Statistics 110 – Practice Midterm Solutions **Summer, 2006**

1. (20 points) One factory has four production lines to produce bicycles. Of the total production, line 1 produces 10%, line 2 produces 20%, line 3 produces 30% and line 4 produces 40%. The rate for defective products for these four production lines are 5%, 4%, 3%, and 2% respectively.

   (a) (5 points) What is the probability, p, that a randomly chosen bicycle is defective?

   $$p = 0.1 \times 0.05 + 0.2 \times 0.04 + 0.3 \times 0.03 + 0.4 \times 0.02 = 0.03$$

   (b) (5 points) If a bicycle is found defective, what is the probability that it comes from production line 4?

   $$P[\text{line 4}|\text{Defective}] = \frac{0.4 \times 0.02}{0.03} = 0.267 = \frac{4}{15}$$

   (c) (5 points) If an independent agency, like Consumers' Report buys 25 bicycles at random, what is the probability that none of them are defective.
   The number of defective bicycles, X, satisfies $X \sim Bin(25, 0.03)$. So

   $$P[X = 0] = 0.97^{25} = 0.467$$

   (d) (5 points) For the 25 bicycles mentioned in the previous part of this question, what are the mean and standard deviation of the number of defective ones?
   Since $X \sim Bin(25, 0.03)$

   $$E[X] = 25 \times 0.03 = 0.75 \qquad \text{and} \qquad \text{SD}(X) = \sqrt{25 \times 0.03 \times 0.97} = 0.853$$

2. (20 points) Let X be the time (in hours) it will take me to grade the first question on this exam and Y be the time (again in hours) to grade the second question. Assume that probability distribution for these times is described by

   $$f_X(x) = \frac{1}{3}(1 + 4x) \qquad \text{and} \qquad f_{Y|X}(y|x) = \frac{2y + 4x}{1 + 4x}$$

   for $0 < x < 1$ and $0 < y < 1$.

   (a) (5 points) Find $f_Y(y)$, the marginal density of Y.

   $$f_{X,Y}(x, y) = \frac{1}{3}(1 + 4x)\frac{2y + 4x}{1 + 4x} = \frac{2y + 4x}{3}$$

   $$f_Y(y) = \int_0^1 \frac{2y + 4x}{3} dx = \left.\frac{2yx + 2x^2}{3}\right|_0^1 = \frac{2y + 2}{3}$$

(b) (5 points) What is density function for the conditional distribution of $X|Y = y$?

$$f_{X|Y}(x|y) = \frac{(2y + 4x)/3}{(2y + 2)/3} = \frac{y + 2x}{y + 1}$$

(c) (5 points) What are the mean and variance of X?

$$E[X] = \int_0^1 x\frac{1}{3}(1 + 4x)dx = \frac{x^2}{6} + \frac{4x^3}{9}\Big|_0^1 = \frac{1}{6} + \frac{4}{9} = \frac{11}{18} = 0.611$$

$$E[X^2] = \int_0^1 x^2\frac{1}{3}(1 + 4x)dx = \frac{x^3}{9} + \frac{x^4}{3}\Big|_0^1 = \frac{1}{9} + \frac{1}{3} = \frac{4}{9} = 0.444$$

$$\text{Var}(X) = \frac{4}{9} - \left(\frac{11}{18}\right)^2 = 0.071$$

(d) (5 points) What is the probability that I can grade the first two questions in less than half an hour?

$$\begin{aligned} P[\text{Time } < 0.5] &= P[X + Y < 0.5] \\ &= \int_0^{0.5}\int_0^{0.5-x} \frac{2y + 4x}{3}dydx \\ &= \int_0^{0.5} \frac{y^2 + 4xy}{3}\Big|_0^{0.5-x} dx \\ &= \int_0^{0.5} \frac{(0.5 - x)^2 + 4x(0.5 - x)}{3}dx \\ &= \int_0^{0.5} \frac{1}{12} + \frac{x}{3} + x^2dx \\ &= \frac{1}{24} + \frac{1}{24} - \frac{1}{24} = \frac{1}{24} = 0.0417 \end{aligned}$$

3. (10 points) Suppose that for a model of car, the highway gas mileage X (in miles per gallon) is described by the density

$$f_X(x) = \frac{28 - x}{4}; \qquad 25 \leq x \leq 27$$

Suppose instead you wish to describe fuel use by

$$Y = \frac{100}{X}$$

the number of gallons needed to go 100 miles. What is the density function of Y?

$$g(x) = \frac{100}{X} \qquad g^{-1}(y) = \frac{100}{y}$$

$$\frac{d}{dy}g^{-1}(y) = -\frac{100}{y^2}$$

$$\begin{aligned} f_Y(y) &= f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| \\ &= \frac{28 - 100/y}{4} \frac{100}{y^2} \\ &= \frac{700y - 2500}{y^3}; \qquad \frac{100}{27} \leq y \leq 4 \end{aligned}$$

4. (25 points) Suppose that there are 10 people, including me, in class today and let X_i be the time (in minutes) it takes person i to get from home to class today. Assume that the times for each person can be described by an exponential distribution with mean 15 minutes and that the times for each person are independent of each other.

   (a) (5 points) What is the probability that it took me more than 15 minutes to get to get to class this morning?
   $X_i \sim Exp(1/15)$ which has a CDF $F_X(x) = 1 - e^{-x/15}$

   $$P[X_1 > 15] = 1 - (1 - e^{-15/15}) = e^{-1} = 0.3679$$

   (b) (5 points) Let V be the shortest travel time of the 10 people. What is $P[V \geq 15]$?

   $$P[V \geq 15] = P[X_1 \geq 15] \times \ldots \times P[X_{10} \geq 15] = (e^{-1})^{10} = 0.3678^{10} = 0.0000454$$

   (c) (5 points) Find $F_V(v)$, the cumulative distribution function of V.

   $$F_V(v) = 1 - (1 - F_X(v))^{10} = 1 - (e^{-v/15})^{10} = 1 - e^{-v/1.5}$$

   (d) (5 points) What is $E[V]$? (Hint: What distribution has the CDF calculated in the previous part.)
   The CDF in the previous part is that of an $Exp(1/1.5)$ distribution so the mean is 1.5. This result can also be derived by

   $$f_V(v) = \frac{d}{dv} 1 - e^{-v/1.5} = \frac{1}{1.5} e^{-v/1.5}$$

   $$\begin{aligned} E[V] &= \int_0^\infty v \frac{1}{1.5} e^{-v/1.5} dv \\ &= ve^{-v/1.5}\Big|_0^\infty + \int_0^\infty e^{-v/1.5} dv \\ &= -1.5e^{-v/1.5}\Big|_0^\infty = 1.5 \end{aligned}$$

(e) (5 points) Describe a situation where the assumption of independence of the travel times would not be reasonable.

One example would be if two people in the class happened to live in the same dormatory and decided to walk to class together. This would lead to them having the same time which clearly violates the assumption of independence.