Student name:

MATH 120: Elementary Statistics
Spring 2022

Test 2
April 13, 2022

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 of any kind is prohibited for the entire duration of the test. This also applies to any restroom breaks taken during the test.
- The time limit for taking this test is 50 minutes from the scheduled start time. Please turn in your test promptly when time is called to avoid late penalties.
- This test adds up to 50 points. It contains questions numbered 1 through 7.

1. [4 points] Given the probabilities P(U)=0.3, P(V)=0.6, find the probability P(U and V) for each of the following situations:

U and V are disjoint	U and V are independent
Answer: P(U and V)=0, since disjoint events can never occur together.	**Answer**: P(U and V)=P(U)×P(V) = 0.18

Grade: 2 pt. for each correct answer. Reasoning is nice, but not essential for credit.

2. [6 points] Given independent random variables X, Y, with means and standard deviations as shown, find the mean and SD of:

	Mean	SD
X	40	8
Y	100	12

(a) $-2Y$

Mean: $E(-2Y) = -2E(Y)\ \boxed{= -200}$

$SD(-2Y) = |-2|\ SD(Y)\ \boxed{= 24}$

(b) $X + Y$

Mean: $E(X+Y) = E(X) + E(Y)\ \boxed{= 140}$

$SD(X+Y) = \sqrt{\text{var}(X+Y)}$

$\text{var}(X+Y) = \text{var}(X) + \text{var}(Y) = 8^2 + 12^2 = 208$

$\therefore\ SD(X+Y) = \sqrt{208}\ \boxed{= 14.422}$

Grade: (a)=(b)=3 pt. each.
For (a) and (b): 1.5 pt each for correct mean + correct SD. Some steps required.

3. [7 points] Public health officials in a state in the U.S. claim that 18% of adults currently smoke cigarettes. Assume that figure is true, and suppose we pick a random sample of 100 adults from that state and count how many of them smoke.

(a) Under what assumptions can we treat this like a binomial experiment?

Answer: Assume each individual in the sample is independent, has the same probability of being a smoker (i.e., the 18% cited by public health officials), and there are only two possible outcomes (smoker, or non-smoker).

(Basically, any answer that captures the essential features of a fixed number of Bernoulli trials is sufficient.)

(b) Assume those conditions are met, define an appropriate random variable, and indicate which specific binomial probability distribution it follows.

Answer:

Let X = the number of smokers in random samples of 100 adults.

Then X follow the binomial distribution $B(100, 0.18)$.

(c) Suppose we want to compute the probability that exactly 25 adults in our sample of 100 smoke. Write an expression for computing that probability.
Hint: It should contain an appropriate form of ${}_nC_r$.

Answer:

The probability that 25 smoke and 75 don't is: $(0.18)^{25} \cdot (1 - 0.18)^{75}$.

There are ${}_{100}C_{25}$ different ways to select 25 out of 100.

Thus, the probability that exactly 25 smoke $= {}_{100}C_{25} \cdot (0.18)^{25}(0.82)^{75}$

(d) Assume your binomial model can be approximated by an appropriate normal model. What is the correct normal model (e.g., its mean, SD)?

Answer:

For the general model $B(n, p)$, the mean$= np$ and the SD$= \sqrt{np(1-p)}$

For our model: mean$= 18$, SD$= \sqrt{(100)(0.18)(0.82)} = 3.842$.

Thus, the correct normal model to use is: $N(18, 3.842)$.

Grade: (a) = (b) = (c) = 2 points; (d) = 1 point.
For (a): Any answer showing awareness of conditions for Bernoulli trials is ok.
For (b): 1+1 pt = correctly define X + correct mean/SD of binomial model.
For (c): 1.5pt = get $(0.18)^{25}(0.82)^{75}$ part; 0.5 pt = multiply by ${}_{100}C_{25}$.
For (d): 0.5 + 0.5 pt = correct mean + correct SD of normal model.

4. [7 points] An insurance policy costs \$150 annually, and will pay policyholders \$20,000 if they suffer major injury, or \$5000 if they suffer a minor injury. The company estimates that each year 1 in every 1000 policyholders will have a major injury and 1 in 500 will have a minor injury.

(a) Create a probability model for the company's profit on a policy.

(b) Find the expected profit.

(c) Find the standard deviation of the profit.

Solution:

(a) Let X = random variable that represents the company's profit on a policy.

The possible values of X are: $-\$19,850, \quad -\$4,850, \quad \$150$.

Reason: Those are the amounts after accounting for policy cost − payout.

The probability model is shown in the table below:

X (in \$)	$-19,850$	$-4,850$	150
$\mathrm{P}(X)$	0.001	0.002	0.997

Reason: $\mathrm{P}(-19,850)$=0.001 because the company expects 1 in every 1000 policyholders will have a major injury. $\mathrm{P}(-4,850)$=0.002 since 1 in 500 are estimated to have a minor injury. $\mathrm{P}(150)=1-0.001-0.002=0.997$.

(b) $E(X)=\sum x\cdot \mathrm{P}(x)=-19,850\cdot(0.001)-4,850\cdot(0.002)+150\cdot(0.997)$

$$\boxed{E(X)=\$120}=\bar{x}$$

(c) $SD(X)=\sqrt{\sum(x-\bar{x})^2\,\mathrm{P}(x)}$

$$=\sqrt{(-19850-120)^2\cdot 0.001+(-4850-120)^2\cdot 0.002+(150-120)^2\cdot 0.997}$$

$$=\sqrt{449100}=\$670.15$$

Answers: The expected profit on a policy = \$120
The standard deviation of profits = \$670.15

Grade: (a) = 2.5 points; (b) = 2.5 points; (c) = 2 points.
For (a): 1 pt = correct values of X; 1.5 pt = correct values of $\mathrm{P}(X)$.
For (b): 1 pt = show correct plug into formula; 1.5 pt = compute answer (with units).
For (c): 1+1 pt = plug into formula + compute correct answer.

5. [7 points] The distribution of voters registered in a state is the following: 40% Democrat, 36% Republican, 10% other parties, and the rest are independent. Suppose 3 people are selected randomly from this list, find the probability that

(a) None of the 3 is a Democrat.

(b) The group includes at least one independent.

(c) Exactly one of the 3 is Republican.

Be sure to show calculation steps and state any assumptions you make.

Solution:

Let D=Democrat, R=Republican, O=Other, I=Independent.

I assume the 3 individuals selected are independent of each other, since they were randomly selected.

(a) $$\begin{aligned} \text{P(none is D)} &= \text{P}(\sim\text{D and} \sim\text{D and} \sim\text{D}) \\ &= \text{P}(\sim\text{D}) \times \text{P}(\sim\text{D}) \times \text{P}(\sim\text{D}) = (0.6)^3 \boxed{= 0.216} \end{aligned}$$

Note that I got $\text{P}(\sim\text{D}) = 1 - 0.4 = 0.6$.

(b) $$\begin{aligned} \text{P(at least one I)} &= 1 - \text{P(none is I)} \\ &= 1 - \text{P}(\sim\text{I and} \sim\text{I and} \sim\text{I}) \\ &= 1 - (0.86 \times 0.86 \times 0.86) = 0.364 \end{aligned}$$

$$\boxed{\text{P(at least one I)} = 0.364}$$

(c) $$\begin{aligned} \text{P(exactly one R)} &= 3 \times \text{P(R and} \sim\text{R and} \sim\text{R)} \\ &= 3 \times (0.36 \times 0.64 \times 0.64) = 3\times 0.1475 \\ &= 0.442 \end{aligned}$$

Reason: To get exactly one R, that person must be the only R, with the other two being $\sim$R. This can occur in 3 different ways, all having the probability $0.36 \times (0.64)^2$.

$$\boxed{\text{P(exactly one R)} = 0.442}$$

Grade: (a) = 2 points; (b) = 2.5 points; (c) = 2 points; plus, 0.5 pt for clearly stating the assumption of independence.

For each of (a), (b), (c): 0.5 pt for correct answers; the rest of the points are for correct steps.

6. [7 points] The career services office at a small liberal arts college has compiled data on the career path of their recent graduating class. The following 2-way table shows these data, organized by type of career path chosen by graduates in different fields of study.

	Major area of study				
	Fine arts	Humanities	Natural sci.	Social sci.	**Total**
Employed	17	41	43	62	163
Grad school	9	19	23	25	76
Total	26	60	66	87	239

(a) Find the probability that a randomly selected graduate is employed or majored in the natural sciences.
(b) What is the probability that an individual who went to graduate school majored in the fine arts?
(c) Are majoring in the humanities and going to graduate school disjoint, independent, neither, or both?
[Solution must show key steps and correct interpretation of each question - e.g., probability of employed and social science, or P(employed and social science), etc.]

Solution:

Probabilities can be computed in 2 different ways: (1) by directly using the numbers given in the table; (2) using an appropriate formula.

(a) $\text{P(Employed or NS)} = \dfrac{\text{all employed + all NS} - \text{double count}}{\text{Total}} = \dfrac{163+66-43}{239}$

$$\boxed{\text{P(Employed or NS)} = \frac{186}{239} \approx 0.7782}$$

Method 2: $\text{P(Employed or NS)} = \text{P(Employed)} + \text{P(NS)} - \text{P(both)} \approx 0.7782$

(b) This question is asking for P(Fine arts | Grad school)

From the table: $\text{P(Fine arts} \mid \text{Grad school)} = \dfrac{9}{76}\ \boxed{\approx 0.1184}$

Method 2: $\text{P(Fine arts} \mid \text{Grad school)} = \dfrac{\text{P(Fine arts and Grad school)}}{\text{P(Grad school)}}$

$$= \frac{9/239}{76/239} = \frac{9}{76} = \text{same as before}$$

(c) From the table, 19 humanities students went to grad school. That clearly means they **are not** disjoint. To check independence:

$\text{P(Humanities)} = \dfrac{60}{239} = 0.251.$ And $\text{P(Humanities|Grad school)} = \dfrac{19}{76} = 0.25$

These probabilities are very close. Thus, it may be reasonable to conclude majoring in the humanities and going to grad school **are independent**.

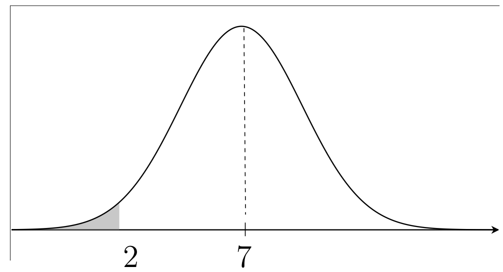
Grade: (a) = (b) = 2.5 points, (c) = 2 points.
For (a)+(b): 1pt = interpret question correctly; 1pt = calculation step; 0.5pt = answer.
For (c): 0.5 + 0.5 pt for correctly addressing disjoint + independent.

7. [4 points ×3] Give brief solutions to the following (unrelated) questions as instructed. It is always helpful to include a relevant sketch, and to show key calculation steps.

(a) Every normal distribution is defined by its parameters, the mean μ, and the standard deviation σ. Suppose for a particular normal distribution we know $\mu = 7$, and 5% of the distribution lies below the value of 2. Find σ.

Solution:

Let $X \sim N(7, \sigma)$
Given 5% is in the left tail when $X = 2$.
z-score for 0.05 in left tail: $z = -1.64$.
This gives: $-1.64 = \dfrac{2-7}{\sigma}$
Therefore, $\sigma = \dfrac{2-7}{-1.64}$ $\boxed{= 3.049}$



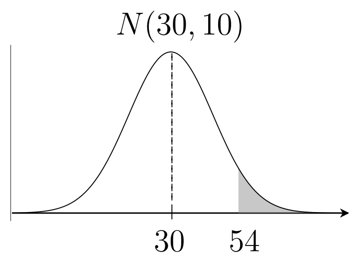
Grade: 1pt = attempt to find z for 5% in left tail.
1pt = do it correctly and find z score; 2pt = compute σ.

(b) Let X be a normally distributed random variable with $\mu = 30$ and $\sigma = 10$.
(i) Find the probability that $X \geq 54$.
(ii) Find the IQR for this distribution.

Solution:

(i) To find the probability that $X > 54$:

$z = \dfrac{54-30}{10} = 2.4$.
From z-table, $P(z > 2.4) = 0.0082$
Thus $\boxed{P(X > 54) = 0.0082}$



(ii) For the IQR, we need to find quartiles. From z-table, 0.25 in left tail corresponds to $z = -0.67$. Thus, $IQR = (2 \times 0.67) \times 10$ $\boxed{= 13.4}$

Grade: 2pt each for (i) and (ii).
For each: 1+1pt for process + answer.

(c) Suppose scores on the math section of the SAT follow the distribution $N(525, 120)$.
(i) Find the z-score for an SAT score of 300.
(ii) Scores on the ACT math test, which is an alternative to the SAT, follow the normal model $N(21, 5.5)$. Suppose person A scored 600 on the math SAT, and person B scored 25 on the ACT math test. Whose performance was better? Why?

Solution:

(i) $z = \dfrac{300-525}{120}$ $\boxed{= -1.875}$

(ii) z-score of person A is: $z_A = \dfrac{600-525}{120} = 0.625$

z-score of person B is: $z_B = \dfrac{25-21}{5.5} = 0.727$

When both scores are scaled by their respective standard deviations, person B scored farther above the ACT mean, than person A scored above the SAT mean. Thus, person B had the better performance.

Grade: 1 pt for (i) and 3 pt for (ii).
For (i): correct answer is sufficient.
For (ii): 1 pt each for z_A, z_B, discussion.