MAT 2379A
Midterm Examination
(with solutions)

October 23, 2013 **Professor Raluca Balan**
Time: 80 minutes

Student Number:________________

Family Name: ________________ **First Name:** ________________

This is a closed book examination. A formula sheet and some statistical tables are included with the exam. Only TI30 and Casio calculators are permitted. Record your answer to each question in the table below.

Question	Answer
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	

NOTE: At the end of the examination, hand in only this page. You may keep the questionnaire.

.**********************.

1. Let X be a discrete random variable with values $0, 1, 2, 3, 4$. The cumulative distribution function of X is given by the table below:

x	0	1	2	3	4
$P(X \leq x)$	0.5	0.6	0.7	0.9	1

Find the expected value of X and the variance of X.

A) $E(X) = 1.3$, $\mathrm{Var}(X) = 0.61$ B) $E(X) = 0.7$, $\mathrm{Var}(X) = 2.21$
C) $E(X) = 1.3$, $\mathrm{Var}(X) = 0.61$ D) $E(X) = 0.5$, $\mathrm{Var}(X) = 1.1$
E) $E(X) = 1.3$, $\mathrm{Var}(X) = 2.21$

Solution: First, we find the table of the probability mass function of X:

x	0	1	2	3	4
$P(X = x)$	0.5	0.1	0.1	0.2	0.1

$$E(X) = 0(0.5) + 1(0.1) + 2(0.1) + 3(0.2) + 4(0.1) = 1.3$$

$$\mathrm{Var}(X) = 0^2(0.5)+1^2(0.1)+2^2(0.1)+3^2(0.2)+4^2(0.1)-(1.3)^2 = 3.9-1.69 = 2.21$$

The answer is E.

2. A new screening test is proposed for tuberculosis. To test its effectiveness, the screening test is applied to 200 tuberculosis patients and to 200 persons selected randomly from the community (called "controls"), who do not have tuberculosis. The following table summarizes the test results:

	Tuberculosis patients	Community controls	Total
Positive tests	180	40	220
Negative tests	20	160	180
Total	200	200	400

What is the sensitivity of this test?

A) 0.90   B) 0.80   C) 0.82   D) 0.89   E) 0.10

Solution:

$$\text{sensitivity} = P(\text{Test} + |\text{True}+) = \frac{P(\text{True} + \text{ and Test}+)}{P(\text{True}+)}$$

$$= \frac{180/400}{200/400} = \frac{180}{200} = 0.90$$

The answer is A. Note that the incorrect answers B, C, D give the specificity, PPV, respectively NPV of the test.

3. Use the data from Question 2. In a certain poor country, the prevalence of tuberculosis is estimated to be 3.3%. If the above screening test was applied to the general population in this country, what proportion of those tested would test positive?
   Hint: Assume that the test has the same sensitivity and false-positive rate as in Question 2.

   A) 0.1492 B) 0.0297 C) 0.2000 D) 0.1934 E) 0.2231

   Solution: We know that $P(\text{True}+) = 0.033$. By the total probability rule,

   $$P(\text{Test}+) = P(\text{Test} + |\text{True}+)P(\text{True}+) + P(\text{Test} + |\text{True}-)P(\text{True}-)$$

   $$= (0.9)(0.033) + (0.2)(0.967) = 0.2231$$

   where we used $P(\text{Test} + |\text{True}-) = 40/200 = 0.2$. The answer is E.

4. In humans, both widow's peak and cleft chin are dominant inherited traits. A widow's peak consists in a V-shaped point in the hairline in the center of the forehead, whereas a cleft chin is a Y-shaped fissure on the chin. In a couple, the woman has a widow's peak and a cleft chin, and she is heterozygous for both traits. The man does not have a widow's peak and he does not have a cleft chin. What is the probability that their offspring has a widow's peak and a cleft chin?

   A) 1/2 B) 1 C) 3/4 D) 0 E) 1/4

   Solution: We denote by P and p the alleles for the widow's peak, respectively the absence of the widow's peak. We denote by C and c the alleles for the cleft chin, respectively the absence of the cleft chin. The woman's genotype is $PpCc$. The man's genotype is $ppcc$. Using a tree diagram we see that the possible genotypes (and phenotypes) for their offspring are: $PpCc$ (widow's peak, cleft chin), $Ppcc$ (widow's peak, no cleft chin), $ppCc$ (no widow's peak, cleft chin), $ppcc$ (no widow's peak, no cleft chin). The probability is 1/4. The answer is E.

5. Assume that the cholesterol levels of adult women can be modeled with a normal distribution with a mean of 188 mg/dL and a standard deviation of 24 mg/dL. We want to find a level l such that 15% of women have a cholesterol level larger than l. Which one of the following R commands gives the correct value for l.

   A) qnorm(0.15,188,24)
   B) qnorm(0.85,188,576)
   C) qnorm(0.85)
   D) qnorm(0.15,188,576)
   E) 24*qnorm(0.85,0,1)+188

   Solution: The cholesterol level X is a normal random variable with mean 188 and standard deviation 24. We want to find l such that $P(X > l) = 0.15$. This means that $P(X \leq l) = 0.85$. By standardization,

$$P(Z < \frac{l-188}{24}) = 0.85$$

   Hence

$$\frac{l-188}{24} = z$$

   where $z = \mathrm{qnorm}(0.85, 0, 1)$. It follows that $l = 24z + 188$. The answer is E.

6. A new rule in Ontario forbids drinking of *any* amount of alcohol while driving. Many people decided to drive on New Year's Eve 2013. Among these, 32% had some alcohol, 9% had an accident and 6% had an accident and were found to have some alcohol in their blood. What is the percentage of drivers who did not drink and did not have an accident on New Year's Eve?

   A) 35% B) 65% C) 47% D) 53% E) 33%

   Solution: We select a driver at random. Let A be the event that this person had some alcohol and B the event that this person had an accident. We know that $P(A) = 0.32$, $P(B) = 0.09$ and $P(A \cap B) = 0.06$. By the addition rule, the probability that a driver drank or had an accident is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.32 + 0.09 - 0.06 = 0.35$$

The desired probability is $P(A' \cap B') = 1 - P(A \cup B) = 1 - 0.35 = 0.65$. The answer is B.

7. Canadian farmers grow a large quantity of genetically modified corn. About 19% of corn is genetically modified to be resistant to insects and about 70% of corn is genetically modified to be resistant to herbicides. Among the corn that are genetically modified to be resistant to herbicides, about 15% is genetically modified to be resistant to insects. If you randomly select an ear of corn that you know has been modified to be resistant to insects, what is the probability it has also been modified to be resistant to herbicides?

   A) 0.8867 B) 0.5526 C) 0.7895 D) 0.1500 E) 0.1050

   Solution: Let A and B be the events that the corn is genetically modified to be resistent to insects, respectively herbicides. We know that $P(A) = 0.19$, $P(B) = 0.70$ and $P(A|B) = 0.15$. By Bayes' rule, the desired probability is:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A)} = \frac{(0.15)(0.70)}{0.19} = 0.5526$$

   The answer is B.

8. In the recent years, a large number of ash trees in Ottawa were infested with an insect called the emerald ash borer. The older the tree is, the larger is its diameter and the higher is its chance of being infested. In a certain neighborhood, all ash trees were planted at about the same time, around the year 1960. Their diameters follow a normal distribution with mean $\mu = 69.9$ cm and standard deviation $\sigma = 5$ cm. Five ash trees are randomly selected from this neighborhood. What is the probability that at most one tree has a diameter greater than 73 cm?

   A) 0.5957 B) 0.2108 C) 0.3850 D) 0.7324
   E) 0.2676

   Solution: Let X be the diameter of a tree. By standardization and Table 17.3,

$$P(X > 73) = P\left(\frac{X - 69.9}{5} > \frac{73 - 69.9}{5}\right) =$$

$$P(Z > 0.62) = 1 - P(Z \le 0.62) = 1 - 0.7324 = 0.2676.$$

Let Y be the number of trees with diameter greater than 73 cm, in the sample of 5. Y has a binomial distribution with $n = 5$ and $p = 0.2676$. The desired probability is:

$$P(Y \leq 1) = P(Y = 0) + P(Y = 1) = (0.7324)^5 + 5(0.2676)(0.7324)^4 = 0.5957$$

The answer is A.

9. The October 2013 issue of the National Geographic Magazine features the survival story of the Kihanzi spray toad from Tanzania, which was saved from extinction due to conservation efforts. In 2009, the species was declared extinct in the wild. Luckily, 500 toads had been captured in 2004, half of them being sent to the Bronx Zoo, and the other half to the Toledo Zoo. Some of these toads became infected with a deadly fungus which was devastating amphibian populations worldwide. Assume that the infection rate is 75.7% among the spray toads at the Bronx Zoo, and 84.3% at the Toledo Zoo. We select randomly a toad among the 500. What is the probability that it is infected? Is the fungus infection of the spray toads independent of the zoo location?

   A) 0.8, yes B) 0.8, no C) 0.5, yes D) 0.5, no
   E) not enough information given

   Solution: Let I be the event that a randomly chosen toad is infected. Let B be the event that the toad comes from the Bronx Zoo and T the event that the toad comes from the Toledo Zoo. We know that $P(B) = P(T) = 0.5$, $P(I|B) = 0.757$ and $P(I|T) = 0.843$. By the total probability rule,

$$P(I) = P(I|B)P(B) + P(I|T)P(T) = (0.757)(0.5) + (0.843)(0.5) = 0.8$$

   Since $P(I) \neq P(I|B)$, I is not independent of B. The answer is B.

10. A forester measured 27 of the trees in a large wooded area that is up for sale. He found a mean diameter of 26 cm and a standard deviation of 12 cm. Suppose that these trees provide an accurate description of the whole forest and that the diameter of a tree is approximately normally distributed. What percent of these trees should be over 35 cm in diameter?

   A) 0.8413 B) 0.7500 C) 0.7734 D) 0.2266 E) 0.1587

   Solution: Let X be the diameter of a tree. We know that X has a normal distribution with mean $\mu = 26$ and standard deviation $\sigma = 12$. By standardization and Table 17.3,

$$P(X > 35) = P\left(\frac{X-26}{12} > \frac{35-26}{12}\right) = P(Z > 0.75)$$

$$= 1 - P(Z \leq 0.75) = 1 - 0.7734 = 0.2266$$

The answer is D.

11. Consider the following R output:

```
> pbinom(20,50,0.3)
[1] 0.9522362
> pbinom(19,50,0.3)
[1] 0.9151974
> pbinom(18,50,0.3)
[1] 0.8594401
> pbinom(15,50,0.3)
[1] 0.5691784
> pbinom(14,50,0.3)
[1] 0.4468316
> dbinom(15,50,0.3)
[1] 0.1223469
```

Let X be a binomial random variable with $n = 50$ and $p = 0.3$. Using the R output above, calculate $P(15 < X \leq 19)$.

A) 0.468 B) 0.3460 C) 0.793 D) 0.944 E) 0.918

Solution: We have

$$P(15 < X \leq 19) = P(X = 16) + P(X = 17) + P(X = 18) + P(X = 19) =$$

$$P(X \leq 19) - P(X \leq 15) = 0.9151974 - 0.5691784 = 0.3460190$$

The answer is B.

12. Canadian Blood Services tests blood donations for West Nile Virus according to the following protocol. Samples from units of blood are pooled, six units at a time, and the pool of six is tested for the presence of West Nile Virus. If any unit comes from a donor with West Nile Virus, the pool tests positive.

a) A blood drive is being held in a particular town. Suppose the actual prevalence of West Nile Virus in the town is 1 in 500. Assuming independence between donors, find the probability that a randomly selected pool of six units test positive.

b) A blood drive is held in the town and 216 people donate. Find the probability that at least one pool of 6 units tests positive from the blood drive.

A) a) 0.00200 ; b) 0.432

B) a) 0.01194 ; b) 0.351

C) a) 0.98806 ; b) 0.282

D) a) 0.01188 ; b) 0.417

E) a) 0.01200 ; b) 0.070

Solution: a) Let X be the number of units which test positives in a pool of 6. X is a binomial random variable with $n = 6$ and $p = 1/500 = 0.002$. The required probability is

$$P(X \geq 1) = 1 - P(X = 0) = 1 - (1 - 0.002)^6 = 0.01194$$

b) There are 216/6=36 pools. By part a), each one has probability 0.01194 of testing positive. Let Y be the number of pools (among the 36) which will test positive. Y is a binomial random variable with $n = 36$ and $p = 0.01194$. The required probability is

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - (1 - 0.01194)^{36} = 0.351$$

The answer is B.