Public Affairs 818
Professor: Geoffrey L. Wallace
October 9th, 2008

Midterm Exam #1

Instructions: You have 120 minutes to complete the examination and there are 6 questions worth a total of 120 points. The questions do not carry equal weight, so allocate your time wisely. You may use a calculator, but you may not use pre-programmed formulas. The standard normal 'Z' table from the second page of your book has been provided to you. Additionally, a table of binomial probabilities for use with question 6 is provided. Partial credit is possible on all questions so be sure to write clearly and show all of your work. If you have questions regarding the wording of a particular question, feel free to ask for clarifications. Good luck!!

<u>Question 1. Water Pollution (25 points)</u>

Low stream flow estimates are crucial for water quality management and issuing and/or renewing National Pollution Discharge Elimination System (NPDES) permits, planning water supplies, hydropower, and irrigation systems, and for assessing the impact of prolonged droughts on aquatic ecosystems. Suppose you have been asked (or appointed) to sit on a committee in charge of managing water and water pollution. Quite possibly, the committee would include scientific experts in the field of water and water pollution, but hopefully your understanding of statistics learned in PA 818 would allow you to engage the experts in a meaningful dialogue on the topics. Scientists will often use the lognormal distribution to describe the behavior of certain variables in environmental problems. You may also see problems presented where the natural log of a variable X is distributed normally with given mean and variance. Let $X =$ the hourly output of water pollution from a paper mill. Suppose the natural log of water pollution in a river under your committee's jurisdiction is normally distributed with a mean (μ) of 3 units of pollution and a variance (σ^2) of 3.

For the questions below you may need the following information:

- $\ln(1) \approx 0$
- $\ln(2.5) \approx 0.92$
- $\ln(1.44) \approx 0.36$
- $\ln(10) \approx 2.30$
- $\ln(30) \approx 3.40$
- $\ln(35) \approx 3.55$
- $\ln(40) \approx 3.69$
- $\ln(90) \approx 4.50$
- $\ln(100) \approx 4.60$

A) **(5 points)** Find the probability that X is greater than 30 units per hour.

B) **(5 points)** Find the probability that X is at most 10 or at least to 100 units per hour.

C) **(5 points)** Find the probability that X exceeds its mean. (**Hint:** The mean of is found through the following expression):

$$E(X) = e^{\mu + \frac{\sigma^2}{2}} \approx 90$$

D) **(10 points)** The committee in your jurisdiction is interested in estimating an upper bound for the natural log of water pollution in your river. They ask you to work out a value for $\ln(x)$ such that the natural log of the hourly output of water pollution will exceed this level less than 2.5% of the time. What is this value of $\ln(x)$?

Question 2. Fundamentals (15 points)

$X \sim N(\mu = 100, \sigma^2 = 9)$. Let $Y = 2 \cdot X + 10$.

A) **(5 points)** What is the distribution of Y ? (A complete answer will specify the distribution, the mean, and the variance of Y.)
B) **(5 points)** What is the probability that Y is smaller than 204?
C) **(5 points)** What is the probability that Y is greater than 204 or smaller than 201?

Question 3. More Fundamentals (10 points)

Use a Venn diagram to derive the formula for $P(A \cup B \cup C)$ where A, B, and C are not mutually exclusive. (**Hint:** We did this in class for $P(A \cup B)$ and were able to show that $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. In this case we had to subtract off the intersection because it was double counted in the sum of $P(A)$ and $P(B)$. The answer for this question is going to take the form $P(A \cup B \cup C) = P(A) + P(B) + P(C) \pm$ *other stuff*, where *other stuff* consist probabilities of intersection involving two or more of the sets A, B, and C)

Question 4. Drunk Drivers (30 points)

On a recent holiday evening, a sample of 500 drivers was stopped by the police. Three hundred were under 30 years of age. A total of 250 were under the influence of alcohol. Of the drivers under 30 years of age, 200 were under the influence of alcohol.

Let A be the event that a stopped driver was under the influence of alcohol.
Let Y be the event that a stopped driver is less than 30 years old.

The following table might help you solve the problem.

	under influence (A)	not under influence (A^C)	Total by age
<30 years old (Y)	200	100	300
≥30 years old (Y^C)	50	150	200
Total by influence	250	250	500

A) **(5 points)** Determine P(A) and P(Y).
B) **(5 points)** What is the probability that a driver is under 30 and not under the influence of alcohol?
C) **(5 points)** Given that a driver is not under 30, what is the probability that he/she is under the influence of alcohol?
D) **(5 points)** What is the probability that a driver is under the influence of alcohol if we know the driver is under 30?
E) **(5 points)** Define mutual exclusivity. Mathematical and english definitions are acceptable. Are A and Y mutually exclusive events? Explain.
F) **(5 points)** Are A and Y independent events? Explain. What is the relationship between the mutual exclusivity of two events and their independence? Mathematical and verbal explanations are acceptable.

Question 5. Drug Sniffing Dogs (20 points)

At the border between the United States and Mexico, United States drug enforcement agents use dogs that are specially trained to identify the presence of narcotics in the automobiles crossing the border. Suppose that there are 18 dogs at a specific border checkpoint, but only 6 have the ability to find cocaine. (Assume that the other 12 dogs have not completed their training and cannot identify cocaine). Due to a tip from an informer at 7:59 a.m., agents at the checkpoint have been told that a large shipment of cocaine will arrive every minute from 8:00 to 8:17 a.m. in a series of purple trucks. Agents only have a brief period to prepare for the shipment and cannot add more dogs trained in cocaine identification. Because of the limited resources at their disposal, agents have decided to target drug enforcement efforts from 8:00 to 8:17 on purple trucks crossing the border. What this means is that only purple trucks will be searched during this time frame.

- Assume that once a dog begins a search on one vehicle, it cannot search other vehicles. The implication of this is that ***no one dog will search more than one vehicle.***
- Assume that if a dog does not find cocaine, then the truck will be released, free to continue into the United States unhindered.
- Assume that the dogs are randomly assigned to purple trucks.

A drug dealer owns 18 purple trucks and decides to send all of his purple trucks filled with cocaine across the border *in succession* for every minute starting at 8:00 and ending at 8:17. You may assume that no other purple trucks cross the border in this time frame.

You could think of the purple trucks attempting to cross the border as an experiment with 18 trials. From the point of view of the customs agents, success on any one trial occurs if a purple truck crossing the border between 8:00 and 8:17 is found to be carrying drugs.

A) (**5 points**) Define statistical independence. Mathematical and english definitions are acceptable. Are the trials in the experiment described above independent? Why or why not?
B) (**5 points**) If the first 7 trucks have driven through and none of them have been caught, what is the probability that both the 8th and 9th trucks will be caught?
C) (**5 points**) Find the probability that the dogs catch the first three trucks that are searched.
D) (**5 points**) If the dogs catch the first 3 trucks, what is the probability that the 5th truck is caught?

Question 6. More Roulette (20 points)

A roulette wheel consists of 38 slots, each containing a color and a number. Eighteen of the slots in a roulette wheel are red, 18 are black and 2 are green. In addition to each slot being colored red, black, or green they are numbered 1 through 38. The likelihood of the ball landing in any of the wheels 38 slots is $1/38$. In roulette gamblers can make several types of bets. They can bet either red or black. These bets pay \$1 for each \$1 wagered. Gamblers can also bet on any specific number 1 through 38. These bets pay \$35 for each \$1 wagered.

A) **(5 points)** If the gambler bets 7, the gambler wins if the ball rolls into a slot numbered 7. If it rolls into any other slot, the gambler loses. If the amount bet in the $i'th$ game is \$10, and the random variable W_i represents the amount of money won or lost on the gamblers bet, find $P(W_i = \$350)$, $P(W_i = -\$10)$, $E(W_i)$ and $Var(W_i)$ assuming that the gambler bets 7 in the $i'th$ game

B) **(5 points)** Assume the gambler bets 7 on 100 rolls of the roulette wheel. Find the gambler's expected winnings and the variance of winnings.

C) **(10 points)** Your close friend Degen returns from Las Vegas. After 2400 \$10 bets on 7 and countless hours of degenerate gambling he managed to lose \$8,880. Shortly after he returns from his trip a roulette wheel scandal breaks. Apparently, the largest manufacture of roulette wheels (Roulette Wheel Inc.) has been making defective wheels. In the defective wheels the probability of the ball landing on the 7 slot is only 1/50. Degen "feels" strongly that he was playing on a defective roulette wheel. Nevada Gaming Commission records indicate that 25% of the large number of roulette wheels at the casino where Degen was playing were defective. Use Bayes' rule and the attached binomial probability table to calculate the probability that he was playing on a defective wheel (**Hint:** Hyperbinbay)