ECON 4750/6750 **TEST #1** **FALL 2009**

Instructions: All questions must be answered on this examination paper; no extraneous sheets of paper will be graded. Point values for each question are in parentheses to the left of the corresponding numerals and should guide the allocation of your time. You may use your textbook and class notes, but you are not permitted to collaborate with others. Show all of your work. Time limit: 75 minutes.

(30) 1. Sandra Brown is a prolific writer of paperback mystery novels. The number of novels she writes each year is a random variable (X) generated by the following probability density function, f(X):

	X			
	2	3	4	5
f(X)	1/3	1/4	1/6	1/4

a. What is the probability that Sandra writes at least 4 novels in a given year?

$$P(X > 3) = P(X = 4) + P(X = 5)$$
$$= \frac{1}{6} + \frac{1}{4}$$
$$= \frac{5}{12}$$

b. What is the mean number of novels that Sandra writes each year?

$$E(X) = \sum_X X.f(X) = 2.\frac{1}{3} + 3.\frac{1}{4} + 4.\frac{1}{6} + 5.\frac{1}{4} = 3 + \frac{1}{3}$$

c. Sandra is paid a $500,000 "retainer" each year by her publisher, for which she is obligated to write two novels. In addition, Sandra receives $1,000,000 for each novel she writes. Denote by Y the random variable defined as Sandra's annual earnings from her novel writing. Calculate E(Y), the mean of her annual earnings.

Let $Y = 500{,}000 + 1{,}000{,}000.X$

$$E(Y) = 500{,}000 + 1{,}000{,}000.E(X)$$

But $E(X) = 3\frac{1}{3}$ (from part b), so:

$$E(Y) = 500{,}000 + 1{,}000{,}000.(3\frac{1}{3})$$
$$E(Y) = 3{,}833{,}333$$

(30) 2. A randomized clinical trial involving 72 former cocaine users was conducted to study the effectiveness of two antidepressants in preventing a relapse in (return to) using cocaine during a six-month period. The subjects were divided into three equal-sized treatment groups. One group received the antidepressant fluoxetine, and another was administered lithium. The third, "control" group was given a placebo (a sugar pill).

Let X denote a random variable specifying the treatment given an individual subject, with X = 1 indicating that fluoxetine was administered, X = 2 representing that the subject received lithium, and X = 3 denoting that a placebo or "control" pill was given to the subject. Let Y denote a random variable describing the outcome of the experiment for an individual subject, with Y = 0 if the individual did not suffer a relapse and Y = 1 if the individual relapsed, returned to using cocaine during the study period. The outcome of this trial is presented in tabular form as a joint probability density function f(X,Y), below:

		X 1	2	3	g(*Y*)
Y	0	2/9	1/12	1/36	1/3
	1	1/9	1/4	11/36	2/3
	h(*X*)	1/3	1/3	1/3	1

a. Calculate the probability that a subject suffered a relapse, given that he or she received fluoxetine during the clinical trial.

$$P(X=1 \mid Y=1) = \frac{f(1,\ 1)}{h(1)} = \frac{1/9}{1/3} = \frac{1}{3} = P(Y=1 \mid X=1)$$

b. What is the probability that an individual suffered a relapse, regardless of the method of treatment?

$$P(Y=1) = g(1) = \sum_X f(X,1) = 2/3$$

c. Is the outcome of the clinical trial independent of the method of treatment? <u>Why or why not?</u>

Statistical independence requires

$$f(X,Y) = h(X) \cdot g(Y)$$

for all possible (X,Y) pairs in the joint p.d.f. $f(X,Y)$.

Let $X = 1$ and $Y = 1$. Then, we test

$$\frac{1}{9} = f(1,1) \neq h(1).g(1) = \frac{1}{3}.\frac{2}{3} = \frac{2}{9}$$

so X and Y are <u>not</u> statistically independent.

(20) 3. The tread life of Bridgestone Dueler H/L Alenza tires has a normal distribution with a mean of 40,000 miles and a standard deviation of 4,000 miles.

a. What is the probability that the tread life of a randomly selected tire of this type will be greater than 46,000 miles?

$$X \sim N(40K, 16M)$$

$$P(X > 46) = P(Z > \frac{46-40}{4}) = P(Z > 1.5)$$

$$\begin{aligned} P(Z > 1.5) &= 1 - P(Z < 1.5) \\ &= 1 - 0.9332 \\ P(Z > 1.5) &= 0.0668 = P(X > 46) \end{aligned}$$

b. What is the probability that the tread life of a randomly selected tire of this type will be between 36,000 and 43,000?

$$P(36 < X < 43) = P(\frac{36-40}{4} < Z < \frac{43-40}{4})$$

$$\begin{aligned} P(-1 < Z < \frac{3}{4}) &= P(Z < 0.75) - P(Z < -1) \\ &= 0.7734 - 0.1587 \\ &= 0.6147 = P(36 < X < 43) \end{aligned}$$

(20) 4. Among all U.S. workers employed full-time, year-round in 2006, the population regression function for the relationship between annual earnings Y (measured in thousands of dollars) and gender X is

E(Y) = 31.2 + 7.4X,

where X = 1 if the worker is male and X = 0 if the worker is female.

a. What are the mean earnings of males? What are the mean earnings of females?

$E(Y \mid X = 1) = 31.2 + 7.4 = 38.6K$ (males)

$E(Y \mid X = 0) = 31.2K$ (females)

b. If males comprise 62% of all of these workers (male and female), what are the mean earnings of all workers?

$$E(Y) = 31.2 + 7.4(0.62) \Leftrightarrow E(Y) = 31.2 + 7.4.E(X)$$
$$E(Y) = 31.2 + 4.588$$
$$E(Y) = 35{,}788$$

(10) Extra Credit

Let X and Y denote two random variables. A unit-free measure of the strength of the linear association between X and Y is the correlation coefficient

$$\rho = \mathrm{Cov}(X,Y)/\sigma_x\ \sigma_y$$

where Cov(X,Y) is the covariance between X and Y, σ_x is the standard deviation of X, and σ_y is the standard deviation of Y. The population regression function describing the linear relationship between Y and X is

$$Y = \beta_0 + \beta_1 X + \mu$$

and is interpreted as the conditional-mean function $E(Y \mid X)$ with $\beta_1 = \mathrm{Cov}(X,Y)/\mathrm{Var}(X)$, where $\mathrm{Var}(X) = (\sigma_x)^2$ is the variance of X, and μ is a random variable with $E(\mu \mid X) = 0$. Derive the relationship between β_1 and ρ in terms of σ_x and σ_y.

(1) $\rho = \mathrm{Cov}(X,Y)/\sigma_x\ \sigma_y \quad \Rightarrow \mathrm{Cov}(X,Y) = \rho.\ \sigma_x\ \sigma_y$

(2) $\beta_1 = \mathrm{Cov}(X,Y)/\mathrm{Var}(X) \quad \Rightarrow \mathrm{COV}(X,Y) = \beta_1.\mathrm{Var}(X)$

Together, (1) and (2) imply:

$$\rho.\ \sigma_x\ \sigma_y = \beta_1.\mathrm{Var}(X)$$

$$\rho.\ \sigma_x\ \sigma_y = \beta_1.\ \sigma_x^2$$

$$\rho.\ \sigma_y / \sigma_x = \beta_1$$