University of California, Davis
Department of Biomedical Engineering

Fall 2014	**Probability and Statistics**	**BIM 105**
David M. Rocke	**for Biomedical Engineers**	**November 6, 2014**

Midterm Examination

NAME _____
PUT YOUR NAME ON EVERY PAGE

1. The time T in milliseconds (ms) until a particular kind of neural signal occurs is exponentially distributed with parameter $\lambda = .25$. What are the following?
   (a) The PDF and the CDF of T.
   (b) The expectation $E(T)$, the variance $\text{Var}(T)$, and the standard deviation of T.
   (c) $\Pr(T > 8)$

2. Given the following data on x and y,

   (a) Compute the mean, variance, and standard deviation of x and y.

   (b) Compute the correlation coefficient.

   (c) Find the least-squares line for predicting y from x.

   (d) Find the predicted value of y when $x = 3$.

   (e) Find the residual for the point $(3, 8)$.

x	3	7	2
y	8	12	4

3. Most cases of gastric cancer and most cases of gastric ulcers are caused by infection with Helicobacter pylori (H. pylori or HP). In a particular South American population, 80% are HP+, meaning that they are infected with H. pylori, and 20% are HP−. Of those who are HP+, 10% develop gastric cancer (GC)and 40% develop gastric ulcers (GU), but for reasons that are not completely understood, nobody develops both ulcers and gastric cancer. Of the HP− population, 2% develop gastric cancer and 2% develop ulcers (and none develop both).

   (a) Make a table with the rows corresponding to disease states and the columns to HP status and each cell of the table holding the probability of that event. Include row and column totals.

   (b) Are the events HP+ and GC statistically independent?

   (c) Find Pr(HP+|GU)

   (d) Find any two events that are mutually exclusive.

4. The data below show 16 spectrophotometric measurements of calcium in arbitrary units (to be later calibrated to standards).

1139	1125	1189	1202	1201	1124	1125	1116
1206	1117	1194	1185	1193	1133	1202	1133

   (a) Compute the five-number summary.

   (b) Compute the inner and outer fences.

   (c) In a box and whisker plot, where would the ends of the whiskers be?

5. A fibroblast cell chosen at random from a large cell culture flask has a mean mass X of 2.0 ng with a standard deviation of 1.8 ng. A keratinocyte cell chosen at random from a large cell culture flask has a mean mass Y of 7.1 ng with a standard deviation of 4.2 ng.

   (a) If 8 fibroblasts and 5 keratinocytes are chosen at random from the flasks and all can be assumed to be statistically independent, what is the mean and standard deviation of the combined mass of the 13 cells?

   (b) If one each of fibroblasts and keratinocytes is chosen at random from the flasks, what is the expected value and standard deviation of the difference in the masses between the two cells?

   (c) If a researcher chooses 83 fibroblasts independently at random, what is the expected value of the sample mean weight $\bar{X}$? What is the standard deviation (aka standard error) of $\bar{X}$?

   (d) Given that the standard normal percentage point $z_{.025} = 1.960$, find the mass M in nanograms such that there is a 97.5% probability that the sample average weight from a sample of 83 fibroblasts is less than M.

   (e) Why can you use the normal distribution to answer the previous question?

6. A batch of parts used in the manufacture of a medical device has a defective rate of 0.1%. A sample of 500 items is taken.

   (a) Find the mean, variance, and standard deviation of the number X of defective items in the sample.

   (b) Find the mean, variance, and standard deviation of the sample proportion $\hat{p} = X/500$ of defective items in the sample.

   (c) What is the probability that $X = 0$?, $X = 1$?, $X = 3$?