NAME: SOLUTIONS RPI ID #: ____________

ENGR 2600 Modeling and Analysis of Uncertainty Section 4 Fall 2008 Exam 2

Rules of the Game:

1. Work entirely alone. Do not give or solicit assistance from any other student.
2. Do not open any IM or email programs or web browsers. Turn off cell phones. Running such software will be regarded as evidence of cheating.
3. This exam is open-book/open-notes/open-computer. Use any written materials you wish.
4. There are 5 questions, weighted equally. Answer all 5. Tip: Do the easier questions first.
5. You must show your work to receive credit, even if the number you write down is correct.
6. Feel free to use the restrooms as necessary.
7. If you have a question, bring it down front so as to minimize disruption.

We don't need no education.
We don't need no thought control.
No dark sarcasm in the classroom.
Teacher, leave them kids alone.
Hey, teacher, leave them kids alone!

-- Pink Floyd "The Wall, Pt.2"

Question 1. A particle travels on a discrete random walk in two dimensions according to the following joint probability mass function.

		Delta Y			
		-1	0	1	row sum
	-1	0.07	0.17	0.10	0.33
Delta X	0	0.27	0.03	0.03	0.33
	1	0.10	0.17	0.07	0.33
	col sum	0.43	0.37	0.20	1.00

a) What is the probability that the particle does not move at all in one time step?

$P(\Delta x=0, \Delta y=0) = 0.03$

b) What is the probability that the particle would remain in the same spot for 2 successive time periods, assuming independence between movements at different times?

$P(2 \text{ non-moves}) = 0.03^2 = 0.0009$

c) If a particle moves right one unit along the X axis, what is the probability distribution of its movement along the Y axis?

$P(\Delta y | \Delta x = +1) = \left\{\frac{0.10}{0.33}, \frac{0.17}{0.33}, \frac{0.07}{0.33}\right\}$

$= \{0.30, 0.50, 0.20\}$

d) If the X and Y movements were independent, what would be the answer to part (b)?

$P(\Delta y) = \{0.43, 0.37, 0.20\}$

e) Compute E[Delta Y | Delta X = -1].

$E[\Delta y | \Delta x = -1] = -1 \times \frac{0.07}{0.33} + 0 \times \frac{0.17}{0.33} + 1 \times \frac{0.10}{0.33}$

$= \frac{0.10 - 0.07}{0.33} = \frac{0.03}{0.33} = 0.0909$

Question 2. Briefly explain, using words and/or equations, what you would do to…

a) Estimate the standard error of a sample mean.

$$SE(\bar{x}) = \frac{s}{\sqrt{n}}$$

b) Estimate the standard error of the reciprocal of a sample mean.

Bootstrap

1. Resample the sample data
2. Compute $1/\bar{x}$ for the bootstrap sample
3. Repeat 1 and 2 many times $(B > 200)$
4. Calculate the sample std dev of all the values of $1/\bar{x}$

c) Estimate the bias in a sample standard deviation. Say sample std dev = S

Bootstrap

1. Resample the sample data
2. Compute S^* for each bootstrap sample
3. Repeat 1 and 2 many times
4. Average the bootstrap values of S^*
5. Bias $\approx$ Avg. of bootstrap values $S^* - S$

d) Summarize the degree of association between two random variables.

Compute the correlation coefficient

$$\rho_{xy} = \frac{Cov[X,Y]}{\sqrt{V[X]V[Y]}}$$

e) Learn about the sampling distribution of a new statistic θ computed from n=7 sample values from a lognormal distribution.

Monte Carlo

1. Generate 7 sample values from the lognormal distribution
2. Compute the point estimate $\hat{\theta}$ from the sample data
3. Repeat 1 and 2 many times.
4. Plot all the values of $\hat{\theta}$ in a histogram.

Question 3. A rectangle has random height H and random width W, where

$$H \sim \text{Normal}(10,2^2),\ W \sim \text{Normal}(20, 3^2), \text{ and } \rho_{LH} = -0.8.$$

Consider the perimeter of this random rectangle.

a) What is the name of the distribution of the perimeter?

Normal

b) What is the expected value of the perimeter?

$E[P] = 2E[H] + 2E[W] = 2\times 10 + 2\times 20 = 60$

c) What is the variance of the perimeter?

$V[P] = 2^2 V[H] + 2^2 V[W] + 2\cdot 2\cdot 2\rho_{HW}\sqrt{V[H]V[W]}$

$= 4\times 4 + 4\times 9 + 4\times(-0.8)\times\sqrt{4\times 9}$

$= 32.8$

Question 4. Consider the coefficient of variation (CV) of samples of size n=5 from an exponential distribution with mean 2. Use Monte Carlo simulation row-wise (i.e., no need for scripts) with 1,000 replications (rows) to answer these questions:

a) From theory, determine the true (infinite data) value of the CV for this situation.

$$CV = \frac{\sigma}{\mu} = \frac{2}{2} = 1.0$$

since $SD[x] = E[x] = \frac{1}{\lambda}$ for exponentials

b) Estimate the mean of the CV using the simulated data.

$\bar{x} = 0.87228$ or 0.85447 in a repeated experiment

c) Estimate the bias in the point estimator.

$$\text{Bias} = E[\hat{\theta}] - \theta \approx \bar{x} - 1.0 = 0.87228 - 1 = -0.13$$

or -0.15

d) Report the standard error of the estimate of the mean.

$$SE(\bar{x}) = \frac{s}{\sqrt{n}} = \frac{0.28927}{\sqrt{1000}} = 0.00915 \quad \text{or} \quad \frac{0.27743}{\sqrt{1000}} = 0.00877$$

e) What would you expect the answer to part (d) to be if you ran the Monte Carlo simulation for 4,000 replications instead of 1,000?

$$SE(\bar{x}) = \frac{s}{\sqrt{4 \cdot n}} = \frac{1}{2} \times \frac{s}{\sqrt{n}} = \frac{1}{2} \times 0.00915 = 0.004575$$

SE will be cut in half.

Question 5. Read in the Devore textbook dataset *Ex01-44.MTW*. The variable is the oxygen consumption rate of firefighters.

a) Determine the 95% CI for the mean.

(26.0066, 36.0534)

or $\bar{X} \pm t_{n-1,\alpha/2} \frac{s}{\sqrt{n}} = 31.03 \pm 2.262 \times \frac{7.0222}{\sqrt{10}}$

b) Report the point estimate of the median.

$\tilde{X} = 29.45$

c) Estimate the standard error of the point estimate of the median using at least 20 bootstrap replications. Copy down the Minitab commands (script) you used.

$SE(\tilde{X}) = 1.385$ by bootstrap