STAT235 MIDTERM EXAM VERSION 111 SOLUTION

Complete the following (please **print**):

Section Number: ___________ Surname, First name: ______________________________

I.D. Number: ___________

Instructions

1. Enter your lecture section number, your last name (the same one that is illustrated on your One Card), first name (the same one that is illustrated on your One Card), and student ID # into the space provided above.

2. Make sure you use ONLY PENCIL to put and mark your name, student ID #, and your exam version number on the scantron sheet. Your name (surname and first name) should be entered into NAME block, your student ID # into IDENTIFICATION NUMBER block and finally the three-digit exam version into SPECIAL CODES block. The exam version is specified at the top of the exam. Make sure to shade in the circle that corresponds to the letter or digit in the box at the top of each column.

3. This is a multiple-choice closed book exam. There are **30** questions in the exam. For each question, carry out the appropriate analysis and put your answer on the scantron sheet by shading the letter A, B, C, D or E that corresponds to your chosen answer. Make sure your answers are clearly marked with ONLY PENCIL. Otherwise, no marks will be given. For each question exactly one of the five answers is correct. If you fill in more than one answer to a question, the question will be scored incorrect. Each question is worth 1 mark. All answers are rounded.

4. Please note that only the answers in the scantron sheet will be considered. If you initially mark your answers in the exam sheet, make sure that you copy them correctly into the scantron sheet. If you change an answer on the scantron sheet, be sure that you erase your first mark completely; then blacken the circle of the answer choice you prefer. Your score will be based on the number of questions you answer correctly. No credit will be given for omitted answers and no credit will be lost for wrong answers. You must mark all your answers on the scantron sheet during the allotted time. No additional time will be allowed at the end of the session for this purpose.

5. You are permitted to use a **non-programmable** calculator approved by the Faculty of Engineering. During the lab exam you are forbidden to use any devices with communication capabilities including cell phones and pagers. You are also forbidden to use any photographically capable devices in the exam room. Copying questions or answers on paper to take from the exam room is prohibited.

6. Note that the formula sheet and the table of the cumulative distribution function of the standard normal distribution are attached to the exam. There are **6** pages in the exam. The exam is graded out of **30** points.

7. You must return your scantron and exam booklet when you finish the exam. You have 80 minutes to complete the exam.

8. Sign the exam booklet in the space provided below.

SIGNATURE: _________________________________

PART 1 (MULTIPLE-CHOICE QUESTIONS, 1 POINT EACH)

1. Which of the following statements is CORRECT?

   (a) For a left-skewed distribution, the mean is usually greater than the median.
   (b) The students in this class only represent a sample.
   (c) Temperature is a categorical variable.
   (d) The population of Alberta can be a population or a sample.
   (e) A bar chart represents numerical data, but a pie chart represents categorical data.

   Solution: See definitions of sample, population, and categorical variables. Also see relationship of measures of center for a left-skewed distribution.

2. Let X and Y denote the amounts of two different types of impurities in a batch of a certain chemical product at Axis Chemicals. If X and Y are independent random variables with standard deviations of 1 and 4, respectively, what is the standard deviation of the random variable $W = 3X - Y + 5$?

   (a) 2 (b) 25 (c) 3.16 (d) 4 **(e) 5**

   Solution: $Var(W) = Var(3X - Y + 5) = 3^2 Var(X_1) + (-1)^2 Var(Y) + Var(5) = 9(1^2) + (1)(4^2) + 0 = 25$
   $\sigma = \sqrt{Var(W)} = \sqrt{25} = 5$

3. Let the probability of throwing a piano, or T, be 0.18. Also, the probability of being Superman's son, or Su, is 0.08. If the probability of throwing a piano, given you are Superman's son, is 0.84, one can conclude that

   (a) T and Su are independent events.
   (b) $P(T \text{ or } Su)$ is greater than 0.25.
   (c) $P(T \text{ and } Su)$ is less than 0.25.
   (d) T and Su are disjoint (mutually exclusive) events.
   (e) $P(Su \mid T)$ is less than 0.25.

   Solution:
   $P(T) = 0.18$, $P(Su) = 0.08$, $P(T \mid Su) = 0.84$ → $P(T) \neq P(T \mid Su)$, so T and Su are not independent.
   $P(T \cap Su) = P(Su) \times P(T \mid Su) = 0.08 \times 0.84 = 0.0672$ → not zero, so T and Su are not disjoint.
   $P(T \cup Su) = P(T) + P(Su) - P(T \cap Su) = 0.18 + 0.08 - 0.0672 = 0.193$
   $$P(Su \mid T) = \frac{P(Su \cap T)}{P(T)} = \frac{0.0672}{0.18} = 0.373$$

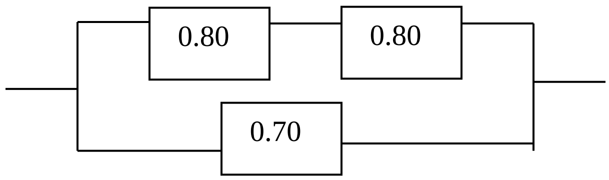
4. A bin contains 5 defective transistors (they immediately fail when put in use), 10 partially defective transistors (they fail after a couple of hours in use), and 25 acceptable transistors. A transistor is chosen at random from the bin and put into use. If it does not immediately fail, what is the probability that it is acceptable?

   (a) 25/40 **(b) 25/35** (c) 15/35 (d) 35/40 (e) 15/40

Solution: Let A be an acceptable transistor and F is a transistor that immediately fails.

$$P(A|F') = \frac{P(A \cap F')}{P(F')} = \frac{P(A)}{1-P(F)} = \frac{25/40}{1-(5/40)} = \frac{25}{35}$$

5. The following circuit, consisting of three devices, operates if and only if there is a path of functional devices from left to right. The probability that each device functions is provided below and the devices operate independently of one another.



The probability that the circuit operates is

(a) 0.328 (b) 0.928 **(c) 0.892** (d) 0.981 (e) 0.898

Solution: The probability that the circuit operates is P(top path ∪ bottom path).
P(top path ∪ bottom path) = P(top path) + P(bottom path) – P(top path ∩ bottom path)
= 0.80*0.80 + 0.70 – 0.80*0.80*0.70 = 0.892

6. Suppose X follows a continuous uniform distribution where $E(X)$ = 4 and $Var(X)$ = 3. What is the value of the upper endpoint?

(a) 7 (b) 5 (c) 22 (d) –1 (e) 1

Solution: $E(X) = \frac{b+a}{2} = 4 \to a = 8-b$

$$Var(X) = \frac{(b-a)^2}{12} = \frac{(b-(8-b))^2}{12} = \frac{(2b-8)^2}{12} = \frac{4(b-4)^2}{12} = 3 \to (b-4)^2 = 9 \to b-4 = \pm 3$$

Thus, b = 1, 7 but since the average, $E(X)$, is 4, it must be 7.

7. Two engineers (married to each other) plan to have children until they get a girl, but they agree that they will not have more than three children, even if all are boys. Assume getting a boy or girl is equally likely. What is the expected number of boys?

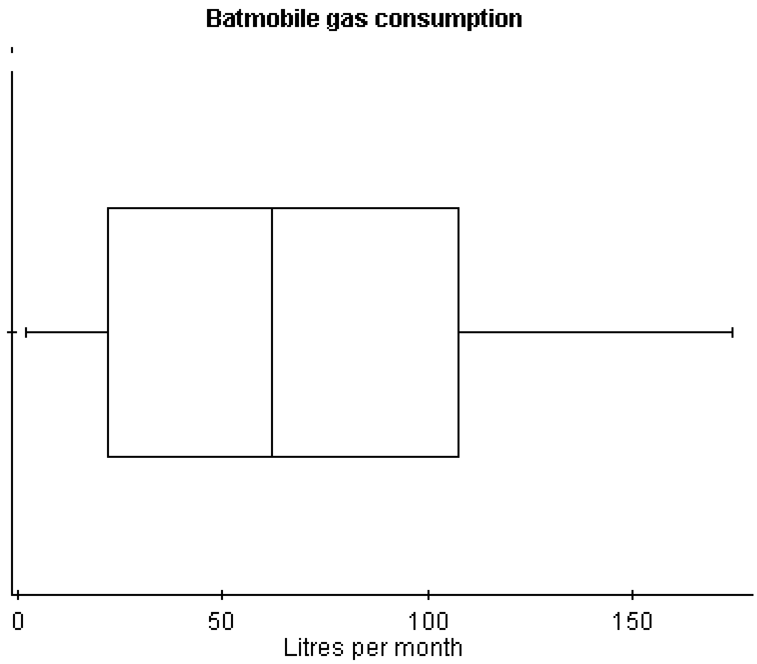
(a) 1.750 **(b) 0.875** (c) 1.375 (d) 1.500 (e) 1.875

Solution: There are 4 possible scenarios for children: 1 girl; 1 boy, 1 girl; 2 boys, 1 girl, and 3 boys.

x	0	1	2	3
$P(X = x)$	1/2	1/4	1/8	1/8

Thus, for X = # of boys, $E(X) = \sum x_i P(X = x_i) = 0(1/2) + 1(1/4) + 2(1/8) + 3(1/8) = 7/8 = 0.875$

8. Determine the most likely shape and approximate third quartile of the following boxplot.



(a) Left-skewed and 60.
(b) Right-skewed and 25.
(c) Left-skewed and 110.
(d) Left-skewed and 25.
(e) Right-skewed and 110.

Solution: Whisker on the right is longer, so right-skewed and third quartile is rightmost edge of the box, so about 110.

9. In comparing probability distributions, which of the following statements is TRUE?

(a) The geometric and Poisson distributions have the lack of memory property.
(b) The uniform and normal distributions have the same distribution shapes.
(c) The Poisson distribution is the limiting form of the negative binomial distribution.
(d) The exponential and Poisson distributions are interchangeable because they have the exact same random variable.
(e) The geometric and negative binomial distributions both have a random number of trials.

Solution: Review properties of distributions.

10. A machine that fills cardboard boxes with cereal has a fill weight whose mean is 12.02 oz, with a standard deviation of 0.03 oz. How many boxes must be randomly sampled from the output of the machine such that the standard deviation of the total weight of the cereal in the sample to be 0.12?

(a) 4 (b) 8 (c) 12 **(d) 16** (e) 20

Solution: Let X_i be the weight of the i^{th} box and $Y = X_1 + X_2 + \ldots + X_n$ be the total weight of n boxes.
$Var(Y) = Var(X_1 + X_2 + \ldots + X_n) = n{*}Var(X_i) \rightarrow (0.12)^2 = n(0.03)^2 \rightarrow n = 16$

11. A speech pathologist creates a universal translator to relay communications between different languages. Let X be the number of languages the universal translator translates on a random day. The following is the probability distribution of X. What is the expected value of X?

x	1	2	3	4	5	6
$P(X = x)$	0.1982	?	0.0428	0.0822	0.0221	0.1987

(a) 0.478 **(b) 2.870** (c) 3.500 (d) 1.958 (e) 2.672

Solution: Since $\sum P(X = x) = 1$,

$P(X = 2) = 1 - (0.1982 + 0.0428 + 0.0822 + 0.0221 + 0.1987) = 0.4560$

$$E(X) = \sum x_i P(X = x_i) = 1(0.1982) + 2(0.4560) + 3(0.0428) + 4(0.822) + 5(0.0221) + 6(0.1987) = 2.870$$

12. A manufacturer of an electronic component is interested in determining the lifetime of a certain type of battery. A sample, in hours of life, is as follows:

123, 116, 122, 110, 175, 166, 125, 85, 118, 117

Which of the following given statements is CORRECT?

(a) The first quartile is 135 and the third quartile is 116.
(b) The interquartile range is 37.
(c) The first quartile is 120 and the third quartile is 135.
(d) The interquartile range is 9.
(e) The first quartile is 122 and the third quartile is 85.

Solution: Putting the values in order: 85, 110, 116, 117, 118, 122, 123, 125, 166, 175
Thus, Q_1 is the 3rd value (116), Q_3 is the 8th value (125), and IQR = $Q_3 - Q_1$ = 125 – 116 = 9.

13. Refer to the previous question. Which of the following given statements is CORRECT?

(a) 85 is an outlier.
(b) 175 is an outlier.
(c) 85 and 175 are outliers.
(d) 85, 166, and 175 are outliers.
(e) There are no outliers in the above data set.

Solution: Lower inner fence = Q_1 – 1.5*IQR = 116 – 1.5*9 = 102.5
Upper inner fence = Q_3 + 1.5*IQR = 125 + 1.5*9 = 138.5
Thus, 85, 166, and 175 are all outliers.

14. Police report that 78% of drivers stopped on suspicion of drunk driving are given a breath test, 36% a blood test, and 22% both tests. What is the probability that a randomly selected suspect is given exactly one of the tests (a blood test or a breath test, but not both)?

(a) 0.64 **(b) 0.70** (c) 0.74 (d) 0.79 (e) 0.82

Solution: Let H be a breath test and B be a blood test.
$P(H) = 0.78$, $P(B) = 0.36$, $P(H \cap B) = 0.22$
$P(\text{exactly one test}) = P(H \cap B') + P(H' \cap B) = (0.78 - 0.22) + (0.36 - 0.22) = 0.70$

15. One out of every 2000 individuals in a population gets bitten by Luis Suarez of the Uruguay World Cup soccer team. A random sample of 1000 individuals is studied. What is the probability that exactly one of the individuals in the sample is bitten by Luis Suarez?

(a) 0.368 (b) 0.607 (c) 0.164 **(d) 0.303** (e) 0.910

Solution: Let *X* be the number of people bitten by Luis Suarez. Then, *X* follows a binomial distribution with $n = 1000$ and $p = 1/2000 = 0.0005$.

$$P(X = 1) = \binom{1000}{1}(0.0005)^1(1-0.0005)^{999} = 1000(0.0005)^1(0.9995)^{999} = 0.303$$

16. Suppose the number of cracks per concrete specimen for a particular type of cement mix has approximately a Poisson distribution. Assume that the average number of cracks per specimen is 2.5. What is the probability that a randomly selected concrete specimen has two or more cracks?

(a) 0.127 (b) 0.287 (c) 0.543 **(d) 0.713** (e) 0.876

Solution: Let *X* be the number of cracks per specimen. Then, *X* follows a Poisson distribution with a mean of 2.5.

$$P(X \geq 2) = 1 - [P(X = 0) + P(X = 1)] = 1 - \left[\frac{e^{-2.5}2.5^0}{0!} + \frac{e^{-2.5}2.5^1}{1!}\right] = 1 - [0.2052 + 0.2873] = 0.713$$

17. Suppose 10% of the engines manufactured on a certain assembly line are defective. If engines are randomly selected one at the time and tested, what is the probability that the first defective engine is found on the fourth trial?

(a) 0.067 **(b) 0.073** (c) 0.083 (d) 0.091 (e) 0.117

Solution: Let *X* be the number of tested engines until the first defective is found. Then, *X* follows a geometric distribution with $p = 0.10$.
$P(X = 4) = (1 - 0.10)^3(0.10) = 0.073$

18. Emmet has constructed 28 double-decker couches. The probability that any couch is defective is 0.0701 and the couches are independent. What is the probability that at least one couch is defective?

(a) 0.407 **(b) 0.869** (c) 0.276 (d) 0.131 (e) 0.724

Solution: Let *X* be the number of defective double-decker couches. Then, *X* follows a binomial distribution with $n = 28$ and $p = 0.0701$.

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{28}{0}(0.0701)^0(1-0.0701)^{28} = 1 - (0.9299)^{28} = 1 - 0.276 = 0.724$$

19. Small aircraft arrive at a certain airport according to a Poisson distribution with a rate of 8 aircraft per hour. What is the probability that at most 2 small aircraft arrive at this airport within 30 minutes?

(a) 0.220 (b) 0.762 (c) 0.014 **(d) 0.238** (e) 0.092

Solution: Let *X* be the number of small aircraft arriving at the airport in 30 minutes. Since small aircraft arrive at the airport at a rate of 8 per *hour*, *X* follows a Poisson distribution with a mean of 4.

$$P(X \le 2) = \frac{e^{-4} \times 4^0}{0!} + \frac{e^{-4} \times 4^1}{1!} + \frac{e^{-4} \times 4^2}{2!} = 0.01832 + 0.07326 + 0.14653 = 0.238$$

20. Refer to the previous question. Suppose you are waiting at that airport. What is the probability that a small aircraft will arrive between 15 minutes and 30 minutes from the time you begin observing?

(a) 0.846 **(b) 0.117** (c) 0.233 (d) 0.883 (e) 0.491

Solution: Let *X* be the time until the next small aircraft arrives. Since small aircraft arrive at the airport at a rate of 8 per hour, *X* follows an exponential distribution where $\lambda = 8$ while 15 minutes and 30 minutes are 0.25 and 0.5 hours, respectively.

$$P(0.25 \le X \le 0.5) = F(0.5) - F(0.25) = (1 - e^{-8(0.5)}) - (1 - e^{-8(0.25)}) = 0.98168 - 0.86466 = 0.1170$$

21. Refer to the previous two questions. What are the expected value and variance, respectively, for the time (in hours) that it takes for a small aircraft to arrive at this airport?

(a) 0.125, 0.0156 (b) 0.125, 0.1250 (c) 0.250, 0.0625 (d) 8.000, 2.8284 (e) 4.00, 2.00

Solution: Let *X* be the time until the next small aircraft arrives. Since small aircraft arrive at the airport at a rate of 8 per hour, *X* follows an exponential distribution where $\lambda = 8$.

$$E(X) = \frac{1}{\lambda} = \frac{1}{8} = 0.125 \qquad Var(X) = \frac{1}{\lambda^2} = \frac{1}{8^2} = 0.0156$$

22. Adult human body weights in a certain population are normally distributed with a mean of 75 kg and a standard deviation of 12 kg. What is the percentage of this population that has body weights between 50 kg and 70 kg?

(a) 33.72% (b) 66.28% **(c) 31.84%** (d) 75.28% (e) 1.88%

Solution: *X* follows a normal distribution with $\mu = 75$ and $\sigma = 12$.

$$P(50 \le X \le 70) = P(50 \le X \le 70) = P\left(\frac{50-75}{12} \le Z \le \frac{70-75}{12}\right) = P(-2.08 \le Z \le -0.42)$$

$$= P(Z \le -0.42) - P(Z \le -2.08) = 0.3372 - 0.0188 = 0.3184$$

23. Refer to the previous question. What is the body weight (in kg) at the 95th percentile?

(a) 55.26 **(b) 94.74** (c) 90.68 (d) 95.44 (e) 64.29

Solution: Let x_0 be the weight at the 95th percentile. Then, $P(X \le x_0) = 0.95$. Since *X* is normal,

$$P(X \le x_0) = P\left(Z = \frac{X-\mu}{\sigma} \le \frac{x_0 - 75}{12} = z_0\right) = 0.95$$, where $P(Z \le z_0) = 0.95$ gives $z_0 = 1.645$.

$$\frac{x_0 - 75}{12} = 1.645 \rightarrow x_0 = 75 + 1.645(12) = 94.74$$

24. Refer to the previous two questions. Suppose that we take a random sample of 25 adults from the population described, what is the probability that the average weight of the 25 adults is greater than 80 kg?

(a) 0.9812 (b) 0.3372 (c) 2.083 **(d) 0.0188** (e) 0.6628

Solution: $P(\bar{X} > 80) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > \frac{80-75}{12/\sqrt{25}}\right) = P(Z > 2.08) = 1 - P(Z < 2.08) = 1 - 0.9812 = 0.0188$

25. It is reported that 14.2% of a certain type of circuit board are defective. In a random sample of 300 circuit boards (sampling with replacement), what is the probability that at least 16% will be defective?

(a) 0.0202 (b) 0.1895 (c) 0.8105 (d) 0.7910 **(e) 0.2090**

Solution: $P(\hat{p} \geq 0.16) = P(0.16 \leq \hat{p}) \approx P\left(\frac{0.16 - {}^{0.5}\!/_{300} - 0.142}{\sqrt{\frac{0.142(1-0.142)}{300}}} \leq \frac{\hat{p} - \frac{0.5}{n} - p}{\sqrt{\frac{p(1-p)}{n}}}\right) = P(0.81 \leq Z)$

$= P(Z \geq 0.81) = 1 - P(Z < 0.81) = 1 - 0.7910 = 0.2090$

26. Thickness of gaskets produced by a certain machine is normally distributed with mean μ = 1.8 mm and standard deviation σ = 0.03 mm. What are the mean and variance of a total of 5 such gaskets?

(a) 1.8, 0.0009 (b) 1.8, 0.15 (c) 9.0, 0.03 **(d) 9.0, 0.0045** (e) 9.0, 0.15

Solution: $E(X_1 + X_2 + X_3 + X_4 + X_5) = E(X_1) + E(X_2) + E(X_3) + E(X_4) + E(X_5) = 5*1.8 = 9.0$

$Var(X_1 + X_2 + X_3 + X_4 + X_5) = Var(X_1) + Var(X_2) + Var(X_3) + Var(X_4) + Var(X_5) = 5*0.03^2 = 0.0045$

27. Suppose that a certain model of DVD players have an 80% chance of operating four years without defect. If a company purchases 900 such DVD players (selecting them randomly and independently), what are the mean and standard deviation of the number of DVD players that would be expected to operate four years without defect?

(a) 180, 12 (b) 720, 144 (c) 720, 461 **(d) 720, 12** (e) 180, 144

Solution: X follows a binomial distribution with n = 900 and p = 0.8. Thus,

$E(X) = np = (900)(0.8) = 720$ Standard deviation $= \sqrt{np(1-p)} = \sqrt{(900)(0.8)(0.2)} = 12.00$

28. Refer to the previous question. What is the probability that at least 730 of these DVD players will operate four years without defect?

(a) 0.8111 (b) 0.2083 **(c) 0.2148** (d) 0.7852 (e) 0.7917

Solution: Since 730 is quite large, the normal approximation to binomial should be used.

$$P(X \geq 730) \approx P\left(\frac{Y - np}{\sqrt{np(1-p)}} \geq \frac{730 - 0.5 - (900)(0.8)}{\sqrt{(900)(0.8)(1-0.8)}}\right) = P(Z \geq 0.79)$$

$= 1 - P(Z \leq 0.79) = 1 - 0.7852 = 0.2148$

29. Suppose a continuous random variable X has the following probability density function:

$$f(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & \textit{elsewhere} \end{cases}$$

Which of the following probabilities approximately equals 0.1353?

(a) $P(1 \leq X \leq 4)$ (b) $P(2 < X \leq 3)$ (c) $P(2 \leq X < 4)$ (d) $P(X \leq 0.5)$ **(e) $P(X > 2)$**

Solution: Determine $F(x) = 1 - e^{-x}$. Use this for every interval to speed up calculations.

30. Suppose a continuous random variable X has the following probability density function:

$$f(x) = \begin{cases} 5x^4 & 0 \leq x \leq 1 \\ 0, & \textit{otherwise} \end{cases}$$

What is the median of X?

(a) 0.6345 (b) 0.6989 (c) 0.7215 **(d) 0.8706** (e) 0.8235

Solution: $F(x) = 0.5$

$$\int_0^x 5u^4 du = 5\left[\frac{1}{5}u^5\right]_0^x = x^5 - 0 = x^5 \rightarrow x^5 = 0.5 \rightarrow x = 0.5^{1/5} = 0.8706$$