Assignment 6

Question 1

The St. Louis encephalitis virus is also spread by mosquitoes. The probability that a person infected with the virus gets encephalitis is 1 in 150. But of the 70 who developed St. Louis encephalitis in and around Monroe, La., in 2001 three died. Using 3/70 as the probability of a sick person dying from the disease, find the probability that in a sample of 10 independent cases:

a. all will recover b. exactly 9 persons will recover c. at least 9 persons recover

Probability that a person with encephalitis dies:

$$P(\text{die}) = \frac{3}{70}$$

Probability that a person recovers:

$$P(\text{recover}) = 1 - \frac{3}{70} = \frac{67}{70}$$

We have 10 independent cases. The number of recoveries follows a binomial distribution:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Where: $n = 10$ (total cases), $p = \frac{67}{70}$ (probability of recovery), k is the number of recoveries.

a) Probability that all will recover $P(X = 10)$:

$$P(X = 10) = \left(\frac{67}{70}\right)^{10}$$

$$P(X = 10) \approx 0.6453 \quad (64.53\%)$$

b) Probability that exactly 9 persons will recover $P(X = 9)$:

$$P(X = 9) = \binom{10}{9} \left(\frac{67}{70}\right)^9 \left(\frac{3}{70}\right)^1$$

$$= 10 \times \left(\frac{67}{70}\right)^9 \times \left(\frac{3}{70}\right)$$

$$P(X = 9) \approx 0.2889 \quad (28.89\%)$$

c) Probability that at least 9 persons recover $P(X \geq 9)$:

$$P(X \geq 9) = P(X = 9) + P(X = 10)$$

$$P(X \geq 9) \approx 0.9343 \quad (93.43\%)$$

Final probabilities:

a) All recover: 0.6453 (64.53%) b) Exactly 9 recover: 0.2889 (28.89%) c) At least 9 recover: 0.9343 (93.43%)

Question 2

A speedrunner is trying to get a world record in Super Mario Bros. 3 in the "any% warpless" category. However, near the end of the run, there are 3 independent coin flips such that if any are heads, a hand will pull you down to play an extra level, which takes up additional time. For a world record to be possible, the speedrunner needs a "no hands run", where all 3 flips are tails.

(a) What is the probability that the speedrunner will get a "no hands run" in at least one of their first 3 attempts?

(b) What is the standard deviation of the number of attempts until the first "no hands run"?

(c) Suppose that after 500 attempts, they have had 43 "no hands runs". What is the expected number of additional attempts until the next "no hands run"?

There are 3 independent coin flips. Each flip lands on tails with probability:

$$P(T) = \frac{1}{2}$$

For a "no hands run", all 3 flips must be tails:

$$P(\text{"no hands run"}) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

The probability of not getting a "no hands run" is:

$$P(\text{failure}) = 1 - \frac{1}{8} = \frac{7}{8}$$

a) Probability of at least one "no hands run" in the first 3 attempts Using the complement rule:

$$P(\text{at least one success}) = 1 - P(\text{all failures})$$

Since the probability of failure in one attempt is $\frac{7}{8}$, the probability of failing 3 times in a row is:

$$P(\text{all failures}) = \left(\frac{7}{8}\right)^3$$

Thus, the probability of at least one success is:

$$P(\text{at least one "no hands run"}) = 1 - \left(\frac{7}{8}\right)^3$$

$$P(\text{at least one success}) \approx 0.3301 \quad (33.01\%)$$

b) Standard deviation of the number of attempts until the first "no hands run" The number of attempts until the first success follows a geometric distribution:

$$X \sim \text{Geom}(p)$$

where $p = \frac{1}{8}$ is the probability of a "no hands run" in one attempt.

The standard deviation of a geometric distribution is given by:

$$\sigma = \frac{\sqrt{1-p}}{p}$$

$$\sigma \approx 7.4833$$

c) Expected number of additional attempts until the next "no hands run" From the problem statement: 500 attempts have resulted in 43 "no hands runs". The empirical probability of a "no hands run" is:

$$\hat{p} = \frac{43}{500}$$

Since the number of attempts until the next success is geometrically distributed with probability $\hat{p}$, the expected number of additional attempts is:

$$E[X] = \frac{1}{\hat{p}}$$

$$E[X] \approx 11.63$$

Final Answers: a) Probability of at least one "no hands run" in the first 3 attempts: 0.3301 (33.01%) b) Standard deviation of attempts until first "no hands run": 7.4833 c) Expected number of additional attempts until next "no hands run": 11.63

Question 3

Show that if X is a discrete random variable, and a and b are constants, then

$$E(aX + b) = aE(X) + b \quad \text{and} \quad \text{var}(aX + b) = a^2 \text{var}(X).$$

By the definition of expectation for a discrete random variable:

$$E(X) = \sum_x x P(X = x)$$

Applying linearity of expectation:

$$E(aX + b) = \sum_x (aX + b) P(X = x)$$

Using summation properties:

$$E(aX + b) = \sum_x [aXP(X = x) + bP(X = x)]$$

Splitting the summation:

$$E(aX + b) = a \sum_x XP(X = x) + b \sum_x P(X = x)$$

Since $\sum_x P(X = x) = 1$, we get:

$$E(aX + b) = aE(X) + b$$

The formula is proved.

We need to prove:

$$\text{var}(aX + b) = a^2\text{var}(X)$$

By the definition of variance:

$$\text{var}(X) = E[X^2] - (E[X])^2$$

Using the variance formula:

$$\text{var}(aX + b) = E[(aX + b)^2] - (E[aX + b])^2$$

Expanding the square:

$$E[(aX + b)^2] = E[a^2X^2 + 2abX + b^2]$$

Using expectation properties:

$$E[a^2X^2 + 2abX + b^2] = a^2E[X^2] + 2abE[X] + b^2$$

Also, from the expectation property:

$$(E[aX + b])^2 = (aE[X] + b)^2 = a^2(E[X])^2 + 2abE[X] + b^2$$

Variance is:

$$\text{var}(aX + b) = \left(a^2E[X^2] + 2abE[X] + b^2\right) - \left(a^2(E[X])^2 + 2abE[X] + b^2\right)$$

Cancel out terms:

$$\text{var}(aX + b) = a^2E[X^2] - a^2(E[X])^2$$

Factor out a^2:

$$\text{var}(aX + b) = a^2\left(E[X^2] - (E[X])^2\right)$$

Since $\text{var}(X) = E[X^2] - (E[X])^2$, we get:

$$\text{var}(aX + b) = a^2\text{var}(X)$$

The formula is proved.

Question 4

Let W be the number of failures until the first success, that is, $W = T - 1$.

Show that W has mean $\frac{1-\pi}{\pi}$ and variance $\frac{1-\pi}{\pi^2}$.

W represents the number of failures before the first success. This follows the geometric distribution with probability mass function (PMF):

$$P(W = k) = (1 - \pi)^k \pi, \quad k = 0, 1, 2, \ldots$$

where π is the probability of success.

Since W is the number of failures, it is related to a geometric random variable T (which represents the trial number of the first success):

$$W = T - 1$$

Given that $T \sim \mathrm{Geom}(\pi)$, we can derive the expectation and variance for W.

a) Expected Value (Mean) of W Use the expectation formula for a geometric distribution:

$$E[T] = \frac{1}{\pi}$$

Since $W = T - 1$, we can use linearity of expectation:

$$E[W] = E[T] - 1 = \frac{1}{\pi} - 1$$

$$E[W] = \frac{1 - \pi}{\pi}$$

The expected number of failures before the first success is:

$$E[W] = \frac{1 - \pi}{\pi}$$

b) Variance of W The variance of a geometric random variable T is:

$$\mathrm{Var}(T) = \frac{1 - \pi}{\pi^2}$$

Since $W = T - 1$, the variance remains unchanged because subtracting a constant does not affect variance:

$$\mathrm{Var}(W) = \mathrm{Var}(T) = \frac{1 - \pi}{\pi^2}$$

Thus, the variance of W is:

$$\mathrm{Var}(W) = \frac{1 - \pi}{\pi^2}$$

These results confirm the expected mean and variance for the number of failures before the first success in a geometric setting.

Question 5

Suppose that you design a new recipe for gerbil food, and are curious whether or not gerbils like the taste better than the most popular brand. Describe a hypothesis test you could run. Make sure to say what H_0

and H_a are for your test. In words, what does it mean if you make a Type I error when evaluating this test? What about a Type II error?

To determine if gerbils prefer the new food over the most popular brand, we set up the hypotheses as follows:

Null Hypothesis (H_0): Gerbils have no preference or prefer the most popular brand equally to the new food.

$$H_0 : \text{Gerbils show no preference for the new food.}$$

Alternative Hypothesis (H_a): Gerbils prefer the new food over the most popular brand.

$$H_a : \text{Gerbils prefer the new food.}$$

1. Design an experiment where a group of gerbils is given a choice between the new food and the most popular brand.
2. Collect data on how many times gerbils choose the new food over the other.
3. Use a statistical test (e.g., a chi-square test for preference or a proportion test) to analyze whether the preference for the new food is significantly higher than expected by random chance.
4. Set a significance level (typically $\alpha = 0.05$) to determine whether to reject H_0.

Type I Error (False Positive):
Rejecting H_0 when it is actually true. Meaning: Concluding that gerbils prefer the new food when in reality, they do not. Consequence: The new food is marketed as "preferred," but gerbils may not actually like it more.

Type II Error (False Negative):
Failing to reject H_0 when H_a is actually true. Meaning: Concluding that gerbils do not prefer the new food when in reality, they do. Consequence: A potentially better product is not recognized, and the new recipe is discarded despite its actual preference.

Question 6

A box of dice contains one that has been altered by changing the side with one dot (ace) to five dots. Thus, the altered die has no aces but two sides with five dots. A die is chosen at random and rolled 24 times. The die is expected to come up ace 4 times in 24 rolls. Consider the hypotheses:

H_0: It is a good die. H_a: It is the altered die.

Use the following decision rule: Reject H_0 if an ace never comes up in the 24 rolls, and retain H_0 otherwise.

a. Find α, the probability of a Type I error.
b. Find β, the probability of a Type II error.

We are rolling a randomly chosen die 24 times and checking if an ace (1 dot) appears at least once.

Good Die (Fair Die) Probabilities: Probability of rolling an ace (1) on a fair die:

$$P(\text{ace}|H_0) = \frac{1}{6}$$

Expected number of aces in 24 rolls:

$$E[\text{aces}|H_0] = 24 \times \frac{1}{6} = 4$$

Altered Die Probabilities: The altered die has no ace (1 dot), so:

$$P(\text{ace}|H_a) = 0$$

Decision Rule: Reject H_0 (conclude die is altered) if no aces appear in the 24 rolls. Retain H_0 (conclude die is fair) if at least one ace appears.

(a) Finding α (Type I Error Probability) A Type I error occurs when we reject H_0 when it is actually true, meaning we conclude the die is altered when it is actually fair.

This happens if we never roll an ace in 24 rolls, given that the die is fair. The probability of not rolling an ace in one roll for a fair die:

$$P(\text{no ace in 1 roll}|H_0) = 1 - \frac{1}{6} = \frac{5}{6}$$

The probability of not rolling an ace in all 24 rolls:

$$P(\text{no ace in 24 rolls}|H_0) = \left(\frac{5}{6}\right)^{24}$$

This value is α, the probability of a Type I error.

$$\alpha = \left(\frac{5}{6}\right)^{24} \approx 0.0126 \quad (1.26\%)$$

There is about a 1.26% chance that we wrongly conclude the die is altered when it is actually fair.

(b) Finding β (Type II Error Probability) A Type II error occurs when we fail to reject H_0 when H_a is actually true, meaning we conclude the die is fair when it is actually altered.

If the die is altered, there is no chance of rolling an ace ($P(\text{ace}|H_a) = 0$). This means we will never roll an ace in 24 rolls. Since our decision rule is to reject H_0 only if no aces appear, and this always happens under H_a, we will never reject H_0. Therefore, the probability of making a Type II error is:

$$\beta = 1$$

If the die is altered, we will always fail to reject H_0, meaning we will never detect that the die is altered.

Question 7

Repeat Exercise 6.3 for the decision rule: Reject H_0 if the coin comes up heads 11 or 12 times in the 12 tosses, and retain H_0 otherwise.

We have two types of coins:

1. Fair Coin (H_0): The probability of heads is:

$$P(H|H_0) = 0.5$$

1. Biased (Altered) Coin (H_a): The probability of heads is:

$$P(H|H_a) = 0.75$$

We toss a randomly selected coin 12 times and count the number of heads.

Decision Rule (for Exercise 6.4):
Reject H_0 (conclude the coin is altered) if heads appears 11 or 12 times. Retain H_0 (conclude the coin is fair) otherwise.

(a) Finding α (Type I Error Probability) A Type I error occurs when we reject H_0 when it is actually true, meaning we conclude the coin is altered when it is actually fair.

The number of heads in 12 tosses follows a binomial distribution:

$$X \sim \text{Binomial}(n = 12, p = 0.5)$$

The probability of a Type I error is:

$$\alpha = P(X = 11 \text{ or } X = 12 \mid H_0)$$

Using the binomial probability formula:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

We compute:

$$\alpha = P(X = 11|p = 0.5) + P(X = 12|p = 0.5)$$

Numerically, we find:

$$\alpha \approx 0.0032 \quad (0.32\%)$$

Interpretation: There is a 0.32% chance of wrongly concluding that the coin is biased when it is actually fair.

(b) Finding β (Type II Error Probability) A Type II error occurs when we fail to reject H_0 when H_a is actually true, meaning we conclude the coin is fair when it is actually biased.

If the coin is biased, the number of heads follows:

$$X \sim \text{Binomial}(n = 12, p = 0.75)$$

The probability of a Type II error is:

$$\beta = P(X \leq 10 \mid H_a)$$

This means we sum the probabilities for $X = 0, 1, ..., 10$ under $p = 0.75$:

$$\beta = \sum_{k=0}^{10} P(X = k \mid p = 0.75)$$

Numerically, we find:

$$\beta \approx 0.8416 \quad (84.16\%)$$

Interpretation: There is an 84.16% chance of failing to detect that the coin is biased when it actually is.

α (Type I Error Probability):

$$\alpha = P(X = 11 \text{ or } X = 12 \mid H_0) \approx 0.0032 \quad (0.32\%)$$

Interpretation: There is a 0.32% chance of wrongly concluding the coin is biased when it is actually fair.

β (Type II Error Probability):

$$\beta = P(X \leq 10 \mid H_a) \approx 0.8416 \quad (84.16\%)$$

Interpretation: The test has a high probability of failing to detect a biased coin.