Probability

Madiba Hudson-Quansah

Contents

Chapter 1

Module 8: Introduction

1.1 Introduction To Probability

Definition 1.1.1: Probability

A mathematical description of randomness and uncertainty / The likelihood of an event occurring.
The notation for Probability is $\mathbb{P}(X)$ where X is the event.
Probability is always between $0 \leqslant \mathbb{P}(X) \leqslant 1$ or $0\% \leqslant \mathbb{P}(X) \leqslant 100\%$.

There are two ways of determining probability:

- Theoretical / Classical - Determined by the nature of the experiment
- Empirical / Observational - Determined by the results of the experiment

1.2 Relative Frequency

Definition 1.2.1: Relative Frequency

Relative frequency is the number of times an event occurs divided by the total number of trials.

$$\mathbb{P}(X) = \frac{\text{Number of times event occurs}}{\text{Total number of trials}}$$

Theorem 1.2.1 The Law of Large Numbers

As the number of trials increases, the relative frequency of an event approaches the theoretical probability of the event.

Chapter 2

Module 9: Find the Probability of Events

2.1 Sample Spaces and Events

Definition 2.1.1: Random Experiment

An experiment whose outcome is determined by chance.

Definition 2.1.2: Sample Space

The list of possible outcomes of a random experiment, denoted by S.

Definition 2.1.3: Event

A statement about the nature of the outcome after the experiment has been conducted, denoted by any capital letter except S.

2.2 Equally Likely Outcomes

$$\mathbb{P}(A) = \frac{\text{Number of outcomes in } A}{\text{Number of outcomes in } S}$$

Where A is an event and S is the sample space.

2.3 Probability Rules

2.3.1 Rule 1: Probability is a Number Between 0 and 1

For any event A, $0 \leq \mathbb{P}(A) \leq 1$.

2.3.2 Rule 2: Addition Rule

$\mathbb{P}(S) = 1$, that is the sum of the probabilities of all possible outcomes is 1.

2.3.3 Rule 3: Complement Rule

$\mathbb{P}(A') = 1 - \mathbb{P}(A)$, that is the probability of the complement of an event is 1 minus the probability the event occurs.

2.3.4 Rule 4: Addition Rule for Mutually Exclusive Events

Definition 2.3.1: Mutually Exclusive / Disjoint events

Events that cannot happen at the same time.

$\mathbb{P}(A \text{ or } B) = \mathbb{P}(\text{ event } A \text{ occurs or event } B \text{ occurs or both occur})$

If A and B are mutually exclusive, then $\mathbb{P}(A \text{ or } B) = \mathbb{P}(A) + \mathbb{P}(B)$

2.3.5 Rule 5: Multiplication Rule for Independent Events

$\mathbb{P}(A \text{ and } B) = \mathbb{P}(\text{ event } A \text{ occurs and event } B \text{ occurs })$

Definition 2.3.2: Independent Events

Two events A and B are said to be independent if the occurrence of one event does not affect the probability of the other event occurring.

Definition 2.3.3: Dependent Events

Two events A and B are said to be dependent if the occurrence of one event affects the probability of the other event occurring.

If A and B are two independent events, then $\mathbb{P}(A \text{ and } B) = \mathbb{P}(A) \times \mathbb{P}(B)$

2.3.6 Rule 6: General Addition Rule

For any two events A and B, $\mathbb{P}(A \text{ or } B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \text{ and } B)$. If the event are mutually exclusive, then $\mathbb{P}(A \text{ and } B) = 0$, giving us $\mathbb{P}(A \text{ or } B) = \mathbb{P}(A) + \mathbb{P}(B)$, i.e. the addition rule for mutually exclusive events.

Chapter 3

Module 10: Conditional Probability and Independence

Definition 3.0.1: Conditional Probability

The probability an event occurs as a result of another event. I.e. Probability of event B, given event event A is,

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \text{ and } B)}{P(A)}$$

3.1 Independence

When two events are independent, the probability of one event occurring does not affect the probability of the other event, i.e.

$$\begin{aligned} \mathbb{P}(B \mid A) &= \mathbb{P}(B) \\ \mathbb{P}(A \mid B) &= \mathbb{P}(A) \\ \mathbb{P}(B \mid A) &= \mathbb{P}(B \mid A') \\ \mathbb{P}(A \text{ and } B) &= \mathbb{P}(A) \times \mathbb{P}(B) \end{aligned}$$

3.2 The General Multiplication Rule

For any two dependent events A and B

$$\mathbb{P}(A \text{ and } B) = \mathbb{P}(A) \times \mathbb{P}(B \mid A)$$

3.3 Probability Trees

Definition 3.3.1: Probability Tree

A diagram that shows the sample space of a random experiment and the probability of each outcome.

3.3.1 Bayes' Theorem

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A) \times P(B \mid A)}{\mathbb{P}(A) \times \mathbb{P}(B \mid A) + \mathbb{P}(A') \times \mathbb{P}(B \mid A')}$$

Chapter 4

Module 11: Random Variables

Definition 4.0.1: Random Variable

Assigns a unique numerical value to the outcome of a random experiment.

Definition 4.0.2: Discrete Random Variable

A random variable that can take on a finite number of values. Discrete random variables are usually counts.

Definition 4.0.3: Continuous Random Variable

A random variable that can take on an infinite number of values. Continuous random variables are usually measurements.

4.1 Discrete Random Variables

4.1.1 Notation

For a given event X, the probability of X is denoted by $\mathbb{P}(X)$. For a given value x, the probability of X is denoted by $\mathbb{P}(X = x)$, i.e. the probability that X takes on the value x.

4.1.2 Probability Distribution

Definition 4.1.1: Probability Distribution

The list of all possible values of a random variable and their corresponding probabilities.

Any probability distribution must satisfy the following two conditions:

- $0 \leqslant \mathbb{P}(X = x) \leqslant 1$ - The probability of any value of X is between 0 and 1.
- $\Sigma_x \mathbb{P}(X = x) = 1$ - The sum of the probabilities of all possible values of X is 1.

4.1.3 Key Words

- At least / No less than - $x \geqslant$
- At most / No more than - $x \leqslant$
- Less than / fewer than - $x <$
- More than / greater than - $x >$

- Exactly - $x =$

4.1.4 Mean and Variance of a Discrete Random Variable

4.1.4.1 Mean

Definition 4.1.2: Mean / Expected value of a Discrete Random Variable

The average value of a random variable, denoted by μ.

For a given random variable X, the mean is given by

$$\mu_X = \Sigma_{i=1}^{n} x_i p_i$$

Where x_i is the value of X and p_i is the probability of X taking on the value x_i.

4.1.4.1.1 Applications of the Mean

- The mean of a random variable is the long-term average value of the random variable.
- The mean of a random variable is the centre of the probability distribution of the random variable.

4.1.4.2 Variance

Definition 4.1.3: Variance

The average of the squared differences between each value of a random variable and the mean of the random variable, denoted by σ^2.

For a given random variable X, the variance is given by

$$\sigma_X^2 = \Sigma_{i=1}^{n} \left(x_i - \mu_X\right)^2 p_i$$

And standard deviation is given by

$$\sigma_X = \sqrt{\sigma_X^2}$$

Where x_i is the value of X and p_i is the probability of X taking on the value x_i.

4.1.4.3 Rules for Mean and Variance of Random Discrete Variables

4.1.4.3.1 Adding or Subtracting a Constant to a Random Variable

If $Y = X + c$, then $\mu_Y = \mu_X + c$, $\sigma_Y^2 = \sigma_X^2$ and $\sigma_Y = \sigma_X$.

If $Y = X - c$, then $\mu_Y = \mu_X - c$, $\sigma_Y^2 = \sigma_X^2$ and $\sigma_Y = \sigma_X$.

4.1.4.3.2 Multiplying a Random Variable by a Constant > 1

If $Y = cX, c > 1$, then $\mu_Y = c\mu_X$, $\sigma_Y^2 = c^2\sigma_X^2$ and $\sigma_Y = c\sigma_X$

4.1.4.3.3 Multiplying a Random Variable by a Constant < 1

If $Y = cX, c < 1$, then $\mu_Y = c\mu_X$, $\sigma_Y^2 = c^2\sigma_X^2$ and $\sigma_Y = c\sigma_X$

4.1.4.3.4 Linear Transformation of a Random Variable

If $Y = a + bX$, then $\mu_Y = a + b\mu_X$, $\sigma_Y^2 = b^2\sigma_X^2$ and $\sigma_Y = |b|\sigma_X$

4.1.4.3.5 Sum of Two Random Variables

If $Z = X + Y$, then $\mu_Z = \mu_X + \mu_Y$, $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$ and $\sigma_Z = \sqrt{\sigma_X^2 + \sigma_Y^2}$. Only if X and Y are independent.

4.1.5 Poisson Random Variables

Definition 4.1.4: Poisson Random Variable

A random variable that counts the number of events that occur in a fixed interval of time or space, denoted by $X \sim \text{Poisson}(\lambda)$. Where λ is the average number of events that occur in the interval.

Definition 4.1.5: Poisson Experiment

Random experiments that satisfy the following conditions:
- The number of trials tends to infinity.
- The probability of success tends to zero.
- $np = 1$ is finite

$$\mathbb{P}(X = x) = \frac{\left(e^{-\lambda} \times \lambda^x\right)}{x!}$$

Where e is the base of the natural logarithm, λ is the average number of events that occur in the interval and x is the number of events that occur in the interval.

If X is Poisson with parameter λ, then

$$\mu_X = \lambda$$

And

$$\sigma_X^2 = \mu = \lambda$$
$$\sigma_X = \sqrt{\sigma^2}$$

4.1.6 Binomial Random Variables

Definition 4.1.6: Binomial Random Variable

A random variable that counts the number of successes in a fixed number of independent trials, denoted by $X \sim \text{Bin}\left(n, p\right)$. Where n is the number of trials and p is the probability of success.

Definition 4.1.7: Binomial Experiment

Random experiments that satisfy the following conditions:

- A fixed number of trials, denoted by n.
- Each trial is independent of the others.
- There are only two possible outcomes for each trial, success or failure.
- There is a constant probability of success, denoted by p, for each trial, which can be expressed as the complement of the probability of failure, $q = 1 - p$.

Note:-

The number (X) of success in a sample of size n taken without replacement from a population with proportion p of successes is approximately binomial with n and p as long as the sample size is at most 10% of the population size (N). I.e.

$$n \leqslant 0.1N$$

Or

$$N \geqslant 10n$$

To calculate the probability of a binomial random variable, we use the formula

$$\mathbb{P}(X = x) = \binom{n}{x} p^x q^{n-x}, \text{ where } x = 0, 1, 2, \ldots, n$$

Where n is the number of trials, x is the number of successes, p is the probability of success and q is the probability of failure.

If X is Binomial with parameters n and p, then

$$\mu_X = np$$

And

$$\sigma_X^2 = np(1-p)$$
$$\sigma_X = \sqrt{np(1-p)}$$

4.2 Continuous Random Variables

4.2.1 Probability Distribution

For a continuous random variable X, the probability distribution is given by the *probability density function*, whose properties are

- $f(x) \geqslant 0$ for all x.
- $\int_{-\infty}^{\infty} f(x)\, dx = 1$
- The probability that X takes on a value between a and b is given by

$$\mathbb{P}(a \leqslant X \leqslant b) = \int_a^b f(x)\, dx$$

Note:-

- The probability that a continuous random variable takes on a specific value is always 0.
- The strictness of the inequality does not matter, i.e. $\mathbb{P}(X \geqslant a) = \mathbb{P}(X > a)$

4.2.2 Normal Random Variables

Definition 4.2.1: Normal Random Variable

A random variable that has a bell-shaped probability distribution, denoted by $X \sim N\left(\mu, \sigma^2\right)$. Where μ is the mean and σ^2 is the variance.

For a normally distributed random variable X:

- There is a 68% chance that X takes on a value within one standard deviation of the mean, i.e. $0.68 = \mathbb{P}\left(\mu - \sigma < X < \mu + \sigma\right)$
- There is a 95% chance that X takes on a value within two standard deviations of the mean, i.e. $0.95 = \mathbb{P}\left(\mu - 2\sigma < X < \mu + 2\sigma\right)$
- There is a 99.7% chance that X takes on a value within three standard deviations of the mean, i.e. $0.997 = \mathbb{P}\left(\mu - 3\sigma < X < \mu + 3\sigma\right)$

4.2.2.1 Finding Probabilities for Normal Random Variables

4.2.2.1.1 Standardizing Values

Definition 4.2.2: z-score

The number of standard deviations a value is from the mean of a normal random variable, denoted by z.

To standardize a normal random variable X, we must find its z-score, given by

$$z = \frac{x - \mu}{\sigma}$$

4.2.2.1.2 Finding Probabilities with the z-score

Definition 4.2.3: Normal Table

A table that shows the probability that a standard normal random variable takes on a value less than a given z-score.

Using the z-score we can find the probability that a normal random variable takes on a value less than a given value x, by tracing the z-score to the normal table.

$$\mathbb{P}(X < x) = \mathbb{P}(Z < z)$$

On a standard normal table z-score are written to two decimal places as row headers and for additional precision the column headers are the first two decimal places of the z-score.

4.2.3 Uniform Distribution

Definition 4.2.4: Uniform Distribution

,denoted by $X \sim U(a, b)$

For a random variable X, if is uniformly distributed over the interval a and b then its *probability distribution density function* is given by

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise} \end{cases}$$

The mean, variance, and standard deviation of a uniformly distributed random variable is given by

$$\mu_X = \frac{a+b}{2}$$
$$\sigma_X^2 = \frac{(b-a)^2}{12}$$
$$\sigma_X = \sqrt{\frac{(b-a)^2}{12}}$$

Chapter 5

Module 12: Sampling Distributions

5.1 Parameters vs. Statistics

Definition 5.1.1: Parameter

A numerical value that describes a characteristic of a population, denoted by a Greek letter, e.g. μ, σ^2.

Definition 5.1.2: Statistic

A numerical value that describes a characteristic of a sample, denoted by a Roman letter, e.g $\bar{x}, s^2$.

Definition 5.1.3: Proportion

A statistic that estimates the proportion of a population or sample that has a certain characteristic, denoted by p for a population and $\hat{p}$ for a sample.

Definition 5.1.4: Sampling Variability

The variability of a statistic from one sample to another.

5.2 Behaviour of Sample Proportion $\hat{p}$

5.2.1 Centre

The mean of the sample proportion is the same as the population proportion, i.e.

$$\mu_{\hat{p}} = p$$

As it is reasonable to expect all the sample proportions in repeated samples to average out to the underlying population.

5.2.2 Spread

The sample size has an effect on the spread of the distribution of the sample proportion, i.e. the **larger the sample size, the less spread out the distribution** of the sample proportion and **more spread for smaller sample sizes**. We can describe the spread of the distribution of the sample proportion more precisely by finding the actual standard deviation of the sample proportion. i.e.

$$\sigma_{\hat{p}} = \sqrt{\frac{p\,(1-p)}{n}}$$

Where p is the population proportion and n is the sample size.

5.2.3 Shape

The shape of the distribution of the sample proportion is approximately normal if the sample size is large enough. I.e. if

$$np \geqslant 10 \text{ and } n\left(1-p\right) \geqslant 10$$

Definition 5.2.1: Sampling of Distribution of $\hat{p}$

The distribution of the values of the sample proportions $\hat{p}$ in repeated samples.

5.3 Behaviour of Sample Mean $\overline{X}$

5.3.1 Centre

The mean of the sample mean is the same as the population mean, i.e.

$$\mu_{\overline{X}} = \mu$$

5.3.2 Spread

The sample size has an effect on the spread of the distribution of the sample mean, i.e. the **larger the sample size, the less spread out the distribution** of the sample mean and **more spread for smaller sample sizes**. We can describe the spread of the distribution of the sample mean more precisely by finding the actual standard deviation of the sample mean. i.e.

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

5.3.3 Shape

The shape of the distribution of the sample mean is approximately normal if the sample size is large enough. I.e. if

$$n \geqslant 30$$

Definition 5.3.1: Sampling of Distribution of $\overline{X}$

The distribution of the values of the sample mean $\overline{X}$ in repeated samples.

Chapter 6

Exercises

Question 1

Three cards are drawn with replacement from a well-shuffled deck of 52 cards. Find the probability distribution of the number of aces drawn. Also, find the mean and variance of the distribution.

Solution:

x	$P(x)$
0	$\binom{3}{0}\left(\frac{4}{52}\right)^{0}\left(1-\frac{4}{52}\right)^{3-0}=0.7865$
1	$\binom{3}{1}\left(\frac{4}{52}\right)^{1}\left(1-\frac{4}{52}\right)^{3-1}=0.1966$
2	$\binom{3}{2}\left(\frac{4}{52}\right)^{2}\left(1-\frac{4}{52}\right)^{3-2}=0.0164$
3	$\binom{3}{3}\left(\frac{4}{52}\right)^{3}\left(1-\frac{4}{52}\right)^{3-3}=0.0005$

$$\begin{aligned}
\mu &= \sum_{i=1}^{4} x_i \times p_i \\
&= (0\times 0.7865)+(1\times 0.1966)+(2\times 0.0164)+(3\times 0.0005) \\
&= 0.2309 \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\sigma^2 &= \sum_{i=1}^{4} \left(x_i-\mu\right)^2 p_i \\
&= \left((0-0.2309)^2\times 0.7865\right)+\left((1-0.2309)^2\times 0.1966\right)+\left((2-0.2309)^2\times 0.0164\right)+\left((3-0.2309)^2\times 0.0005\right) \\
&= 0.21338519 \\
&= 0.2134
\end{aligned}$$

Question 2

Paper clips are produced in a variety of colours The proportion of red paper clips produced is 0.20, Determine the probability that, in a random sample of 50 coloured paper clips, the number of red clips is:

1. Fewer than 10
2. At least 8 but at most 12

Solution:

$$X \sim B\left(50, 0.20\right)$$

1.

$$\begin{aligned} P\left(X < 10\right) &= \sum_{i=0}^{9} \binom{50}{i} (0.2)^i \left(1 - 0.20\right)^{50-i} \\ &= 0.4437 \end{aligned}$$

2.

$$\begin{aligned} P\left(8 \leqslant X \leqslant 12\right) &= P\left(X \leqslant 12\right) - P\left(X \leqslant 8\right) \\ &= \sum_{i=8}^{12} \binom{50}{i} (0.20)^i \left(1 - 0.20\right)^{50-i} \\ &= 0.6235 \end{aligned}$$

Question 3

A recent large-scale survey established that 15 percent of cars have fully functioning brake lights

1. Calculate the probability that, in a random sample of 18 cars, exactly 2 cars have faulty brake lights.
2. Determine the probability that, in a random sample of 50 cars, more than 5 cars but fewer than 10 cars have faulty brake lights.

Solution:

1.

$$\begin{aligned} &X \sim B\left(18, 0.15\right) \\ &P\left(X = 2\right) = \binom{18}{2} (0.15)^2 \left(1 - 0.15\right)^{18-2} \\ &\qquad = 0.2556 \end{aligned}$$

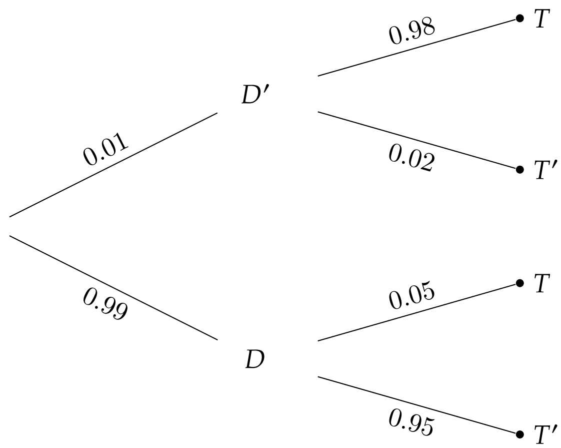
2.

$$\begin{aligned} &X \sim B\left(50, 0.20\right) \\ &P\left(5 < X < 10\right) = P\left(X < 10\right) - P\left(X < 5\right) \\ &\qquad = \sum_{i=6}^{9} \binom{50}{i} (0.15)^i \left(1 - 0.15\right)^{50-i} \\ &\qquad = 0.5717 \end{aligned}$$

Question 4

You are diagnosed with an uncommon disease. You know that there only is a 1% chance. Use the letter D for the event "you have a disease" and T for "the test says so". It is known that the test is perfect. $P(T \mid D) = 0.98$ and $P(T' \mid D') = 0.95$

1. Given that you test positive, what is the probability that you really have the disease?
2. You obtain a second opinion: in an independent repetition of the test. You test positive again. Given this, what is the probability that you really have the disease.



Solution:

1.

$$
\begin{aligned}
P(D \mid T) &= \frac{P(D \cap T)}{P(T)} \\
&= \frac{P(D) \times P(T \mid D)}{P(D) \times P(T \mid D) + P(D') \times P(T \mid D')} \\
&= \frac{0.01 \times 0.98}{(0.01 \times 0.98) + (0.99 \times 0.05)} \\
&= 0.1653
\end{aligned}
$$

2.

$$
\begin{aligned}
P((D \mid T) \cap (D \mid T)) &= 0.1653 \times 0.1653 \\
&= 0.0273
\end{aligned}
$$

Question 5

Selorm, arriving at a bus stop, just misses the bus. Suppose that he decides to walk if the (next) bus takes longer than 5 minutes to arrive. Suppose also that the time in minutes between the arrivals of buses at the bus stop is a continuous random variable with a $U(4, 6)$. Let X be the time Selorm will wait.

1. What is the probability that X is less than $4\frac{1}{2}$ minutes
2. What is the probability that X equals 5 minutes?
3. Is X a discrete random variable or a continuous random variable?

Solution:

1.

$$
\begin{aligned}
f(x) &= \frac{1}{6-4} \\
f(x) &= 0.5 \\
P\left(X < \frac{9}{2}\right) &= (4.5-4) \times 0.5 \\
&= 0.25
\end{aligned}
$$

2.

$$P(X = 5) = 0$$

Question 6

If random variable X follows a Poisson distribution with mean 3.4. Find $P(X = 6)$

Solution:

$$
\begin{aligned}
X &\sim \text{Poisson}(3.4) \\
P(X = 6) &= \frac{e^{-3.4} \times 3.4^{6}}{6!} \\
&= 0.0716
\end{aligned}
$$

Question 7

Cretan Airlines services which arrive late to Athens Airport on a typical week can be modelled by a Poisson distribution with mean of 4.5

1. Determine the probability that on a given week there will be
   (a) four late arrivals
   (b) less than four late arrivals
   (c) at least seven late arrivals
2. Determine the probability that on a given two week period there will be between eight and thirteen (inclusive) late arrivals.

Solution: $X \sim \text{Poisson}(4.5)$

1. (a)

$$
\begin{aligned}
P(X = 4) &= \frac{e^{-4.5} \times 4.5^{4}}{4!} \\
&= 0.18980
\end{aligned}
$$

(b)

$$
\begin{aligned}
P(X < 4) &= \sum_{i=0}^{3} \frac{e^{-4.5} \times 4.5^{i}}{i!} \\
&= 0.3423
\end{aligned}
$$

(c)

$$\begin{aligned} P(X \geqslant 7) &= 1 - P(X < 7) \\ &= 1 - \sum_{i=0}^{6} \frac{e^{-4.5} \times 4.5^{i}}{i!} \\ &= 1 - 0.8311 \\ &= 0.1689 \end{aligned}$$