Methods Summary Notes

2022, Year 12

Contents

1 Differentiation Rules **3**
1.1 The Chain Rule . . . 3
1.2 The Product Rule . . . 3
1.3 The Quotient Rule . . . 3

2 Euler's Number **4**
2.1 Differentiating Exponential Functions . . . 4
2.2 General Exponential Model . . . 4

3 The Differential Relationship of Trigonometric Functions **4**
3.1 Using the Differentiation Rules with Trig Functions . . . 4
3.2 Solving Practical Problems with Trig Functions . . . 5

4 The Second Derivative **5**

5 Optimisation **5**
5.1 The Process . . . 5

6 Rectilinear Motion **6**

7 Change **6**
7.1 Incremental Change . . . 6
7.1.1 Example . . . 6
7.2 Approximate Percentage Change . . . 6

8 Anti-differentiation **7**
8.1 Power Rule . . . 7
8.2 Integrals involving e . . . 7
8.3 Harder Integration . . . 7
8.3.1 Integrating $(ax + b)^n$ . . . 7
8.3.2 Products Requiring Expansion . . . 7
8.3.3 Products Using Reverse Chain Rule . . . 8
8.3.4 Derivatives Before Integrals . . . 8

8.4 Integrals of Trigonometric Functions 8
8.4.1 Type 1 8
8.4.2 Type 2 8
8.4.3 Type 3 8
8.4.4 Type 4 8
8.5 Riemann Sums 8
8.6 The Fundamental Theorem of Calculus 9
8.7 Particular Solutions Using Indefinite Integrals 9
8.8 Marginal and Total Change 9
8.9 Rectilinear Motion with Integrals 9

9 Logarithms **10**
9.1 Laws 10
9.2 Graphing Logarithmic Functions 10
9.2.1 Transformation Generalisations 10
9.2.2 Logarithmic Scales 11

10 Probability **12**
10.1 Review 12
10.2 Discrete Random Variables 12
10.2.1 Combinations 12
10.2.2 Probability Mass Functions 13
10.2.3 Expected Value and Variance 13
10.2.4 Linear Changes of Origin and Scale 13
10.3 Continuous Random Variables 14
10.3.1 Probability Density Functions 14
10.3.2 Cumulative Distribution Functions 14
10.3.3 Mean, Median and Variance 14
10.3.4 Linear Changes of Origin and Scale 15
10.3.5 Uniform CRVs 15
10.4 Distributions 15
10.4.1 Uniform + Non-Uniform Distributions 15
10.4.2 Bernoulli Distributions 16
10.4.3 Binomial Distributions 16
10.4.4 The Normal Distribution 17
10.5 Sampling Methods 18
10.5.1 Vocabulary 18
10.5.2 Bias 18
10.5.3 Methods 18
10.6 Suitability of the Normal Distribution Model 19
10.7 Capture-Recapture Sampling Method 19
10.8 Sample Proportions 20
10.8.1 Confidence Intervals 20

1 Differentiation Rules

1.1 The Chain Rule

Leibniz notation:

$$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}$$

Function notation:

$$f'(x) = f'(g(x)) \cdot g'(x)$$

To calculate (eg for $(4x-3)^4$):

1. Differentiate with whole bracket: $4(4x-3)^3$
2. Then, multiply by derivative of part inside brackets: $4(4x-3)^3 \times 4 = 16(4x-3)^3$

Reciprocal property of derivatives:

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}$$

1.2 The Product Rule

For a function defined in the form $y = f(x) \cdot g(x)$, the product rule states that

$$\frac{dy}{dx} = f'(x) \cdot g(x) + f(x) \cdot g'(x)$$

Example:

$$\begin{aligned} y &= (3x+8)(7-2x) \\ \frac{dy}{dx} &= (3x+3)(-2) + (7-2x)(3) \\ &= -12x+5 \end{aligned}$$

1.3 The Quotient Rule

For a function in the form $y = \frac{f(x)}{g(x)}, g(x) \neq 0$, the quotient rule states that

$$\frac{dy}{dx} = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$$

When using the quotient rule, it can be useful to first write $g(x)$ on the numerator.

2 Euler's Number

Euler's number is obtained by a continuous rate of growth and is irrational.

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = 2.71828128\ldots$$

Also,

$$\lim_{h\to 0}\frac{e^h-1}{h} = 1$$

2.1 Differentiating Exponential Functions

$$\frac{d}{dx}\left(e^{f(x)}\right) = f'(x)\cdot e^{f(x)}$$

2.2 General Exponential Model

A situation in which there is continuous exponential change can be modelled by $y = y_0 e^{kx}$

- y is dependent variable
- x is independent variable
- y_0 is initial value (when $x = 0$)
- k is rate constant (proportion of change per time unit)

A positive k value represents growth and a negative value represents decay.

3 The Differential Relationship of Trigonometric Functions

$$\frac{d}{dx}(sin(x)) = cos(x)$$

$$\frac{d}{dx}(cos(x)) = -sin(x)$$

$$\frac{d}{dx}(tan(x)) = \frac{1}{cos^2(x)}$$

3.1 Using the Differentiation Rules with Trig Functions

Given a trig function in the form $y = a[f(x)]^n$, according to the chain rule,

$$\frac{dy}{dx} = an[f(x)]^{n-1}\cdot f'(x)$$

3.2 Solving Practical Problems with Trig Functions

When solving problems with trig functions, consider:

- y-intercept
- equilibrium points
- maximum, minimum

4 The Second Derivative

Simply the derivative of the first derivative, it describes how the gradient of the function is changing.

$$\frac{d}{dx}\left(\frac{dy}{dx}\right) \Rightarrow \frac{d^2y}{dx^2} = f''(x)$$

Normal: the line perpendicular to the tangent at a point.

Concavity: the shape explaining how the gradient of a function is changing. Concave down means the function is below its tangent lines, and concave up means it is above its tangents over a certain interval.

Point of Inflection: a point at which concavity changes. There is no concavity at that point.

Given the function $y = f(x)$, the value of the second derivative shows the concavity of the function at that point.

$f''(x) < 0$: function is concave down
$f''(x) = 0$: insufficient information, concavity test required
$f''(x) > 0$: function is concave up

5 Optimisation

Optimisation: a calculus technique used to find the "optimal solution" to a situation $\Rightarrow$ either need to maximise or minimise.

5.1 The Process

1. Interpret problem, draw diagrams
2. Make a function with one variable
3. Differentiate
4. Solve for stationary points
5. Test nature
6. Check endpoints if finite domain provided
7. Answer in the context of the situation

6 Rectilinear Motion

Rectilinear Motion: moving in or along a straight line path.

Acceleration: rate of change of velocity. Second derivative of displacement. If acceleration has the same sign as the velocity, the object is speeding up, else it is slowing down.

7 Change

7.1 Incremental Change

Incremental/approximate change is calculated using incrementally small values of change in $f(x)$ over change in x.

$$\frac{\delta y}{\delta x} \approx \frac{dy}{dx}$$

$$\delta y \approx \frac{dy}{dx}\delta x$$

7.1.1 Example

Find the square of 20.01.

$$\begin{aligned}
Let\ y &= x^2 \\
\frac{dy}{dx} &= 2x \\
\frac{dy}{dx}|_{x=20} &= 40 \\
\delta y &\approx \frac{dy}{dx}|_{x=20} \cdot \delta x \\
&\approx 40 \times 0.01 \\
&\approx 0.4 \\
\therefore (20.01)^2 &\approx 400 + 0.4 \\
&\approx 400.4
\end{aligned}$$

7.2 Approximate Percentage Change

If δy represents an approximate/incremental change in y then $\frac{\delta y}{y}$ represents the approximate percentage change in y.

$$\frac{\delta y}{y} \approx \frac{\frac{dy}{dx}\delta x}{y}$$

8 Anti-differentiation

Anti-differentiation is the process of finding $f(x)$ with a given gradient function.

$$\frac{dy}{dx} = f'(x) \Rightarrow y = f(x) + c$$

Family of curves: all curves that have the same derivative. A **general solution** involves adding c to the end, and a **particular solution** is one where c is known.

Linearity property: scalar multiplication and addition can occur in integrals.

$$\int kf(x)dx = k\int f(x)dx$$

$$\int [f(x) \pm g(x)]dx = \int f(x)dx \pm \int g(x)dx$$

8.1 Power Rule

If

$$\frac{d}{dx}(ax^n) = nax^{n-1}$$

then

$$\int ax^n dx = \frac{ax^{n+1}}{n+1} + c$$

8.2 Integrals involving e

$$\int e^{kx}dx = \frac{1}{k}e^{kx} + c$$

$$\int f'(x)e^{f(x)}dx = e^{f(x)} + c$$

8.3 Harder Integration

8.3.1 Integrating $(ax+b)^n$

$$\int (ax+b)^n dx = \frac{1}{a(n+1)}(ax+b)^{n+1} + c$$

8.3.2 Products Requiring Expansion

Suppose

$$\int f(x) \cdot [g(x)]^n dx$$

If $f(x)$ is not related to $g'(x)$, expand the integrand and use reverse power rule.

8.3.3 Products Using Reverse Chain Rule

If

$$\int f(x) \cdot [g(x)]^n dx$$

and $f(x)$ is related to $g'(x)$, let $y = [g(x)]^{n+1}$ and compare.

8.3.4 Derivatives Before Integrals

The derivative of a product might tell you something about a related integral. (TODO)

8.4 Integrals of Trigonometric Functions

8.4.1 Type 1

$$\int asin(b(x-c))dx = -\frac{a}{b}cos(b(x-c)) + k$$

$$\int acos(b(x-c))dx = \frac{a}{b}sin(b(x-c)) + k$$

8.4.2 Type 2

Where $g(x)$ is sin or cos:

$$\int kf'(x) \cdot g(f(x))dx$$

8.4.3 Type 3

$$\int f(x) \cdot [g(x)]^n dx$$

If $f(x)$ is related to $g'(x)$, let $y = [g(x)]^{n+1}$ and compare.

8.4.4 Type 4

8.5 Riemann Sums

Riemann sum: numerical method of approximating area under a curve over an interval $a \leq x \leq b$ by partitioning area into finite number of rectangles and summing their areas together. Left Riemann sum has rectangles touch curve at top-left corner,

$$A_{left} \approx \sum_{i=0}^{n-1} f(x_i)\Delta x$$

right Riemann sum has rectangles touch curve at top-right corner.

$$A_{right} \approx \sum_{i=1}^{n} f(x_i)\Delta x$$

Creates rectangles of width $\Delta x = \frac{b-a}{n}$.

Accuracy of estimation can be improved by decreasing width of rectangles (increases n) or by using middle Riemann sums (midpoint of rect touches curve).

8.6 The Fundamental Theorem of Calculus

From Riemann sums:

$$\lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x = \lim_{\delta x\to 0} \sum_{i=1}^{n} f(x_i)\delta x = \int_a^b f(x)dx$$

Definite integral represents sum of all products over interval.

$$F(x) = \int_a^x f(t)dt$$

$$F'(x) = \frac{d}{dx}\left(\int_a^x f(t)dt\right) = f(x)$$

$$\int_a^x f(x)dx = F(b) - F(a)$$

Properties:

$$\int_a^b kf(x)dx = k\int_a^b f(x)dx$$

$$\int_a^b f(x) \pm g(x)dx = \int_a^b f(x)dx \pm \int_a^b g(x)dx$$

$$\int_a^c f(x)dx = \int_a^b f(x)dx \int_b^c f(x)dx$$

8.7 Particular Solutions Using Indefinite Integrals

Particular solution: when c is found using a specific condition. These are sometimes called initial value problems.

8.8 Marginal and Total Change

Total/net change: change in $f(x)$ when x changes from a to b.

8.9 Rectilinear Motion with Integrals

Change in displacement

$$\int_a^b v(t)dt$$

Total distance travelled

$$\int_a^b |v(t)|dt$$

Change in velocity

$$\int_a^b a(t)dt$$

9 Logarithms

Logarithm: an inverse operation that expresses a relationship between a base and its exponent.

$$a^x = b \Rightarrow log_a(b) = x$$

The *common logarithm* is most applicable in real world; leaving out base implies use of base 10.

Properties of $log_a b$:

- $a > 0$
- so $b > 0$
- $log_a b \in \Re$

9.1 Laws

1. Product-sum law: $log_a(mn) = log_a(m) + log_a(n)$
2. Quotient-difference law: $log_a(\frac{m}{n}) = log_a(m) - log_a(n)$
3. Power-to-scalar law: $log_a(m^n) = nlog_a(m)$
4. $log_a(a) = 1$
5. $log_a(1) = 0$
6. Negative Index law: $log_a(\frac{1}{b}) = -log_a(b)$
7. Change of base law: $log_a(b) = \frac{log_c(b)}{log_c(a)}$

9.2 Graphing Logarithmic Functions

Logarithmic function: inverse function of exponential relationship, base r.

$$y = r^x \to y = log_r x$$

9.2.1 Transformation Generalisations

A vertical dilation of scale factor k of exponential function base r corresponds to a horizontal dilation of scale factor k on logarithmic function.

$$y = k \times r^x \to y = log_r\left(\frac{x}{k}\right)$$

A horizontal dilation of scale factor $\frac{1}{k}$ of exponential function base r corresponds to a vertical dilation of scale factor $\frac{1}{k}$ on logarithmic function.

$$y = r^{kx} \to y = \frac{1}{k} log_r x$$

A horizontal translation of exponential function base r corresponds to a vertical translation of k units on logarithmic function.

$$y = r^{x-k} \to y = log_r x + k$$

A vertical translation of k units of exponential function base r corresponds to a horizontal translation k units on logarithmic function.

$$y = r^x + k \to y = log_r (x - k)$$

A reflection of exponential function about the x-axis corresponds to a reflection about the y-axis of the log function.

$$y = -r^x \to y = log_r (-x)$$

A reflection of exponential function about the y-axis corresponds to a reflection about the x-axis of the log function.

$$y = r^{-x} \to y = -log_r (-x)$$

General form of logarithmic function:

$$y = alog_r(b(x - c)) + d$$

9.2.2 Logarithmic Scales

Logarithmic scale: a scale of measurement for which relatively large or small quantities are represented using more manageable numbers. Examples include earthquake magnitudes, pH, sound levels in decibels.

Semi-log scale: graphical representation of points such that one of the axes is a linear scale while the other is a logarithmic scale. **Log-linear:** y-axis is logarithmic, x-axis is linear. Equation of line has form $log_r y = mx + c$. **Linear-log:** y-axis is linear, x-axis is logarithmic. Equation of line has form $y = mlog_r x + c$.

Log-log scale: graphical representation of points such that both axes are logarithmic scales. Equation of line has form $log_r y = mlog_r x + c$ where m is power of polynomial and c is power of base r that gives dilation factor. Changing a line from a linear-linear graph to a log-log graph retains its general shape; i.e. exponential functions, linear functions and logarithmic functions keep their respective shape on both graphs.

10 Probability

10.1 Review

Sample Space Displays

- Venn diagram
- Two-way table
- Tree diagram
- Array

Conditional Probability

$$P(A|B) = \frac{P(A \cup B)}{P(B)}$$

Independent Events

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \\ \Rightarrow P(A|B) &= \frac{P(A)P(B)}{P(B)} \\ &= P(A) \end{aligned}$$

10.2 Discrete Random Variables

Variable: a characteristic of a population or sample that can vary between members of said population or sample. Usually denoted by a capital letter, e.g. X, Y, Z.

Discrete variables can take on individual, distinct values: integers, associated to counting. **Continuous variables** can take on any value within a given range: usually associated with measuring.

Random variable: numerical values are outcomes of probability experiment. The outcome is unpredictable, governed by the possibilities of a random process, e.g. number of apples picked from any tree in an orchard, height of randomly selected person.

Notation: If N is random variable, $P(N = x)$ represents the probability that the outcome of N is x.

10.2.1 Combinations

Multiplication Principle (AND)

$${}^{n_1}C_{r_1} \times {}^{n_2} C_{r_2}$$

Addition Principle (OR)

$${}^{n_1}C_{r_1} + {}^{n_2} C_{r_2}$$

10.2.2 Probability Mass Functions

Probability Mass Function: function that describes the likelihood for each possible value of the discrete random variable. $P(X = x)$ for discrete random variable X. Can draw probability histogram to represent function.

Axioms of probability. Function must abide by axioms of probability to be considered valid.

$$0 \leq P(X = x) \leq 1$$

$$\sum P(X = x) = 1$$

Cumulative distribution function: calculates cumulative probability (probability of event up to and including some value x)

$$F(x) = P(X \leq x)$$

for all x in the domain of the discrete random variable. Also,

$$P(X < x) = P(X \leq x) - P(X = x)$$

10.2.3 Expected Value and Variance

Expected value of a DRV, $E[X]$ or μ, refers to the mean value of X. Acts as a weighted average as each value of $X = x$ has its own probability.

Mode: observe the outcome with the greatest probability. Median: observe cumulative distribution when 0.5.

Variance of a DRV, $Var[X]$, is measure of spread of outcomes taking into account each likelihood. It is the weighted average of the squared deviations from the mean.

$$Var[X] = E[(X - \mu)^2] = \sigma^2$$

Standard deviation is the square root of the variance. Takes into account the fact that the distances were squared in variance calculation. Returns statistic to same dimension as variable originally. σ^2 and σ cannot be negative.

$$SD[X] = \sigma_x = \sqrt{Var[X]}$$

$$Var[X] = E[X^2] - E[X]^2$$

10.2.4 Linear Changes of Origin and Scale

Transforming a set of outcomes by adding, subtracting, multiplying or dividing by a constant is considered a linear transformation. **Linear change of origin** occurs when each outcome of X is changed by the same addition or subtraction of a constant value. **Linear change of scale** occurs when each outcome of X is changed by the same multiplication or division of a constant value.

Generalisation for linear transformation from $X \to aX \pm b$ where a and b are constants:

$$E[aX \pm b] \Rightarrow aE[X] \pm b$$
$$Var[aX \pm b] \Rightarrow a^2 Var[X]$$
$$SD[aX \pm b] \Rightarrow |a| SD[X]$$

10.3 Continuous Random Variables

Continuous Variable: a variable that can take on any value withing a given range of values. Usually associated with measuring. Has an infinite number of possible values over an interval.

For any given continuous random variable X, $P(X = k) = 0$. The continuous nature of the random variable means that the probability at a specific value k is 0. $P(a \leq X \leq b) = P(a < X < b) = P(a < X \leq b) = P(a \leq X < b)$

10.3.1 Probability Density Functions

Represents an expected smooth continuous curve of the probability distribution of the data. **Probability Density Function:** a function that describes the likelihood for a range of possible values of a continuous random variable. It must abide by the axioms of probability. Values of this function are able to be greater than 1.

10.3.2 Cumulative Distribution Functions

Cumulative Distribution Function: a function $F(x) = P(X \leq x)$ which describes the likelihood from the lower bound up to and/or including the value of x. Can also be written as a piecewise-defined function; convention is to include 0 and 1 case.

$$F(x) = P(X \leq x) = \begin{cases} 0 & x < a \\ \int_a^x f(t)dt & a \leq x \leq b \\ 1 & x > b \end{cases}$$

Notice the connection to the fundamental theorem of calculus.

10.3.3 Mean, Median and Variance

Mean

$$E[X] = \int_a^b x f(x) dx$$

Variance

$$Var[X] = \int_a^b (x - \mu)^2 f(x) dx = \int_a^b x^2 f(x) dx - \mu^2$$

Median

$$\int_a^m f(x)dx = 0.5$$

Percentile: a set of scores that is divided into 100 subgroups. kth percentile identifies the score for which $k\%$ of the data lies below.

nth percentile: the value of k such that $P(X \leq k) = \frac{n}{100}$

10.3.4 Linear Changes of Origin and Scale

The generalisations for discrete random variables also hold true for continuous random variables.

10.3.5 Uniform CRVs

Continuous uniform distribution: a continuous variable such that each outcome is equally likely and the distribution is symmetrical, forming a rectangular-shaped distribution.

For a continuous random variable X denoted as $X \sim U(a, b)$ then

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

$$E[X] = med[X] = \frac{a+b}{2}$$

$$Var[X] = \frac{(b-a)^2}{12}$$

10.4 Distributions

10.4.1 Uniform + Non-Uniform Distributions

Discrete Uniform Distribution: each outcome of probability experiment is equally likely and distribution is symmetrical.

$$X \sim U(n)$$

$$P(X = x) = \frac{1}{n}$$

where n is the parameter of the distribution.

Shape of probability histogram for discrete uniform distribution is **rectangular**, all columns are the same height.

$$E[X] = \frac{b+a}{2}$$

If $a = 0$, then

$$E[X] = \frac{n-1}{2}$$

If $a = 1$, then

$$E[X] = \frac{n+1}{2}$$

$$Var[X] = \frac{(b-a+1)^2 - 1}{12} = \frac{n^2 - 1}{12}$$

10.4.2 Bernoulli Distributions

Bernoulli trial: a trial with only two possible outcomes: success or failure. The outcomes are complementary. **Bernoulli-distributed random variable:** $X \sim Ber(p)$ where p is the probability of success for the single trial, $0 \leq p \leq 1$. Conditions for a Bernoulli distribution:

1. There are only two possible outcomes, success or failure.
2. Probability of success, p, is constant so probability of failure is $1 - p$ or q.

$$E[X] = p$$

$$Var[X] = pq$$

$$SD[X] = \sqrt{pq}$$

10.4.3 Binomial Distributions

A **binomially distributed random variable** has outcomes which are formed by a series of n identical and independently distributed (i.i.d) Bernoulli trials. The variable represents the number of successes.

$X \sim Bin(n, p)$ where n is the number of i.i.d trials and p is the probability of success for the single Bernoulli trial. Probability mass function for binomially distributed DRV, where x is num of successes:

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

For $X \sim Bin(n, p)$,

$$E[X] = np$$

$$Var[X] = np(1-p)$$

Graphically, for binomial distributions,

- Increasing trials decreases spread and height
- Positively skewed when $p < 0.5$
- Negatively skewed when $p > 0.5$

10.4.4 The Normal Distribution

The normal distribution is a distribution that is symmetrical around the mean, where the frequency decreases further away from it. They form a bell-shaped curve. A continuous random variable X is normally distributed with two parameters; the mean μ and variance σ^2. It can be notated using the notation

$$X \sim N(\mu, \sigma^2)$$

Note that σ^2 is the variance and not the standard distribution, so the square root must be found before performing calculations with it.

Probability density function for normal distribution, $f(x)$:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

Properties of the normal distribution curve:

1. Curve is symmetrical about given value of x
2. $f(x) \to 0$ as $x \to \pm\infty$
3. There are two points of inflection either side of the mean, at $x = \pm\sigma$
4. Total area under the curve is equal to 1
5. $P(x = k) = 0$

Standard score: (aka z-score) numerical value representing how many standard deviations above or below the mean a particular score in a data set is.

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

Standard normal distribution: normal distribution model with mean of 0 and standard deviation of 1.

$$Z \sim N(0, 1)$$

All properties are maintained for the Z distribution.

The probability of a score being one, two and three standard deviations from the mean is approximately 68%, 95% and 99.7% respectively.

ClassPad: Distribution → Continuous → normCDf. OR, (statistics mode) Calc → Type: distribution, Normal CD.

Tails: left tail when $P(X < k)$, right tail when $P(X > k)$, center tail when $P(-k < X < k)$.

Quantile: the equivalent decimal representation of a percentile. For example, 99th percentile is equivalent to the 0.99 quantile. Always equivalent to left tail where $P(X \leq x)$.

10.5 Sampling Methods

Sampling methods are the way in which data is collected from a sample.

10.5.1 Vocabulary

- **Population:** the set of all eligible group members.
- **Census:** data collected from every member of the population.
- **Sample:** a select subset of the population.
- **Survey:** data collected from each member of a sample.
- **Statistic:** a numerical measure of a sample.
- **Point estimate:** an estimate of data for an entire population obtained from a sample.

10.5.2 Bias

A biased sample will not be representative of the population, as it will favour some section of the population.

1. Sampling Bias
   - Spatial bias – location based
   - Temporal bias – time based
   - Under-coverage bias – under- or over-representation of population
   - Self-selection bias – results from opt-in process
2. Response Bias
   - Voluntary response bias – randomly selected but member chooses whether or not to respond
   - Non-response bias – unwillingness or inability to respond
   - Leading-question bias – question is asked in a way to prompt a certain response

The best way to reduce the likelihood of bias is to use some type of random process.

10.5.3 Methods

The two types of methods are **probability (random) sampling** and **non-probability (non-random) sampling**.

- Probability Sampling
  1. Simple Random Sampling

2. Systematic Sampling
3. Stratified Sampling
4. Cluster Sampling

- Non-Probability Sampling

1. Convenience Sampling
2. Quota Sampling
3. Volunteer Response Sampling
4. Purposive / Judgement Sampling
5. Snowball Sampling

When discussing methods and bias, be sure to answer concisely and in context of the question.

10.6 Suitability of the Normal Distribution Model

Many naturally occurring random (not necessarily continuous) variables can be modelled with a random distribution.

- Symmetry of data: the data should be symmetrical about the mean
- Proportionality of data: the proportion of data values lying within ranges like the 68%, 95% and 99.7% rule.

The binomial distribution is used for discrete random variables, but it can be approximated using a normal distribution. When using a normal distribution in this case, a continuity adjustment must be used to approximate probabilities from a binomial distribution.

Discrete (X)	**Continuous (X_N)**
$P(X = k)$	$P(k - 0.5 < X_N < k + 0.5)$
$P(X < k)$	$P(X_N < k - 0.5)$
$P(X \leq k)$	$P(X_N < k + 0.5)$

and so on.

10.7 Capture-Recapture Sampling Method

A controlled investigation of an object (eg animal) where objects are

1. captured, tagged and released
2. recaptured over a period of time

The method uses the proportion of the tagged objects in the second sample to estimate the population size.

$$\frac{t_1}{N} = \frac{t_2}{n} \Rightarrow N = \frac{nt_1}{t_2}$$

where

- t_1 is the number captured and tagged in first observation
- n is the number recaptured in second observation
- t_2 is the number already tagged in second observation
- N is unknown population size

10.8 Sample Proportions

The fraction, decimal or percentage that satisfies the success condition.

$$sample\ proportion = \frac{number\ of\ successes}{total\ sample\ size}$$

The sample proportion $\hat{p}$ is a random variable such that $\hat{p} = \frac{X}{n}$ where X is a binomial variable.

10.8.1 Confidence Intervals

Considers multiple samples and the likelihood of a sample satisfying certain proportions. Most, but not all, confidence intervals will contain the true value of p. If sample data has been observed and a single $C\%$ confidence interval $[a, b]$ has been constructed using the sample data, then the probability that the confidence interval contains the population mean is *either 0 or 1.* Upon the repeated construction of $C\%$ confidence intervals using an identical process, we can expect that $C\%$ of the time, the confidence intervals will contain the true population mean.

Sampling distribution of sample proportions: the distribution of the values of $\hat{p}$ for repeated samples of sample size n.

Central limit theorem: for a sufficiently large sample size ($n \geq 30$) and a $p \approx 0.5$, the sampling distribution of sample proportions is approximately normally distributed where $\hat{p} \sim N(p, \frac{p(1-p)}{n})$. As $n \to \infty$, the distribution of $\hat{p}$ becomes more normal.

Effects of changed conditions on confidence intervals.

- When the confidence level increases the z-value increases and so the margin of error increases, hence increasing the width of the confidence interval. *The inverse is also true.*
- When the sample size increases, the standard error decreases and so the margin of error decreases, hence decreasing the width of the confidence interval. *The inverse is also true.*

Calculating Unknowns: Using a confidence interval, we can find critical information about confidence intervals. Given a confidence interval,

1. **Calculating $\hat{p}$.** When given bounds $[a, b]$, the symmetry of the interval means
$$\hat{p} = \frac{a+b}{2}$$

2. **Calculating margin of error.**

$$e = \frac{w}{2} = \frac{b-a}{2}$$

3. **Calculating standard error.** Given that $\hat{p}$ is known and n is given,

$$SD[\hat{p}] = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Alternatively, if the confidence level and margin of error are known,

$$SD[\hat{p}] = \frac{e}{z}$$

4. **Calculating confidence level.** First find the z-score, then $P(-z < Z < z) = C\%$

$$e = z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \Rightarrow z = \frac{e}{SD[\hat{p}]}$$

5. **Calculating sample size.**

$$n = \left(\frac{z\sqrt{\hat{p}(1-\hat{p})}}{e}\right)^2$$

To result in a particular margin of error, a restriction is required on n.

$$n \geq \left(\frac{z\sqrt{\hat{p}(1-\hat{p})}}{e}\right)^2$$

Carefully consider the rounding of n as it may result in a greater margin of error than expected. Generally, rounding up should ensure the restriction is met.

Solution 1 use historical data as a population proportion or another given point estimate from a different sample. **Solution 2** use $\hat{p} = 0.5$ as the value that gives the most conservative value for the sample size if historical data is unavailable. This will maximise the margin of error.

Problems Involving Containment and Precision

Precision: a qualitative measure of how close the estimate is to the true value of the parameter. To obtain a better interval estimate, the width of the confidence interval should be decreased while preserving the confidence level, requiring the sample size to increase.

When the value of $\hat{p}$ tends towards 0.5, the standard error increases and so the margin of error increases.

Comparing Samples to a Population or Other Samples For a claimed value or data-based observation of p, we can comment on three cases.

1. Observed interval is completely below p. There is sufficient evidence to suggest that the claimed p is not supported by the sample.

2. Observed interval contains p. Sufficient evidence to suggest that claimed p cannot be rejected by sample. (the sample does not verify the claimed p)

3. Observed interval is completely above p. There is sufficient evidence to suggest that the claimed p is not supported by the sample.

Level of Significance: complement of the confidence level. For example, a 90% confidence level implies a 10% level of significance.

Statistically different: whether or not a sample is unusually different to the expected statistic, justified by statistical significance levels.