**Student name:**

MATH 120: Elementary Statistics
Section 1

Test 2
March 28, 2019

**Instructions:**

- This is a regular "closed-book" test, and is to be taken without the use of notes, books, or other reference materials. Collaboration or group work is not permitted.
- Cell-phone usage, in any form, is prohibited for the entire duration of the test. This also applies to any restroom breaks taken during the test.
- Important: For full credit, all solutions must show relevant calculation steps, reasons, sketch, and/or justifcation, as needed.
- This test adds up to 50 points. It contains questions numbered 1-7.

**1.** [6 points] Given the probabilities P(A) = 0.6, P(A and B) = 0.2, P(A or B) = 0.7, find:

(a) P(B)

(b) P(A | B)

(a) we know that $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$
using the given values to plug into this, we get

$$0.7 = 0.6 + P(B) - 0.2$$

$$\therefore \boxed{P(B) = 0.7 - 0.6 + 0.2 = 0.3}$$

(b) $P(A|B) = \dfrac{P(A \text{ and } B)}{P(B)}$. $P(A \text{ and } B)$ is given, and we know $P(B)$ from part (a)

$$= \frac{0.2}{0.3} = \frac{2}{3}$$

$$\boxed{P(A|B) = \frac{2}{3}}$$

Grade: 1.5pt = correctly write or apply formula for P(A or B)
1.5pt = correctly compute P(B)
1.5pt = Know/show/apply correct P(A|B) formula
1.5pt = correctly compute P(A|B)



Other approaches:
2 pt = Venn diagram showing P(A) − P(A and B) = 0.4
1 pt = use it to get P(B)
3 pt = use diagram to correctly figure out P(A|B)

**2.** [6 points] The radioactive gas radon, found in some homes, poses a health risk to residents. To assess the level of contamination in their area, a county health department wants to test a few homes. If the risk seems high, they will publicize the results to emphasize the need for home testing. Officials plan to use the local property tax lists to randomly select 25 homes in the county for testing. Identify the following as precisely as possible:

* The population:

1.5 pt → All homes in that county or region

* The sample:

1.5 pt → The 25 homes randomly selected from the property tax lists.

* The parameter(s) of interest:

1.5 pt → The percent of all homes in the county with unsafe levels of radon concentration.

OR The mean level of radon contamination in homes in the area.

* The statistic(s):

1.5 pt → The percent of homes in the sample of 25 with unsafe levels of radon concentration.

**3.** [6 points] Each of the following questions describes a method for choosing a sample. Write the letter corresponding to the sampling strategy in the blank next to each.

B i. An insurance research organization wants to study driving patterns among different age groups. They randomly select 200 drivers under the age of 25 years, 400 drivers between 25 and 60, and 150 drivers over the age of 60.

C ii. A pollster uses a computer to randomly select 5000 names from the list of registered voters in an area.

D iii. An apple orchard is divided into 18 regions. Management randomly selects 2 regions, counts all the apples produced in those, and uses that to estimate the number of apples produced in the whole orchard.

E iv. A government auditor is tasked with choosing and auditing 36 companies in a state. She lists all the company names in alphabetical order, randomly selects 3 letters of the alphabet, and randomly picks 12 companies with names from each of those alphabets.

A. convenience
B. stratified random
C. simple random
D. cluster
E. multistage
F. systematic
G. voluntary response

Grade: 1.5 points for each correct answer.

Grade: (a) = 1 point (b) = 4 points (c) = 3 points. For (a): 0.5 + 0.5 = answer + reason
For (b): 1 pt = attempt complement. 1 pt = set it up correctly. 2 pt = correct computation.
For (c): 1 pt = correctly compute P(Both Green). 1 pt = correct P(Both Blue). 1 pt = add both

**4.** [8 points] Box A and Box B are filled with green and blue marbles as shown. Each box is shaken thoroughly.

| BOX A | BOX B |
|---|---|
| 6 Green<br>4 Blue | 60 Green<br>40 Blue |

(a) We are allowed to randomly pick one marble from one of the two boxes. We want to get a green marble. Which box will give us the best chance? Why?

(b) Suppose we are allowed to randomly pick three marbles from Box B. Find the probability of getting at least one green.

(c) If we pick 3 marbles from Box B, what is the probability of getting the same color?

(a) Box A: $P(\text{Green}) = \frac{6}{6+4} = 0.6$

Box B: $P(\text{Green}) = \frac{60}{60+40} = 0.6$

Because the probabilities are equal, both boxes give the same chance of randomly picking a Green marble.

(b) since there are 100 marbles total, it will give very similar answers with, or without replacement.

<u>With replacement</u>

$P(\text{at least one Green})$
$= 1 - P(\text{none Green})$
$= 1 - \left(\frac{40}{100}\right)^3 = 1 - (0.4)^3$

$\boxed{P(\text{at least one Green}) = 0.936}$

<u>Without replacement</u>

$P(\text{at least one Green})$
$= 1 - P(\text{none Green})$
$= 1 - \left(\frac{40}{100}\right)\left(\frac{39}{99}\right)\left(\frac{38}{98}\right)$

$\boxed{P(\text{at least one Green}) = 0.9389}$

we have assumed independence because the marbles are picked randomly.

(c) Again, with or without replacement will give similar results.

<u>With replacement</u>

$P(\text{all Green}) = P(\text{Green}) \times P(\text{Green}) \times P(\text{Green})$
$= (0.6)^3 = 0.216$

$P(\text{all Blue}) = P(\text{Blue}) \times P(\text{Blue}) \times P(\text{Blue})$
$= (0.4)^3 = 0.064$

$P(\text{all same color}) = 0.216 + 0.064$
$\boxed{= 0.28}$

<u>Without replacement</u>

$P(\text{all Green}) = P(1^{st}\,G) \times P(2^{nd}\,G) \times P(3^{rd}\,G)$
$= \left(\frac{60}{100}\right)\left(\frac{59}{99}\right)\left(\frac{58}{98}\right) = 0.2116$

$P(\text{all Blue}) = P(1^{st}\,B) \times P(2^{nd}\,B) \times P(3^{rd}\,B)$
$= \left(\frac{40}{100}\right)\left(\frac{39}{99}\right)\left(\frac{38}{98}\right) = 0.0611$

$P(\text{all same Color}) = 0.2116 + 0.0611$
$\boxed{= 0.2727}$

P(all same) is just the sum of each, because they are disjoint.

**5.** [8 points] To play a game, you must pay \$10 per play. There is a 10% chance you will win \$20, and a 40% chance you will win \$15. The rest of the time you win nothing.

(a) Create a probability model for your net winnings per play.
(b) Find the expected amount you'll win.
(c) Find the standard deviation of the amount.

If you are unsure whether you got the right answer to (a), assume the probability model is

| X (in \$) | -2 | 5 | 8 | 12 |
|---|---|---|---|---|
| P(X) | 0.15 | 0.2 | 0.3 | 0.35 |

Note, this is NOT the correct answer to part (a); however, feel free to use it to solve parts (b)-(c) of this question.

(a) Let X denote my net winnings per play.
It costs \$10 to play. The amount won can be \$20, \$15, \$0.
Thus, the possible values of X are \$10, \$5, −\$10.
The probability model is shown in the table below

| X (\$) | 10 | 5 | −10 |
|---|---|---|---|
| P(X) | 0.1 | 0.4 | 0.5 |

(b) Expected Value, $E(X) = \sum x \cdot P(x) = 10 \times (0.1) + 5 \times (0.4) - 10 \times (0.5)$
$\boxed{= -\$2}$
The expected amount I'll win is $-\$2$ = a loss of \$2.

(c) standard deviation of the amount: $SD(X) = \sqrt{\sum (x-\bar{x})^2 P(X)}$
$= \left[(10+2)^2(0.1) + (5+2)^2(0.4) + (-10+2)^2(0.5)\right]^{1/2} \boxed{= \$8.124}$

---

If alt. prob. model: (b) $E(X) = -2 \times 0.15 + 5 \times 0.2 + 8 \times 0.3 + 12 \times 0.35$
$\therefore E(X) = \$7.3$
(c) $SD(X) = \sqrt{(-2-7.3)^2(0.15) + (5-7.3)^2(0.2) + (8-7.3)^2(0.3) + (12-7.3)^2(0.35)}$
$\therefore SD(X) = \$4.681$

---

Grade: (a) = 3 points, (b) = 3 points. (c) = 2 points
For (a): 1.5pt = correct values of X; 1.5pt = correct P(X)
For (b): 2pt = correct formula & calculation steps
1pt = correct answer
For (c): 1pt = correct calculation step. 1pt = correct answer

**6.** [8 points] A company's records show that on any given day about 1% of their day shift employees and 2% of their night shift employees will be absent. Sixty percent of the employees work the day shift.

(a) Find the probability that a randomly selected employee will be absent on any given day.

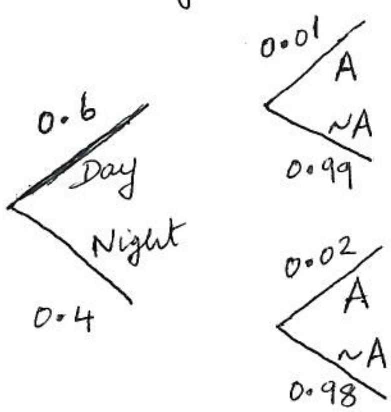
(b) Find the probability that an employee who is absent, works the day shift.

(c) Find the probability that an employee who is not absent, works the night shift.

(d) Are employees' absence and the shift they work disjoint, independent, neither, or both?

2 points

A tree diagram would work well for organizing this info.




P(Day and Absent) = $0.6 \times 0.01 = 0.006$

P(Night and Absent) = $0.4 \times 0.02 = 0.008$

1 point

(a) P(Absent) = P(Day and Absent) + P(Night and Absent)

$= 0.006 + 0.008$ [as shown in the tree diagram]

$\Rightarrow \boxed{P(\text{Absent}) = 0.014}$

1.5 point

(b) This question is asking for P(Day | Absent)

$P(\text{Day} \mid \text{Absent}) = \dfrac{P(\text{Day and Absent})}{P(\text{Absent})} = \dfrac{0.006}{0.014} = 0.4286$

$\boxed{P(\text{Day} \mid \text{Absent}) = 0.4286}$

1.5 point

(c) This question is asking for $P(\text{Night} \mid \sim\text{Absent}) = \dfrac{P(\text{Night and } \sim\text{Absent})}{P(\sim\text{Absent})}$

$P(\text{Night} \mid \sim\text{Absent}) = \dfrac{0.4 \times 0.98}{(1 - 0.014)}$ ← From tree diagram / ← From part (a)

$\boxed{P(\text{Night} \mid \sim\text{Absent}) = 0.3976}$

2 points

(d) Employees' absence and shift are certainly not disjoint, since it is possible to be absent and work either day or night shift.

i.e, $P(\text{Absent and Day}) \neq 0$, $P(\text{Absent and Night}) \neq 0$

They are also not independent, since $P(\text{Absent}) \neq P(\text{Absent} \mid \text{Day})$

$0.014 \neq 0.01$

Also, $P(\text{Day}) \neq P(\text{Day} \mid \text{Absent})$

$0.6 \neq 0.4286$

Thus, they are neither disjoint, nor independent.

**7.** [8 points] A nutrition lab wants to study the effect of storage time on the amount of Vitamin C that is present in freeze dried fruit. They conduct an experiment in which 12 packages of freeze dried fruit are randomly divided into groups of 4, and stored for the following periods (in months): 0, 6, 12. At the end of each storage period, the amount of Vitamin C present in the corresponding packages is estimated (in milligrams of Vitamin C per 1000 milligrams of fruit). Identify each of the following as clearly & precisely as possible in this study:

* Experimental units:

12 packages of freeze dried fruit

1.5 pt →

* Factor(s) and levels:

There is one factor = Storage time
It has 3 levels = 0, 6, 12 months

2 pt →

* How many treatments? Name them.

There are 3 treatments → they are the same as the factor levels: 0, 6, 12 months of storage

1.5 pt →

* Response variable(s):

The amount of Vitamin C present (in milligrams of Vit. C, per 1000 mg. of fruit).

1.5 pt →

* Control of variability - name at least one form employed in the study:

(i) The experimental units were randomly assigned to treatment groups

1.5 pt →

(ii) Perhaps the fact that each treatment group consisted of multiple packages of fruit can also be considered a form of controlling variability.