AP Stats Topic 4: Randomness and Probability, Probability Distributions

A simple random sample of adults living in a suburb of a large city was selected. The age and annual income of each adult in the sample were recorded. The resulting data are summarized in the table below.

	Annual Income			
Age Category	$25,000-$35,000	$35,001-$50,000	Over $50,000	Total
21-30	8	15	27	50
31-45	22	32	35	89
46-60	12	14	27	53
Over 60	5	3	7	15
Total	47	64	96	207

(a) What is the probability that a person chosen at random from those in this sample will be in the 31-45 age category?

(b) What is the probability that a person chosen at random from those in this sample whose incomes are over $50,000 will be in the 31-45 age category? Show your work.

(c) Based on your answers to parts (a) and (b), is annual income independent of age category for those in this sample? Explain.

Solution

Part (a): $P(age\ 31-45) = \frac{89}{207} = 0.42995$

Part (b): $P(age\ 31-45 | income\ over\ 50{,}000) = \frac{35}{96} = 0.36458$

Part (c):

If annual income and age were independent, the probabilities in (a) and (b) would be equal. Since these probabilities are not equal, annual income and age category are not independent for adults in this sample.

Airlines routinely overbook flights because they expect a certain number of no-shows. An airline runs a 5 P.M. commuter flight from Washington, D.C., to New York City on a plane that holds 38 passengers. Past experience has shown that if 41 tickets are sold for the flight, then the probability distribution for the number who actually show up for the flight is as shown in the table below.

Number who actually show up	36	37	38	39	40	41
Probability	0.46	0.30	0.16	0.05	0.02	0.01

Assume that 41 tickets are sold for each flight.

(a) There are 38 passenger seats on the flight. What is the probability that all passengers who show up for this flight will get a seat?

(b) What is the expected number of no-shows for this flight?

(c) Given that not all passenger seats are filled on a flight, what is the probability that only 36 passengers showed up for the flight?

Solution

Part (a):

P (everyone gets a seat) = P ($X \leq 38$) $= .46 + .30 + .16 = .92$

OR $= 1 - (.05 + .02 + .01) = .92$

Part (b):

Y = number of no shows

y	0	1	2	3	4	5
p(y)	.01	.02	.05	.16	.30	.46

$E(Y) = 0(.01) + 1(.02) + 2(.05) + 3(.16) + 4(.30) + 5(.46) = 4.1$

OR

$E(X) = 36(.46) + 37(.30) + 38(.16) + 39(.05) + 40(.02) + 41(.01) = 36.9$

$E(Y) = 41 - E(X) = 4.1$

Part (c):

$$P(X=36 \mid \text{not all seats are filled}) = P(X = 36 \mid X < 38) = \frac{P(X = 36)}{P(X < 38)} = \frac{.46}{.76} = .605$$

Two antibiotics are available as treatment for a common ear infection in children.

- Antibiotic A is known to effectively cure the infection 60 percent of the time. Treatment with antibiotic A costs \$50.
- Antibiotic B is known to effectively cure the infection 90 percent of the time. Treatment with antibiotic B costs \$80.

The antibiotics work independently of one another. Both antibiotics can be safely administered to children. A health insurance company intends to recommend one of the following two plans of treatment for children with this ear infection.

- Plan I: Treat with antibiotic A first. If it is not effective, then treat with antibiotic B.
- Plan II: Treat with antibiotic B first. If it is not effective, then treat with antibiotic A.

(a) If a doctor treats a child with an ear infection using plan I, what is the probability that the child will be cured?

If a doctor treats a child with an ear infection using plan II, what is the probability that the child will be cured?

(b) Compute the expected cost per child when plan I is used for treatment.

Compute the expected cost per child when plan II is used for treatment.

(c) Based on the results in parts (a) and (b), which plan would you recommend?

Explain your recommendation.

Solution

Part (a):

Let A be the event "antibiotic A works."
Let B be the event "antibiotic B works."

The probability that a child will be cured with Plan I is:

$$P(\text{Cure}_{\text{I}}) = P(\text{A}) + P(\text{not A})P(\text{B})$$
$$= 0.6 + (0.4 \times 0.9)$$
$$= 0.96$$

The probability that a child will be cured with Plan II is:

$$P(\text{Cure}_{\text{II}}) = P(\text{B}) + P(\text{not B})P(\text{A})$$
$$= 0.9 + (0.1 \times 0.6)$$
$$= 0.96$$

Part (b):

Treatment with antibiotic A costs \$50, and treatment with antibiotic B costs \$80.

The expected cost per child when Plan I is used for treatment is:

$$E(\text{Cost}_{\text{I}}) = \$50 \times 0.6 + \$130 \times 0.4$$
$$= \$30 + \$52$$
$$= \$82$$

The expected cost per child when Plan II is used for treatment is:

$$E(\text{Cost}_{\text{II}}) = \$80 \times 0.9 + \$130 \times 0.1$$
$$= \$72 + \$13$$
$$= \$85$$

Part (c):

Since the probability that a child will be cured is the same under either plan, some other criterion must be used to make a recommendation. From a financial point of view, Plan I should be recommended because the expected cost per child is less than Plan II.

A shopping mall has three automated teller machines (ATMs). Because the machines receive heavy use, they sometimes stop working and need to be repaired. Let the random variable X represent the number of ATMs that are working when the mall opens on a randomly selected day. The table shows the probability distribution of X.

Number of ATMs working when the mall opens	0	1	2	3
Probability	0.15	0.21	0.40	0.24

(a) What is the probability that at least one ATM is working when the mall opens?

(b) What is the expected value of the number of ATMs that are working when the mall opens?

(c) What is the probability that all three ATMs are working when the mall opens, given that at least one ATM is working?

(d) Given that at least one ATM is working when the mall opens, would the expected value of the number of ATMs that are working be less than, equal to, or greater than the expected value from part (b) ? Explain.

Solution

Part (a):

The probability that at least one ATM is working when the mall opens is:

$$P(X \geq 1) = 0.21 + 0.40 + 0.24 = 0.85.$$

Part (b):

The expected value of the number of ATMs that are working when the mall opens is:

$$E(X) = 0(0.15) + 1(0.21) + 2(0.40) + 3(0.24) = 1.73 \text{ machines.}$$

Part (c):

The probability that all three ATMs are working when the mall opens, given that at least one ATM is working is:

$$P(X = 3 \mid X \geq 1) = \frac{P(X = 3 \text{ and } X \geq 1)}{P(X \geq 1)} = \frac{P(X = 3)}{P(X \geq 1)} = \frac{0.24}{0.85} \approx 0.282$$

Part (d):

Given that at least one ATM is working when the mall opens, the expected value of the number of working ATMs would be greater than the expected value calculated in part (b). By eliminating the possibility of 0 working ATMs, the probabilities for 1, 2, and 3 working ATMs all increase proportionally, so the expected value must increase.

At an archaeological site that was an ancient swamp, the bones from 20 brontosaur skeletons have been unearthed. The bones do not show any sign of disease or malformation. It is thought that these animals wandered into a deep area of the swamp and became trapped in the swamp bottom. The 20 left femur bones (thigh bones) were located and 4 of these left femurs are to be randomly selected without replacement for DNA testing to determine gender.

(a) Let X be the number out of the 4 selected left femurs that are from males. Based on how these bones were sampled, explain why the probability distribution of X is not binomial.

(b) Suppose that the group of 20 brontosaurs whose remains were found in the swamp had been made up of 10 males and 10 females. What is the probability that all 4 in the sample to be tested are male?

(c) The DNA testing revealed that all 4 femurs tested were from males. Based on this result and your answer from part (b), do you think that males and females were equally represented in the group of 20 brontosaurs stuck in the swamp? Explain.

(d) Is it reasonable to generalize your conclusion in part (c) pertaining to the group of 20 brontosaurs to the population of all brontosaurs? Explain why or why not.

Solution

Part (a):

X is not binomial since the trials are not independent and the conditional probabilities of selecting a male change at each trial depending on the previous outcome(s), due to the sampling without replacement.

Part (b):

$$P(X=4)=\left(\frac{10}{20}\right)\left(\frac{9}{19}\right)\left(\frac{8}{18}\right)\left(\frac{7}{17}\right)=\frac{5040}{116280}=0.043$$

Part (c):

No. If males and females were equally represented, the probability of observing four males is small (0.043).

Part (d):

No, we can't generalize to the population of all brontosaurs because it is not reasonable to regard this sample as a random sample from the population of all brontosaurs; there is reason to suspect that this sampling method might cause bias.

Die A has four 9's and two 0's on its faces. Die B has four 3's and two 11's on its faces. When either of these dice is rolled, each face has an equal chance of landing on top. Two players are going to play a game. The first player selects a die and rolls it. The second player rolls the remaining die. The winner is the player whose die has the higher number on top.

a. Suppose you are the first player and you want to win the game. Which die would you select? Justify your answer.

b. Suppose the player using die A receives 45 tokens each time he or she wins the game. How many tokens must the player using die B receive each time he or she wins in order for this to be a fair game? Explain how you found your answer.

(A fair game is one in which the player using die A and the player using die B both end up with the same number of tokens in the long run.)

Solution

Possible Outcomes

Die A	Die B	Winner	Prob
9	3	A	(2/3)(2/3) = 4/9
9	11	B	(2/3)(1/3) = 2/9
0	3	B	(1/3)(2/3) = 2/9
0	11	B	(1/3)(1/3) = 1/9

OR

		DIE A					
		0	0	9	9	9	9
DIE B	3	B	B	A	A	A	A
	3	B	B	A	A	A	A
	3	B	B	A	A	A	A
	3	B	B	A	A	A	A
	11	B	B	B	B	B	B
	11	B	B	B	B	B	B

Winner	Prob
A	16/36 = 4/9
B	20/36 = 5/9

a. Choose die B, because the probability of winning is higher (5/9 compared to 4/9 for die A)

b. Let X be the number of tokens the player using die B should receive. For the game to be fair, we need

$45(4/9) = X(5/9)$

Solving this equation for X gives X = 36. Player B should receive 36 tokens.

The ELISA tests whether a patient has contracted HIV. The ELISA is said to be positive if it indicates that HIV is present in a blood sample, and the ELISA is said to be negative if it does not indicate that HIV is present in a blood sample. Instead of directly measuring the presence of HIV, the ELISA measures levels of antibodies in the blood that should be elevated if HIV is present. Because of variability in antibody levels among human patients, the ELISA does not always indicate the correct result.

As part of a training program, staff at a testing lab applied the ELISA to 500 blood samples known to contain HIV. The ELISA was positive for 489 of those blood samples and negative for the other 11 samples. As part of the same training program, the staff also applied the ELISA to 500 other blood samples known to not contain HIV. The ELISA was positive for 37 of those blood samples and negative for the other 463 samples.

(a) When a new blood sample arrives at the lab, it will be tested to determine whether HIV is present. Using the data from the training program, estimate the probability that the ELISA would be positive when it is applied to a blood sample that does not contain HIV.

(b) Among the blood samples examined in the training program that provided positive ELISA results for HIV, what proportion actually contained HIV?

(c) When a blood sample yields a positive ELISA result, two more ELISAs are performed on the same blood sample. If at least one of the two additional ELISAs is positive, the blood sample is subjected to a more expensive and more accurate test to make a definitive determination of whether HIV is present in the sample. Repeated ELISAs on the same sample are generally assumed to be independent. Under the assumption of independence, what is the probability that a new blood sample that comes into the lab will be subjected to the more expensive test if that sample does not contain HIV?

Solution

Part (a):

The estimated probability of a positive ELISA if the blood sample does not have HIV present is

$$\frac{37}{500} \quad OR \quad \frac{37}{500} = 0.074$$

Part (b):

A total of 489 + 37 = 526 blood samples resulted in a positive ELISA. Of these, 489 samples actually contained HIV. Therefore the proportion of samples that resulted in a positive ELISA that actually contained HIV is

$$\frac{489}{526} \quad OR \quad \frac{489}{526} \approx 0.9297$$

Part (c):

From part (a), the probability that the ELISA will be positive, given that the blood sample does not actually have HIV present, is 0.074. Thus, the probability of a negative ELISA, given that the blood sample does not actually have HIV present, is $1 - 0.074 = 0.926$.

P(new blood sample that does not contain HIV will be subjected to the more expensive test)

= *P*(*1st ELISA positive* and *2nd ELISA positive* OR *1st ELISA positive* and *2nd ELISA negative* and *3rd ELISA positive* | *HIV not present in blood*)
= *P*(*1st ELISA positive* and *2nd ELISA positive* | *HIV not present in blood*)
+ *P*(*1st ELISA positive* and *2nd ELISA negative* and *3rd ELISA positive* | *HIV not present in blood*)
= (0.074)(0.074) + (0.074)(0.926)(0.074)
= 0.005476 + 0.005070776
= 0.010546776
≈ 0.0105

OR

A local arcade is hosting a tournament in which contestants play an arcade game with possible scores ranging from 0 to 20. The arcade has set up multiple game tables so that all contestants can play the game at the same time; thus contestant scores are independent. Each contestant's score will be recorded as he or she finishes, and the contestant with the highest score is the winner.

After practicing the game many times, Josephine, one of the contestants, has established the probability distribution of her scores, shown in the table below.

Josephine's Distribution				
Score	16	17	18	19
Probability	0.10	0.30	0.40	0.20

Crystal, another contestant, has also practiced many times. The probability distribution for her scores is shown in the table below.

Crystal's Distribution			
Score	17	18	19
Probability	0.45	0.40	0.15

(a) Calculate the expected score for each player.

(b) Suppose that Josephine scores 16 and Crystal scores 17. The difference (Josephine minus Crystal) of their scores is -1. List all combinations of possible scores for Josephine and Crystal that will produce a difference (Josephine minus Crystal) of -1, and calculate the probability for each combination.

(c) Find the probability that the difference (Josephine minus Crystal) in their scores is -1.

(d) The table below lists all the possible differences in the scores between Josephine and Crystal and some associated probabilities.

Distribution (Josephine minus Crystal)						
Difference	-3	-2	-1	0	1	2
Probability	0.015			0.325	0.260	0.090

Complete the table and calculate the probability that Crystal's score will be higher than Josephine's score.

Solution

Part (a):

The expected scores are as follows:

Josephine
$\mu_J = 16(0.1) + 17(0.3) + 18(0.4) + 19(0.2) = 17.7$
Crystal
$\mu_C = 17(0.45) + 18(0.4) + 19(0.15) = 17.7$

Part (b):

J	C	Probability
16	17	(0.1)(0.45) = 0.045
17	18	(0.3)(0.40) = 0.12
18	19	(0.4)(0.15) = 0.06

Part (c):

The probability is
0.045 + 0.12 + 0.06 = 0.225

Part (d):

P(difference = –1) = 0.225 (from part c)
P(difference = –2) = 1 – 0.015 – 0.225 – 0.325 – 0.260 – 0.90 = 0.085

Distribution of Josephine – Crystal

Differences	–3	–2	–1	0	1	2
Probability	0.015	**0.085**	**0.225**	0.325	0.260	0.090

The probability that Crystal's score is higher than Josephine's score is
P(difference < 0) = 0.015 + 0.085 + 0.225 = 0.325

Let the random variable X represent the number of telephone lines in use by the technical support center of a software manufacturer at noon each day. The probability distribution of X is shown in the table below.

x	0	1	2	3	4	5
$p(x)$	0.35	0.20	0.15	0.15	0.10	0.05

(a) Calculate the expected value (the mean) of X.

(b) Using past records, the staff at the technical support center randomly selected 20 days and found that an average of 1.25 telephone lines were in use at noon on those days. The staff proposes to select another random sample of 1,000 days and compute the average number of telephone lines that were in use at noon on those days. How do you expect the average from this new sample to compare to that of the first sample? Justify your response.

(c) The median of a random variable is defined as any value x such that $P(X \leq x) \geq 0.5$ and $P(X \geq x) \geq 0.5$. For the probability distribution shown in the table above, determine the median of X.

(d) In a sentence or two, comment on the relationship between the mean and the median relative to the shape of this distribution.

Solution

Part (a):

The expected number of telephone lines in use by the technical support center at noon is:

$$E(X) = 0 \times 0.35 + 1 \times 0.2 + 2 \times 0.15 + 3 \times 0.15 + 4 \times 0.1 + 5 \times 0.05$$
$$= 1.6$$

Part (b):

We would expect the average based on 1,000 days to be closer to 1.6 than the first average based on 20 days. Both averages have the same expected value (1.6), but the variability for sample averages based on 1,000 days is smaller than the variability for sample averages based on 20 days.

Part (c):

The median of X is 1.

x	$P(X \leq x)$	$P(X \geq x)$
0	0.35	1.0
1	0.55	0.65
2	0.70	0.45
3	0.85	0.30
4	0.95	0.15
5	1.0	0.05

OR

The median of X is 1 because $P(X \leq 1) = 0.55 \geq 0.50$ and $P(X \geq 1) = 0.65 \geq 0.50$.

Part (d):

The probability histogram is clearly skewed to the right (or toward the larger values) so the mean (1.6) is larger than the median (1), as is typical for a right-skewed distribution.

A test consisting of 25 multiple-choice questions with 5 answer choices for each question is administered. For each question, there is only 1 correct answer.

(a) Let X be the number of correct answers if a student guesses randomly from the 5 choices for each of the 25 questions. What is the probability distribution of X?

This test, like many multiple-choice tests, is scored using a penalty for guessing. The test score is determined by awarding 1 point for each question answered correctly, deducting 0.25 point for each question answered incorrectly, and ignoring any question that is omitted. That is, the test score is calculated using the following formula.

Score = (1 × number of correct answers) – (0.25 × number of incorrect answers) + (0 × number of omits)

For example, the score for a student who answers 17 questions correctly, answers 3 questions incorrectly, and omits 5 questions is

$$\text{Score} = (1 \times 17) - (0.25 \times 3) + (0 \times 5) = 16.25.$$

(b) Suppose a student knows the correct answers for 18 questions, answers those 18 questions correctly, and chooses randomly from the 5 choices for each of the other 7 questions. Show that the expected value of the student's score is 18 when using the scoring formula above.

(c) A score of at least 20 is needed to pass the test. Suppose a student knows the correct answers for 18 questions, answers those 18 questions correctly, and chooses randomly from the 5 choices for each of the other 7 questions. What is the probability that the student will pass the test?

Solution

Part (a):

Let X denote the number of correct guesses, assuming that a student guesses randomly among the five options on all 25 questions. Then X has a binomial probability distribution with $n = 25$ and $p = \frac{1}{5} = 0.20$.

Part (b):

Let Y denote the number of correct responses on the seven questions for which the student guesses randomly from among the five options. Then Y has a binomial probability distribution with $n = 7$ and $p = 0.20$. Then the expected value of Y, $E(Y) = np = 7(0.20) = 1.4$ correct responses.

Next, using the scoring formula provided,
$\text{Score} = (18 + Y) - 0.25(7 - Y) + 0(0) = 16.25 + 1.25Y$.

The expected exam score is therefore:
$E(\text{Score}) = E(16.25 + 1.25Y) = 16.25 + 1.25E(Y) = 16.25 + 1.25(1.4) = 16.25 + 1.75 = 18$ correct responses.

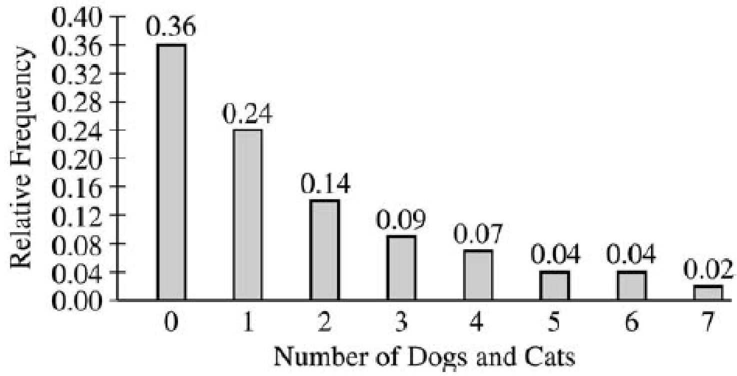
Part (c):

Let Y be defined as in part (b). The student passes when Score ≥ 20, which means that $16.25 + 1.25Y \geq 20$, which means that $Y \geq \frac{20 - 16.25}{1.25} = 3$. In other words, in order to pass, the student must get three or more correct from the seven questions on which the student guesses.

Y has a binomial probability distribution with $n = 7$ and $p = 0.20$, so

$$P(Y \geq 3) = 1 - P(Y \leq 2) = 1 - \left[\binom{7}{0}(.2)^0(.8)^7 + \binom{7}{1}(.2)^1(.8)^6 + \binom{7}{2}(.2)^2(.8)^5\right] = 1 - 0.852 = 0.148.$$

The graph below displays the relative frequency distribution for X, the total number of dogs and cats owned per household, for the households in a large suburban area. For instance, 14 percent of the households own 2 of these pets.



(a) According to a local law, each household in this area is prohibited from owning more than 3 of these pets. If a household in this area is selected at random, what is the probability that the selected household will be in violation of this law? Show your work.

(b) If 10 households in this area are selected at random, what is the probability that exactly 2 of them will be in violation of this law? Show your work.

(c) The mean and standard deviation of X are 1.65 and 1.851, respectively. Suppose 150 households in this area are to be selected at random and $\overline{X}$, the mean number of dogs and cats per household, is to be computed. Describe the sampling distribution of $\overline{X}$, including its shape, center, and spread.

<u>Solution</u>

Part (a):

$$P(X > 3) = 0.07 + 0.04 + 0.04 + 0.02 = 0.17.$$

Part (b):

Y = number of households in violation.

Y has a binomial distribution with $n = 10$ and $p = 0.17$.

$$P(Y = 2) = \binom{10}{2}(0.17)^2(0.83)^8 = 0.2929.$$

Part (c):

The distribution of $\overline{X}$ will:

1. be approximately normal;
2. have mean $\mu_{\overline{X}} = \mu = 1.65$;
3. have standard deviation $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{1.851}{\sqrt{150}} = 0.1511.$

The table below shows the political party registration by gender of all 500 registered voters in Franklin Township.

PARTY REGISTRATION–FRANKLIN TOWNSHIP

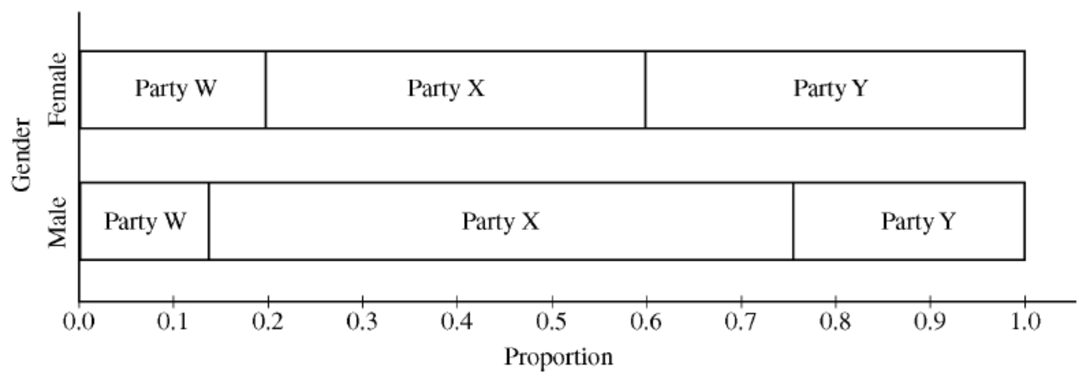
	Party W	Party X	Party Y	Total
Female	60	120	120	300
Male	28	124	48	200
Total	88	244	168	500

(a) Given that a randomly selected registered voter is a male, what is the probability that he is registered for Party Y?

(b) Among the registered voters of Franklin Township, are the events "is a male" and "is registered for Party Y" independent? Justify your answer based on probabilities calculated from the table above.

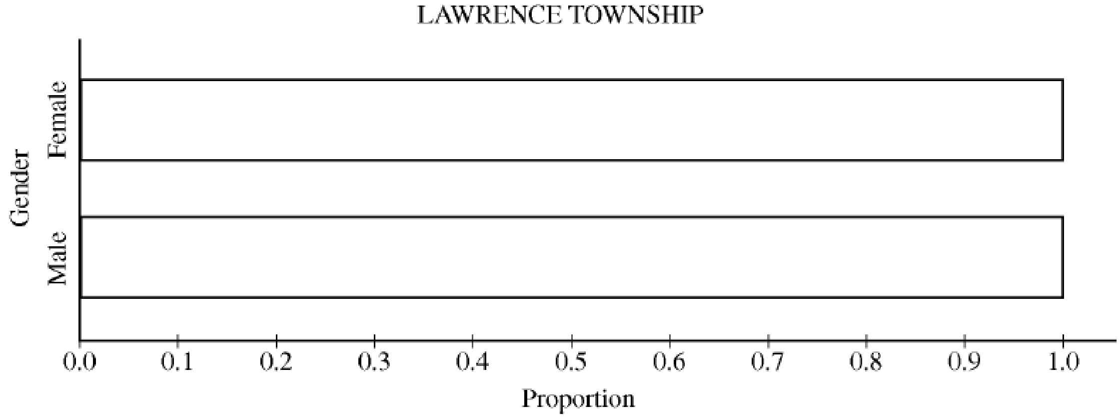
(c) One way to display the data in the table is to use a segmented bar graph. The following segmented bar graph, constructed from the data in the party registration–Franklin Township table, shows party-registration distributions for males and females in Franklin Township.

FRANKLIN TOWNSHIP



In Lawrence Township, the proportions of all registered voters for Parties W, X, and Y are the same as for Franklin Township, and party registration is independent of gender. Complete the graph below to show the distributions of party registration by gender in Lawrence Township.

LAWRENCE TOWNSHIP



Solution

Part (a):

Of the 200 male registered voters in Franklin Township, 48 are registered for Party Y. Therefore the conditional probability that a randomly selected voter is registered for Party Y, given that the voter is a male, is $\frac{48}{200} = 0.24$.

Part (b):

No, the events "is a male" and "is registered for Party Y" are not independent. One justification of this conclusion is to note that the conditional probability of the event "is registered for Party Y" given the event "is a male" — which was computed in part (a) — is not equal to the probability of the event "is registered for Party Y," as shown below.

$$P(\text{is registered for Party Y} \mid \text{is a male}) = 0.24$$
$$P(\text{is registered for Party Y}) = \frac{168}{500} = 0.336$$

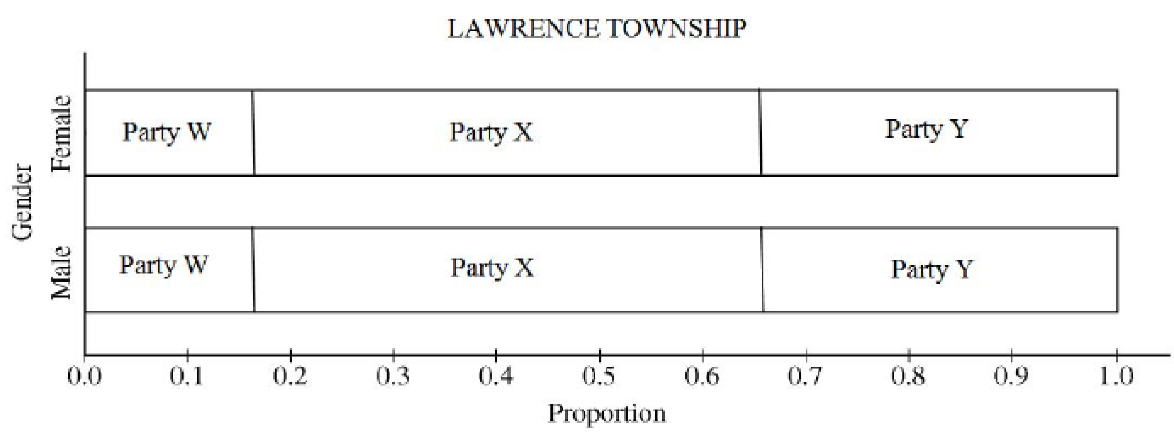
Because $0.24 \neq 0.336$, the two events are not independent.

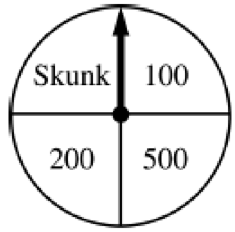
Part (c):

The marginal proportions of voters registered for each of the three political parties (without regard to gender) are given below.

$$\text{Party W: } \frac{88}{500} = 0.176$$
$$\text{Party X: } \frac{244}{500} = 0.488$$
$$\text{Party Y: } \frac{168}{500} = 0.336$$

Because party registration is independent of gender in Lawrence Township, the proportions of males and females registered for each party must be identical to each other and also identical to the marginal proportion of voters registered for that party. Using the order Party W, Party X, and Party Y, the graph for Lawrence Township is displayed below.





Contestants on a game show spin a wheel like the one shown in the figure above. Each of the four outcomes on this wheel is equally likely and outcomes are independent from one spin to the next.

- The contestant spins the wheel.
- If the result is a skunk, no money is won and the contestant's turn is finished.
- If the result is a number, the corresponding amount in dollars is won. The contestant can then stop with those winnings or can choose to spin again, and his or her turn continues.
- If the contestant spins again and the result is a skunk, all of the money earned on that turn is lost and the turn ends.
- The contestant may continue adding to his or her winnings until he or she chooses to stop or until a spin results in a skunk.

(a) What is the probability that the result will be a number on all of the first three spins of the wheel?

(b) Suppose a contestant has earned $800 on his or her first three spins and chooses to spin the wheel again. What is the expected value of his or her total winnings for the four spins?

(c) A contestant who lost at this game alleges that the wheel is not fair. In order to check on the fairness of the wheel, the data in the table below were collected for 100 spins of this wheel.

Result	Skunk	$100	$200	$500
Frequency	33	21	20	26

Based on these data, can you conclude that the four outcomes on this wheel are not equally likely? Give appropriate statistical evidence to support your answer.

Solution

Part (a):

P(a number on all 3 spins) = [P(number)]^3 since the outcomes are independent from spin to spin

$$= \left(\frac{3}{4}\right)^3 = 0.4219$$

Part (b):

Winnings	0	900	1000	1300
Probability	0.25	0.25	0.25	0.25

E(winnings) = $\sum x_i p_i$ = 0(0.25)+900(0.25)+1000(0.25)+1300(0.25) = 800

Or

E(winnings on 4th spin) = -800(0.25) + 100(0.25) + 200(0.25) + 500(0.25) = 0

So E(winnings) = initial amount + E(winnings on 4th spin) = 800 + 0 = 800

Part (c):

Element 1: States a correct pair of hypotheses

H_0: The four outcomes are equally likely $\left(\text{or } p_1 = p_2 = p_3 = p_4 = \frac{1}{4}\right)$

H_a: The four outcomes are not equally likely (or at least one p_i differs from $\frac{1}{4}$)

Element 2: Identifies a correct test (by name or by formula) and checks appropriate conditions.

Chi-square test (for goodness of fit) $\chi^2 = \sum \frac{(Obs - Exp)^2}{Exp}$

Conditions: Outcomes of spins of the wheel are independent and large sample size.

The problem states that successive spins of the wheel are independent.

The expected counts are all equal to 25, which is greater than 5 (or 10), so the sample size is large enough to proceed.

3. A medical researcher surveyed a large group of men and women about whether they take medicine as prescribed. The responses were categorized as never, sometimes, or always. The relative frequency of each category is shown in the table.

	Never	Sometimes	Always	Total
Men	0.0564	0.2016	0.2120	0.4700
Women	0.0636	0.1384	0.3280	0.5300
Total	0.1200	0.3400	0.5400	1.0000

(a) One person from those surveyed will be selected at random.

(i) What is the probability that the person selected will be someone whose response is never and who is a woman?

(ii) What is the probability that the person selected will be someone whose response is never or who is a woman?

(iii) What is the probability that the person selected will be someone whose response is never given that the person is a woman?

(b) For the people surveyed, are the events of being a person whose response is never and being a woman independent? Justify your answer.

(c) Assume that, in a large population, the probability that a person will always take medicine as prescribed is 0.54. If 5 people are selected at random from the population, what is the probability that at least 4 of the people selected will always take medicine as prescribed? Support your answer.

a. i) P(Never and Woman) = 0.0636

ii) P(Never or Woman) = P(Never) + P(Woman) − P(Never and Woman)

= 0.1200 + 0.5300 − 0.0636

= 0.5864

P(Never or Woman) = 0.5864

iii) $P(\text{Never} \mid \text{Woman}) = \frac{P(\text{Never and Woman})}{P(\text{Woman})} = \frac{0.0636}{0.5300}$

$P(\text{Never} \mid \text{Woman}) = 0.12$

b. For independence:

$P(\text{Woman}) = P(\text{Woman} \mid \text{Never})$

$P(\text{Woman}) = 0.53$

$P(\text{Women} \mid \text{Never}) = \frac{0.0636}{0.1200} = 0.53$

Because P(Woman) = P(Woman | Never), we can conclude being a person who answers "Never" and being a woman are independent.

c. P(Medicine) = 0.54

X = # who take prescribed medicine out of 5.

Binomial, p = 0.54, n = 5

$P(X \geq 4) = P(X=4) + P(X=5)$

$= \binom{5}{4}(0.54)^4(0.46)^1 + \binom{5}{5}(0.54)^5(0.46)^0$

$= 0.1956 + 0.0459$

$= 0.2415$

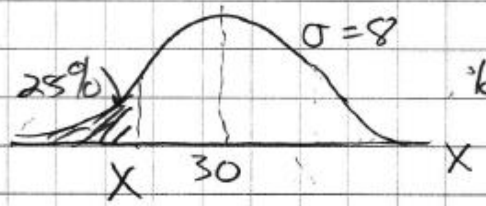
5. A company that manufactures smartphones developed a new battery that has a longer life span than that of a traditional battery. From the date of purchase of a smartphone, the distribution of the life span of the new battery is approximately normal with mean 30 months and standard deviation 8 months. For the price of $50, the company offers a two-year warranty on the new battery for customers who purchase a smartphone. The warranty guarantees that the smartphone will be replaced at no cost to the customer if the battery no longer works within 24 months from the date of purchase.

(a) In how many months from the date of purchase is it expected that 25 percent of the batteries will no longer work? Justify your answer.

(b) Suppose one customer who purchases the warranty is selected at random. What is the probability that the customer selected will require a replacement within 24 months from the date of purchase because the battery no longer works?

(c) The company has a gain of $50 for each customer who purchases a warranty but does not require a replacement. The company has a loss (negative gain) of $150 for each customer who purchases a warranty and does require a replacement. What is the expected value of the gain for the company for each warranty purchased?

a.



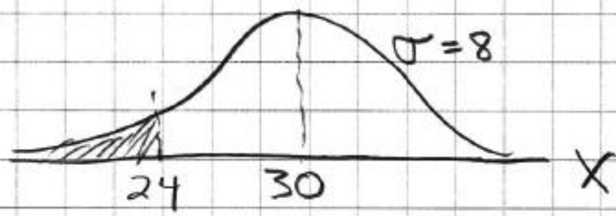
X= months of battery life

z for the 25th percentile = −0.6745

$X = 30 + (-0.6745)(8) = 24.604$ months

We expect at about 24.6 months 25% of the phones have failed.

b.



$P(X \leq 24) = 0.2266$

C.

Expected Gain $= (-150)(0.2266) + (50)(1 - 0.2266)$

$= -33.99 + 38.67$

$= \$4.68$

For each warranty purchased, the expected gain is \$4.68