New York City College of Technology MAT1372 Practice Midterm sol’ns, Halleck, Sp 2018

* Closed book and notes (except for a single, handwritten front & back formula sheet worth up to 5 bonus points for its usefulness and thoroughness). No use of computers, laptops, tablets or other handhelds.
* You may use scientific or graphing calculator.
* A selection of problems similar to those below will be chosen (total below is 160; points for selected problems will sum to 100).
1. (10 pts) The formula **=B$2\*$A2** is located in cell **B1**.
	1. What does cell B1 evaluate to? **3\*3=9**
	2. If this was copied and pasted into cell D3, what would the resulting formula be? **=D$2\*$A4**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **A** | **B** | **C** | **D** |
| **1** |  2 | **=B$2\*$A2** | 4 |  5 |
| **2** | 3 | 3 | 8 | 6 |
| **3** | 5 | 4 | 3 | ????????? |
| **4** | 4 | 3 | 4 | 9 |

1. (10 pts) A histogram uses bins, each of which represents a(n) \_interval\_\_ of **possible** data **values** **(outcomes)**. If the range of our

data is 21-48 and we want 7 bins, then the width of each bin is (48-21+1)/7=4 (show your calculation.)

For 10 bins, the width is (48-21+1)/10=2.8~3. For 5 bins, the width is (48-21+1)/5=5.6~6. If our bin width is 5,

how many bins will we have? (48-21+1)/5=5.6~6. For this last choice, the first and last #’s

for the right endpoints in Excel are \_\_24\_\_ and \_\_\_49\_\_\_ or if using a stem and leaf plot, the

stems would be \_2, 2, 3, 3, 4, 4.

1. **(**15 pts) The following are weights of chickens that have been put out on display for sale in a market.
	1. Make a stem and leaf plot using every .05 for the stem:

(3.7, 3.8, 3.8, 3.9, 3.9, 4.0, 4.0, 4.1, 4.1).

* 1. What can you say about the distribution? Is it evenly distributed?

Bell-shaped? Symmetric? **Roughly bell-shaped. Roughly symmetric.**

* 1. Find the median and mode. The mean is 3.97. **mode 3.96**

 **median (3.96+3.96)/2=3.96**

* 1. Does this indicate right or left skewing? Why?

**Slightly to the Right: Mean> median**

3.75 3.88 3.94 3.96 4.02 4.09

3.82 3.92 3.94 3.98 4.03 4.10

3.84 3.93 3.96 3.99 4.06 4.12

3.86 3.93 3.96 4.02 4.06 4.17

1. (15 pts) Consider an experiment that consists of withdrawing a ball from the box, replacing it, and withdrawing a second ball. Draw a tree diagram. Be sure to include labels and probabilities. Use the tree diagram to make a table with outcomes and probabilities. **Express all probabilities as fractions in lowest terms.** There are 2 red, 2 blue and 1 green ball in the box.

****

* 1. What is the sample space of this experiment?

**1st row of table above**

* 1. As a set, what is the event A: the first ball drawn is red?

What is its probability?

 **A={rr, rb, rg}, P(A)=10/25=2/5**

* 1. As a set, what is the event B: the same color ball is drawn twice?

What is its probability?

 **B={rr, bb, gg}, P(B) = 9/25**

* 1. Find A or B and its probability. Draw a Venn diagram with outcomes.

Shade appropriately. A or B ={**rr, rb, rg, bb, gg} P(A or B)=P(A)+ P(B) – P(A and B) = 2/5+9/25-4/25=(10+9-4)/25=15/25=3/5**



* 1. Find A and (not B) and its probability. Draw a Venn diagram with outcomes. Shade appropriately. A and (not B) = {rb, rg}, P(A and (not B))=4/25+2/25=6/25
1. (15 pts) Repeat Prob. 4 when the second ball is drawn without replacement of the first ball.

Consider an experiment that consists of withdrawing a ball from the box, NOT replacing it, and withdrawing a second ball. Draw a tree diagram. Be sure to include labels and probabilities. Use the tree diagram to make a table with outcomes and probabilities. **Express all probabilities as fractions in lowest terms.** There are 2 red, 2 blue and 1 green ball in the box.

****

* 1. What is the sample space of this experiment?

**1st row of table above**

* 1. As a set, what is the event A: the first ball drawn is red?

What is its probability?

 **A={rr, rb, rg}, P(A)=4/10=2/5**

* 1. As a set, what is the event B: the same color ball is drawn twice?

What is its probability?

 **B={rr, bb}, P(B) = 2/10 = 1/5**

* 1. Find A or B and its probability. Draw a Venn diagram with outcomes.

Shade appropriately. **A or B ={rr, rb, rg, bb}**

**P(A or B)=P(A)+ P(B) – P(A and B) = 2/5+1/5-1/10=(4+2-1)/10=5/10=1/2**

****

* 1. Find A and (not B) and its probability. Draw a Venn diagram with outcomes. Shade appropriately. **A and (not B) = {rb, rg}, P(A and (not B))=1/5+1/10=3/10**
1. (10 pts) Of the families in a certain community, all have a parent or grandparent, 93 percent have at least one parent, and 18 percent have at least one grandparent.
	1. If a family is chosen at random, what is the probability it has both a parent and a grandparent? Draw a Venn diagram with probabilities. Shade appropriately.

**P(Pa and G) = P(Pa) + P(G) – P(G or Pa) = .93+.18 – 1 = 11%**



* 1. If the community consists of 1000 families, how many of them have only a grandparent? Draw a Venn diagram with frequencies. Shade appropriately.

**|G and PaC| = |G| – |Pa and G| = 180 – 110 = 70**

1. (10 pts) It is estimated that 20 percent of all adolescents in the United States are obese, 1 percent of all adolescents suffer from diabetes and 0.8 percent of all adolescents both are obese and suffer from diabetes. Determine the conditional probability that a randomly chosen adolescent
	1. suffers from diabetes given that he or she is obese;

**P(O and DC) = P(O) – P(O and D) = 20% - 0.8%**

**P(D and OC) = P(D) – P(O and D) = 1% - 0.8%**

**P(D|O)=P(D and O)/P(O)=.8%/20%=8/200=1/25**

* 1. is obese given that she or he suffers from diabetes.

**P(O|D)=P(D and O)/P(D)=.8%/1%=8/10=4/5**

1. (10 pts) A committee is formed of 6 randomly selected adult residents. What is the probability that it will consist of
	1. 4 women and 2 men? **6 nCr 2/26 = 15/26**
	2. 5 women and 1 men? **6 nCr 1/26= 6/26**
	3. 6 women (and 0 men)? **1/ 26**
	4. Use a, b, c to find at least 4 women? **22/ 26 = 11/25 = 11/32**
	5. Use d and complement to find at most 3 women (which is equivalent to at least 3 men).

**1-11/32= 21/32**

8alt. (10 pts) A committee is formed of 6 randomly selected adult residents taken from a pool of 15 men and 5 women. What is the probability that it will consist of

* 1. 4 women and 2 men? **5 nCr 4\*15 nCr 2/20 nCr 6 = 5\*105/5005 = 105/1001**
	2. 5 women and 1 men? **5 nCr 5\*15 nCr 1/20 nCr 6 = 15/5005=3/1001**
	3. 6 women (and 0 men)? **0**
	4. Use a, b, c to find at least 4 women? **108/1001**
	5. Use d and complement to find at most 3 women (which is equivalent to at least 3 men).

**1-108/1001= 893/1001**

1. (10 pts) An ordinary license plate from New York has 3 letters followed by 4 digits, e.g. KAK-3310
	1. How many different ordinary license plates can be made? **263\*104**
	2. If the letters must be distinct and the first digit cannot be a zero, how many plates can be made?

**26nPr3\*9\*103=26\*25\*24\*9\*103**

* 1. If exactly 2 letters must be the same and the last digit must be odd, how many plates can be made? **26\*25\*3nCr2\*103\*5=26\*25\*3\*103\*5**
1. **pick the letter which is repeated**
2. **pick the remaining letter**
3. **place the letters which are repeated**
4. **pick the first 3 numbers**
5. **pick the last number**
6. (10 pts) A random variable has the following probability distribution:

 X -2 -1 1 2

 P(X) 0.3 0.2 0.3 ?

Find and **interpret**

1. probability when X=2: **P(2)=1-(0.3+0.2+0.3) = 1-0.8=0.2**
2. mean **= -2\*.3+-1\*.2+1\*.3+2\*.2=-.6-.2+.3+.4= -.1 On average, the outcome will be -.1**
3. standard deviation

 X^2 4 1 1 4

 P(X) 0.3 0.2 0.3 .2

**=√(4\*.3+1\*.2+1\*.3+4\*.2)-(-.1)^2) =√(2.5-.01) =√2.49=1.58**

**On average an outcome will be about 1 ½ units from the mean (~0), which makes perfect sense if you look at the distribution which is close to uniform.**

1. (20 pts) A shipment of parts contains 40 items of which 5 are defective. Ten of these items are randomly chosen and inspected. Let X denote number that are defective. Let be the random variable that is 1 if the ith part selected is defective and 0 otherwise. Note that .
	1. Find expressions for probability distribution of *X* (leave as a product and/or ratio of binomial coefficients):

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| x | 0 | 1 | 2 | 3 | 4 | 5 |
| P(x) |  |  |  |  |  |  |

* 1. Expand the probability for your first outcome so that it is a product and/or ratio of integers.

$$\frac{\frac{35\*34\*33\*32\*31\*30\*29\*28\*27\*26}{10\*9\*8\*7\*6\*5\*4\*3\*2\*1}}{\frac{40\*39\*38\*37\*36\*35\*34\*33\*32\*31}{10\*9\*8\*7\*6\*5\*4\*3\*2\*1}}=\frac{35\*34\*33\*32\*31\*30\*29\*28\*27\*26}{40\*39\*38\*37\*36\*35\*34\*33\*32\*31}$$

* 1. Write the Excel command to find probability for your first outcome.

**=combin(35,10)/combin(40,10) (Calculator: 35 nCr 10/40 nCr 10 ≈ 22%)**

* 1. Find the distribution and expectation for .

|  |  |  |
| --- | --- | --- |
| **Xi** | **0** | **1** |
| **P(Xi)** | **7/8** | **1/8** |

**E() = 1/8**

* 1. Use part d) and the fact that expectation distributes over a sum regardless of independence **to find E[X]= 1/8+ … +1/8=10/8=1.25**
	2. Expectation is a measure of a distribution’s center. In light of your calculation from e), discuss aspects that the graph of X might have. Will it be symmetric? skewed? If so, in which direction?

 **Distribution will be skewed to right (longer tail). Likely outcomes are concentrated on left.**

1. (25 pts) The table below displays for each of 10 students in a statistics course the number of days he or she was absent (X) and his or her final grade (Y).



1. Use graph paper to draw the scatter plot.

Use as window x:[0,10] y:[60,100] x-scale of 1 and y-scale of 5.

(4.5, 81)

(8.5, 70)

1. Put in what you think is the trend line. Roughly half the data points should be above and below the line. Find and label grid points (not data points) close to the ends and use them to calculate the slope.

**m = (70−­81)/(8.5−4.5) = −11/4 = −2.75**

1. Find and interpret the coefficient of correlation.

r=sum dev prods/√ (sum of xdev2\* sum of ydev2) = -231/√ (82.5\*726) = -.944

**Since r is negative, the correlation is negative (trendline goes down from left to right).**

1. Find and interpret the square of the coefficient of correlation.

**Since r2 is .891, at most 89% of student’s final exam score may be attributed to attendance (in actuality, it is much less; other factors are preparation, aptitude and work outside of class).**

1. Find and interpret slope of trend line:.**b = sum dev prods/sum of xdev2 = -231/82.5 ≈ -2.8**

**For every absence, the final grade decreases on average by 2.8 points.**

1. Find y-intercept of trend line. $a=\overbar{y}-b\overbar{x}=81-\left(-2.8\right)4.5=93.6$

**A person with 0 absences will have on average a score of 93.6.**