CE 5690 – Descriptive Modeling of Data
Problem Set No. 1
Due Friday 1/25/08

1. Consider randomly selecting a student from the population of Michigan Tech undergraduates. Let A denote the event that the selected individual has a Visa credit card and B be the analogous event for a MasterCard. Suppose that $P(A) = 0.5$, $P(B) = 0.4$, and $P(A \cap B) = 0.25$.

   (a) Compute the probability that the selected individual has at least one of the two types of credit cards.

   (b) Compute the probability that the selected individual has neither type of card.

   (c) Compute the probability that the selected individual has a Visa card but not a MasterCard.

2. For the credit card scenario of problem 1, calculate the following probabilities:

   (a) $P(B \mid A)$

   (b) $P(B^c \mid A)$

   (c) $P(A \mid B)$

   (d) $P(A^c \mid B)$

   (e) Given that the selected individual has at least one card, what is the probability that he or she has a Visa card?

3. Seventy percent of the light aircraft that disappear while in flight in a certain country are subsequently discovered. Of the aircraft that are discovered, 60% have an emergency locator, whereas 90% of the aircraft NOT discovered do not have such a locator.

   Suppose a light aircraft has disappeared.

   (a) If it has an emergency locator, what is the probability that it will not be discovered?

   (b) If it does not have an emergency locator, what is the probability that it will be discovered?

4. One method used to distinguish between granitic (G) and basaltic (B) rocks is to examine a portion of the infrared spectrum of the sun's energy reflected from the rock surface. Let R_1, R_2, and R_3 denote measured spectrum intensities at three different wavelengths; typically, for granite $R_1 < R_2 < R_3$, whereas for basalt $R_3 < R_1 < R_2$. When measurements are made remotely (using aircraft), various orderings of the R_i's may arise whether the rock is basalt or granite. Flights over regions of known composition have yielded the following information:

	Granite	Basalt
$R_1 < R_2 < R_3$	60%	10%
$R_1 < R_3 < R_2$	25%	20%
$R_3 < R_1 < R_2$	15%	70%

Suppose that for a randomly selected rock in a certain region, $P(G) = 0.25$ and $P(B) = 0.75$.

(a) Show that $P(G \mid R_1 < R_2 < R_3) > P(B \mid R_1 < R_2 < R_3)$. If measurements yielded R1 < R2 < R3, would you classify the rock as granite or basalt?

(b) If measurements yielded $R_1 < R_3 < R_2$, how would you classify the rock? Answer the same question for $R_3 < R_1 < R_2$.

(c) Using the classification rules indicated in parts (a) and (b), when selecting a rock from this region, what is the probability of an erroneous classification? [Hint: Either G could be classified as B or B as G, and P(B) and P(G) are known.]

(d) If $P(G) = p$ rather than 0.25, are there values of p (other than 1) for which one would always classify a rock as granite?