Homework 1

Student Name: Vo, Huynh Quang Nguyen

CWID: A20446163

Problem 2.5.3

The following table summarizes the number of deceased beetles under autolysis (the destruction of a cell after its death by the action of its own enzymes) and putrefaction (decomposition of organic matter, especially protein, by microorganisms, resulting in production of foul-smelling matter):

	Autolysis-High	Autolysis-Low
Putrefaction-High	14	59
Putrefaction-Low	18	9

a. If the autolysis of a sample is high, what is the probability that the putrefaction is low?

Answer:

The total number of deceased beetles is:

$$\text{Total} = 14 + 59 + 18 + 9 = 100$$

The number of deceased beetles whose autolysis is high is:

$$\text{Autolysis-High} = 14 + 18 = 32$$

The number of deceased beetles whose putrefaction is low is:

$$\text{Putrefaction-Low} = 18 + 9 = 27$$

Therefore, the probability that the putrefaction is low given the autolysis is high can be calculated as:

$$P(\text{Putrefaction-Low}|\text{Autolysis-High}) = \frac{P(\text{Putrefaction-Low, Autolysis-High})}{P(\text{Autolysis-High})} = \frac{\frac{18}{100}}{\frac{32}{100}} = \frac{18}{32} = \frac{9}{16} = 0.5625 = 56.25\%$$

b. If the putrefaction of a sample is high, what is the probability that the autolysis is high?

Answer:

The number of deceased beetles whose putrefaction is high is:

$$\text{Putrefaction-High} = 14 + 59 = 73$$

The probability that the autolysis is high given the putrefaction is high can be calculated as:

$$P(\text{Autolysis-High}|\text{Putrefaction-High}) = \frac{P(\text{Autolysis-High, Putrefaction-High})}{P(\text{Putrefaction-High})} = \frac{\frac{14}{100}}{\frac{73}{100}} = \frac{14}{73} \approx 0.1918 = 19.18\%$$

c. If the putrefaction of a sample is low, what is the probability that the autolysis is low?

Answer:

The probability that the autolysis is low given the putrefaction is low can be calculated as:

$$P(\text{Autolysis-Low}|\text{Putrefaction-Low}) = \frac{P(\text{Autolysis-Low}, \text{Putrefaction-Low})}{P(\text{Putrefaction-Low})} = \frac{\frac{9}{100}}{\frac{27}{100}} = \frac{9}{27} = \frac{1}{3} \approx 0.3333 = 33.33\%$$

Problem 2.6.1

Suppose that $P(A|B) = 0.4$ and $P(B) = 0.5$. Determine the following:

a. $P(A \cap B)$

Answer:

According to the multiplication rule, $P(A \cap B)$ can be computed as follows:

$$P(A \cap B) = P(A|B) \times P(B) = 0.4 \times 0.5 = 0.2$$

b. $P(A' \cap B)$

Answer:

According to the multiplication rule, $P(A' \cap B)$ can be calculated as follows:

$$P(A' \cap B) = P(A'|B) \times P(B)$$

Since $P(A'|B)$ and $P(A|B)$ are complementary events, $P(A'|B)$ is calculated as:

$$P(A'|B) = 1 - P(A|B) = 1 - 0.4 = 0.6$$

Therefore, $P(A' \cap B)$ is calculated as:

$$P(A' \cap B) = 0.6 \times 0.5 = 0.3$$

Problem 2.7.4

Disks of polycarbonate plastic from a supplier are analyzed for scratch and shock resistance. The results from 100 disks are summarized as follows:

	Shock-High	Shock-Low
Scratch-High	70	9
Scratch-Low	16	5

Let A denote the event that a disk has high shock resistance, and let B denote the event that a disk has high scratch resistance. Are events A and B independent?

Answer:

To determine whether these two events are independent, we have to prove that:

$$P(A|B) = P(A) \quad \cup \quad P(B|A) = P(B)$$

In this scenario, $P(A)$ and $P(B)$ are calculated as follows:

$$P(A) = \frac{70 + 16}{100} = \frac{86}{100}$$
$$P(B) = \frac{70 + 9}{100} = \frac{79}{100}$$

Meanwhile, $P(A|B)$ and $P(B|A)$ are calculated as follows:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{70}{100}}{\frac{79}{100}} = \frac{70}{79}$$
$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{70}{100}}{\frac{86}{100}} = \frac{70}{86}$$

Since $P(A|B) \neq P(A)$ and $P(B|A) \neq P(B)$, these two events are dependent to each other.

Problem 2.8.4

Software to detect fraud in consumer phone cards tracks the number of metropolitan areas where calls originate each day. It is found that 1% of the legitimate users originate calls from two or more metropolitan areas in a single day. However, 30% of fraudulent users originate calls from two or more metropolitan areas in a single day. The proportion of fraudulent users is 0.01%. If the same user originates calls from two or more metropolitan areas in a single day, what is the probability that the user is fraudulent?

Answer:

Let F be the event that a user is fraudulent, and let M be the event that a user originates calls from two or more metropolitan areas in a single day. Hence, we define that:

- $P(F|M)$: the probability that a user is fraudulent given they origniate calls from two or more metropolitan areas in a single day.
- $P(M|F)$: the probability that a fraudulent user originates calls from two or more metropolitan areas in a single day. In this scenario, 30% of fraudulent users originate calls from two or more metropolitan areas in a single day means that $P(M|F) = 0.3$.
- $P(F)$: the total proportion of fraudulent users. In this scenario, the proportion of fraudulent users is 0.01% means that $P(F) = 0.0001$.
- $P(M)$: the total proportion of users from two or more metropolitan areas.
- $P(M|F')$: the probability that a user is legitimate given they origniate calls from two or more metropolitan areas in a single day. In this case, 1% of the legitimate users originate calls from two or more metropolitan areas in a single day means that $P(M|F') = 0.01$.

Using the Bayes' Rule, $P(F|M)$ can be calculated as:

$$P(F|M) = \frac{P(M|F) \times P(F)}{P(M)}$$

To calculate $P(M)$, we can use the law of total probability:

$$P(M) = P(M|F) \times P(F) + P(M|F') \times P(F') = 0.3 \times 0.0001 + 0.01 \times (1 - 0.0001) \approx 0.01003$$

Therefore, $P(F|M) = 0.3 \times 0.0001/0.01003 \approx 0.00299 = 0.299$

Problem 3.1.10

verify that the following functions are probability mass functions, and determine the requested probabilities. $f(x) = \frac{8}{7}(\frac{1}{2})^x$ given $x = 1, 2, 3$.

Answer:

To verify whether $f(x)$ is a probability mass function, we must verify that:

- $f(x = x_i) \geq 0 \quad \forall x_i = \{1, 2, 3\}$.
- $\sum_{x_i} f(x_i) = 1 \quad \forall x_i = \{1, 2, 3\}$.

In this scenario:

$$f(x=1)=\frac{8}{7}(\frac{1}{2})^1=\frac{4}{7}$$
$$f(x=2)=\frac{8}{7}(\frac{1}{2})^2=\frac{2}{7}$$
$$f(x=3)=\frac{8}{7}(\frac{1}{2})^3=\frac{1}{7}$$

Since $f(x=1)\geq 0, f(x=2)\geq 0, f(x=3)\geq 0$, and $f(x=1)+f(x=2)+f(x=3)=1$, we can conclude that $f(x)$ is a probability mass function.

(a) $P(X\leq 1)=\frac{8}{7}(\frac{1}{2})^1=\frac{4}{7}$.

(b) $P(X>1)=1-P(X\leq 1)=1-\frac{4}{7}=\frac{3}{7}$.

(c) $P(2<X<6)=P(X=3)=\frac{1}{7}$

(d) $P(X\leq 1 \text{ or } X>1)=\frac{8}{7}(\frac{1}{2})^1+\frac{8}{7}(\frac{1}{2})^2+\frac{8}{7}(\frac{1}{2})^3=1$

Problem 3.1.17

An assembly consists of three mechanical components. Suppose that the probabilities that the first, second, and third components meet specifications are 0.95, 0.98, and 0.99, respectively. Assume that the components are independent. Determine the probability mass function of the number of components in the assembly that meet specifications.

Answer:

We have the following table:

Component	Meet Specifications	Does Not Meet Specifications
1st	0.95	1 - 0.95 = 0.05
2nd	0.98	1 - 0.98 = 0.02
3rd	0.99	1 - 0.99 = 0.01

Therefore, the probability mass function of the number of components in the assembly that meet specifications is defined as:

$$\begin{cases} f(X=0)=0.05\times 0.02\times 0.01=0.00001 \quad [1] \\ f(X=1)=0.95\times 0.02\times 0.01+0.98\times 0.05\times 0.01+0.99\times 0.05\times 0.02=0.00167 \quad [2] \\ f(X=2)=0.95\times 0.98\times 0.01+0.95\times 0.02\times 0.99+0.05\times 0.98\times 0.99=0.07663 \quad [3] \\ f(X=3)=0.95\times 0.98\times 0.99=0.92169 \quad [4] \end{cases}$$

where:

- [1]: in this scenario, all components do not meet the specifications.
- [2]: in this scenario, one out of three components meets the specification - either the 1st, the 2nd, or the 3rd component.
- [3]: in this scenario, two out of three components meet the specification - either the 1st and the 2nd, the 1st and the 3rd, or the 2nd and the 3rd components.
- [4]: in this scenario, all components meet the specification.

Problem 3.4.4

Thickness measurements of a coating process are made to the nearest hundredth of a millimeter. The thickness measurements are uniformly distributed with values 0.15, 0.16, 0.17, 0.18, and 0.19. Determine the mean and variance of the coating thickness for this process.

Answer:

Denote X as a random variable that describes the thickness measurments. Since the thickness measurements are uniformly distributed, the probability associating with each measurement will be $1/N$ with N is the number of measurments. Hence, the mean of coating is computed as follows:

$$\mathbb{E}[X] = \frac{1}{N}\sum_i^N p_i x_i = \frac{0.15 \times 0.2 + 0.16 \times 0.2 + 0.17 \times 0.2 + 0.18 \times 0.2 + 0.19 \times 0.2}{5} = 0.34$$

The variance is computed as follows: $$ \mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = (\frac{0.15^2\times0.2 + 0.16^2\times0.2 + 0.17^2\times0.2 + 0.18^2\times0.2 + 0.19^2\times0.2}{5}) - 0.34^2 = 0.1034

Problem 3.5.6

An electronic product contains 40 integrated circuits. The probability that any integrated circuit is defective is 0.01, and the integrated circuits are independent. The product operates only if there are no defective integrated circuits. What is the probability that the product operates?

Answer: Since all integrated circuits are independent from one to another, the probability that the product operates can be computed as follows:

$$P(\text{operate}) = \Big(1 - P(\text{defective})\Big)^{40} = (1 - 0.01)^{40} \approx 0.669$$

Problem 3.6.3

Consider a sequence of independent Bernoulli trials with $p = 0.2$.

a. What is the expected number of trials to obtain the first success?

b. After the eighth success occurs, what is the expected number of trials to obtain the ninth success?

Answer:

a. Assuming that the probability for a trial to be succeeded is $p = 0.2$. We know the number of trials that are succeeded can be found using the geometric distribution:

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

with k is the number of trials up to the first success. Hence, the expected number of trials to obtain the first success $E(X)$ is given as:

$$E(X) = \frac{1}{p} = \frac{1}{0.2} = 5$$

b. The expected number of trials to obtain the ninth success after the eighth success occurs is also a geometric distribution problem. Since the trials are independent, the probability of success in each trial remains the same ($p = 0.2$). Hence, the expected number of trials to obtain the ninth success will be 5.

Problem 3.6.7

Heart failure is due to either natural occurrences (87%) or outside factors (13%). Outside factors are related to induced substances or foreign objects. Natural occurrences are caused by arterial blockage, disease, and infection. Assume that causes of heart failure for the individuals are independent.

a. What is the probability that the first patient with heart failure who enters the emergency room has the condition due to outside factors?

b. What is the probability that the third patient with heart failure who enters the emergency room is the first one due to outside factors?

c. What is the mean number of heart failure patients with the condition due to natural causes who enter the emergency room before the first patient with heart failure from outside factors?

Answer:

a. The probability that the first patient with heart failure who enters the emergency room has the condition due to outside factors:

$$P(A) = P(\text{outside factors}) = 0.13$$

b. By saying the third patient with heart failure who enters the emergency room is the first one due to outside factors, we mean that the first two patients have a heart attack caused by natural occurrences. Hence, this probability can be computed as:

$$P(B) = P(\text{natural occurences})^2 \times P(\text{outside factors}) = 0.87^2 \times 0.13 = 0.0984$$

c. We can consider that the number of patients with the condition due to outide factors has a geometric distribution:

$$P(X = k) = p(1 - p)^k$$

with $p = 0.13$, and k is the number of patients with heart failure up to the first one is caused by outside. Hence, the mean number is:

$$E(X) = \frac{1}{p} = 1/0.13 \approx 7.69$$

Problem 3.8.3

Suppose that the number of customers who enter a store in an hour is a Poisson random variable, and suppose that $P(X = 0) = 0.05$. Determine the mean and variance of X.

Answer:

The Poisson random variable - the number of customers who enter a store in an hour in this scenario - is generally described as:

$$f(x) = \frac{e^{-\lambda T}(\lambda T)^x}{x!}$$

In this scenario, $P(X = 0) = 0.05$. Hence,

$$P(X = 0) = \frac{e^{-\lambda T}(\lambda T)^0}{0!} = 0.05 \to (\lambda T) \approx 2.9957$$

Therefore,

- Mean: $E(X) = \lambda T = 2.9957$
- Variance: $V(X) = E(X) = 2.9957$

Problem 3.8.5

Astronomers treat the number of stars in a given volume of space as a Poisson random variable. The density in the Milky Way Galaxy in the vicinity of our solar system is one star per 16 cubic light-years.

a. What is the probability of two or more stars in 16 cubic light-years?

b. How many cubic light-years of space must be studied so that the probability of one or more stars exceeds 0.95?

Answer:

The Poisson random variable - the number of stars in a given volume of space in this scenario - is generally described as:

$$f(x) = \frac{e^{-\lambda T}(\lambda T)^x}{x!}$$

Given that the density is one star per 16 cubic light years, we can agree that $\lambda T = \frac{1}{16}$.

a. The probability of two or more stars in 16 cubic light-years can be computed as follows:

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - \frac{e^{-1}(1)^0}{0!} - \frac{e^{-1}(1)^1}{1!} \approx 0.2642$$

b. In this case, we need to find the λT since our volume under interest is now different. The new λT can be computed as follows:

$$P(X \geq 1) = 0.95 \rightarrow 1 - \frac{e^{-\lambda_1 T_1}(\lambda_1 T_1)^0}{0!} = 0.95 \rightarrow \lambda_1 T_1 \approx 3$$

Since $\lambda_0 T_0 = \lambda_1 T_1$, the new volume (in cubic light-years) is computed as: $T_1 = (16 \times 3)/1 = 48$

Problem 3.8.9

The number of surface flaws in plastic panels used in the interior of automobiles has a Poisson distribution with a mean of 0.05 flaw per square foot of plastic panel. Assume that an automobile interior contains 10 square feet of plastic panel.

a. What is the probability that there are no surface flaws in an auto’s interior?

b. If 10 cars are sold to a rental company, what is the probability that none of the 10 cars has any surface flaws?

c. If 10 cars are sold to a rental company, what is the probability that at most 1 car has any surface flaws?

Answer:

The Poisson random variable - in this scenario, the number of surface flaws in plastic panels - has the following distribution:

$$f(x) = \frac{e^{-\lambda T}(\lambda T)^x}{x!}$$

Since the mean (λ) is 0.05 per square foot and the area under interest is 10 square feet, we can agree that $\lambda T = \frac{0.05}{1} \times 10 = 0.5$.

a. The probability that there are no surface flaws in an auto’s interior can be computed as follows:

$$P(X = 0) = \frac{e^{-0.5}(0.5)^0}{0!} \approx 0.606531$$

b. We assume that the number of surface flaws in a car is an independent event, then the probability that none of the 10 cars sold to a rental company has any surface flaws can be computed as:

$$P(\text{no car has flaws}) = P(X = 0)^{10} = 0.606531^{10} \approx 0.006737$$

c. The probability that at most 1 car amomg 10 cars sold to a rental company has any surface flaws can be computed as:

$$P(\text{at most 1 car has flaws}) = P(\text{no car has flaws}) + P(\text{only one car has flaws}) = 0.006737$$
$$+ \binom{10}{1}(1 - 0.006737)(0.006737)^9 \approx 0.05045$$