Y, and the weight, denoted by the variable X. Calculate the correlation coefficient between X and Y and perform a hypothesis test to determine if there is a relationship at the 0.05 level. (4 points)

iii. Perform a linear regression to derive the relationship between X and Y of the form: (4 points)

$$Y = bX + a$$

iv. Perform a test of the hypothesis that the new drug was effective in lowering the blood pressure at the 0.01 level. (4 points)

Q6. The average temperature in Fahrenheit for the first ten days of 2015 for Boston is given below.

Day	1	2	3	4	5	6	7	8	9	10
Temp (^oF)	28	36	30	44	34	17	15	9	25	21

i. Calculate the range, median and inter-quartile range (IQR) of the temperature in the ten days. (4 points)

ii. We want to estimate the overall (population) average and standard deviation of the temperature in Boston. Use the data in the ten days to calculate an estimate of the mean and standard deviation of the temperature in Boston in January. (4 points)

iii. We may assume that the temperature is normally distributed with the mean and standard deviation derived in part ii. What is the probability that the temperature is < 30 ^oF on any given day in Jan? (2 points)

iv. A study has determined that when the temperature falls below 30 ^oF, the probability of a flight from Logan Airport being delayed is 0.8. When the temperature is above 30 ^oF, the probability of the flight being delayed is 0.4 (yes, Logan is a pretty bad airport). Your flight in Jan is delayed. What is the probability that the temperature is above 30 ^oF? (4 points)

v. The historical average temperature in Boston in Jan (from the National Oceanic and Atmospheric Administration) is 29^oF. Perform a hypothesis test at the 5% level to determine whether the temperature for the first ten days of Jan 2015 is significantly different from the historical average. (4 points)