1 Functions

A relation $f : X \to Y$ where $X, Y \subseteq \mathbb{R}$ is a function if, $\forall\ a \in X$, the vertical line $x = a$ cuts the graph of f at exactly one point.

A relation f can be proven to be not a function if there is some $a \in X$ for which the vertical line $x = a$ cuts the graph of f zero times or more than once.

A function f is one-one if, $\forall\ a \in \mathbb{R}$, the horizontal line $y = a$ cuts the graph of f at most once.

A function f can be proven to be not one-one if there is some $a \in \mathbb{R}$ for which the horizontal line $y = a$ cuts the graph of f more than once.

2 APGP

For a series S_n, the nth term of the corresponding progression

$$T_n = S_n - S_{n-1}$$

For an arithmetic progression T_n and an arithmetic series S_n the nth term

$$\begin{aligned} T_n &= a + (n-1)d \\ S_n &= \frac{n}{2}(a + l) \\ &= \frac{n}{2}(2a + (n-1)d) \end{aligned}$$

where a is the first term, l is the last term and d is the common difference. To prove that a series S_n is an arithmetic series,

$$T_n - T_{n-1} \text{ is a constant} \iff S_n \text{ is an arithmetic series}$$

For a geometric progression T_n and a geometric series S_n the nth term

$$\begin{aligned} T_n &= ar^{n-1} \\ S_n &= a\frac{r^n - 1}{r - 1} \end{aligned}$$

where a is the first term and r is the common ratio. To prove that a series S_n is a geometric series,

$$\frac{T_n}{T_{n-1}} \text{ is a constant} \iff S_n \text{ is a geometric series}$$

For a geometric progression with $|r| < 1$ the sum to infinity

$$S_\infty = \frac{a}{1 - r}$$

3 Induction format

Let P_n be the statement $\boxed{\text{LHS}} = \boxed{\text{RHS}}$, $n \in \boxed{X}$.

When $n = \boxed{n}$, LHS $= \boxed{\text{LHS}} = \boxed{\text{LHS value}}$.
RHS $= \boxed{\text{RHS}} = \boxed{\text{RHS value}} =$ LHS.
$\therefore P_{\boxed{n}}$ is true.

Assume P_k is true for some $k \in \boxed{X}$ i.e. $\boxed{\text{LHS}} = \boxed{\text{RHS}}$.

To prove P_{k+1} is also true i.e. $\boxed{\text{LHS}} = \boxed{\text{RHS}}$

$$\begin{aligned} \text{LHS} &= \boxed{\text{LHS}} = \cdots \\ &= \boxed{\text{RHS}} = \text{RHS} \end{aligned}$$

$\therefore P_k$ is true $\implies P_{k+1}$ is true.

Since $P_{\boxed{n}}$ is true, and P_k is true $\implies P_{k+1}$ is true, by mathematical induction, P_n is true $\forall\ n \in \boxed{X}$.

4 Vectors

Let there be two points A and B which have position vectors $\mathbf{a}$ and $\mathbf{b}$.

For the vectors $\mathbf{a}$ and $\mathbf{b}$, their dot product where θ is the angle between them

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= |\mathbf{a}||\mathbf{b}| \cos\theta \\ &= \begin{pmatrix} x_\mathbf{a} \\ y_\mathbf{a} \\ \vdots \end{pmatrix} \cdot \begin{pmatrix} x_\mathbf{b} \\ y_\mathbf{b} \\ \vdots \end{pmatrix} = x_\mathbf{a}x_\mathbf{b} + y_\mathbf{a}y_\mathbf{b} + \cdots \end{aligned}$$

Their cross product where θ is the angle between them and $\hat{\mathbf{n}}$ is a unit vector perpendicular to both in the direction of a right-handed screw turned from $\mathbf{a}$ to $\mathbf{b}$

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= \hat{\mathbf{n}}|\mathbf{a}||\mathbf{b}| \sin\theta \\ &= \begin{pmatrix} x_\mathbf{a} \\ y_\mathbf{a} \\ z_\mathbf{a} \end{pmatrix} \times \begin{pmatrix} x_\mathbf{b} \\ y_\mathbf{b} \\ z_\mathbf{b} \end{pmatrix} = \begin{pmatrix} y_\mathbf{a}z_\mathbf{b} - y_\mathbf{b}z_\mathbf{a} \\ z_\mathbf{a}x_\mathbf{b} - z_\mathbf{b}x_\mathbf{a} \\ x_\mathbf{a}y_\mathbf{b} - x_\mathbf{b}y_\mathbf{a} \end{pmatrix} \end{aligned}$$

From the above definitions

$$\begin{aligned} \mathbf{a} \perp \mathbf{b} &\iff \mathbf{a} \cdot \mathbf{b} = 0 \\ \mathbf{a} \parallel \mathbf{b} &\iff \exists\ \lambda \in (\mathbb{R} \setminus \{0\}) : \mathbf{b} = \lambda\mathbf{a} \end{aligned}$$

The angle between the vectors

$$\theta = \cos^{-1} \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}$$

The length of projection and the projection vector of $\mathbf{a}$ onto $\mathbf{b}$ where F is the foot of the perpendicular from A to OB

$$\begin{aligned} OF &= |\mathbf{a} \cdot \hat{\mathbf{b}}| \\ \overrightarrow{OF} &= (\mathbf{a} \cdot \hat{\mathbf{b}})\hat{\mathbf{b}} \end{aligned}$$

The perpendicular distance from A to OB

$$AF = |\mathbf{a} \times \hat{\mathbf{b}}|$$

A vector normal to $\mathbf{a}$ and $\mathbf{b}$ is simply $\mathbf{a} \times \mathbf{b}$.

If $ABCD$ is a parallelogram then

$$\begin{aligned} \text{area of } ABD &= \frac{1}{2}\left|\overrightarrow{AB} \times \overrightarrow{AD}\right| \\ \text{area of } ABCD &= \left|\overrightarrow{AB} \times \overrightarrow{AD}\right| \end{aligned}$$

For three points A, B and C,

$$A, B \text{ and } C \text{ are collinear} \iff \exists\ \lambda \in (\mathbb{R} \setminus \{0\}) : \overrightarrow{AB} = \lambda\overrightarrow{AC}$$

The general equation of a line is

$$\mathbf{r} = \mathbf{a} + \lambda\mathbf{d} \quad \lambda \in \mathbb{R}$$

where $\mathbf{d}$ is the direction vector of the line.

The parametric equation of a plane is

$$\mathbf{r} = \mathbf{a} + \lambda\mathbf{m}_1 + \mu\mathbf{m}_2 \quad \lambda, \mu \in \mathbb{R}$$

The vector equation of the same plane is

$$\mathbf{r} \cdot \mathbf{n} = D \quad D = \mathbf{a} \cdot \mathbf{n}$$

The cartesian equation of the plane is

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \wedge \mathbf{a} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \implies ax + by + cz = D$$

5 Complex Numbers

For a complex number

$$\begin{aligned} z &= x + \mathrm{i}y \\ &= |z|(\cos(\arg z) + \mathrm{i}\sin(\arg z)) \\ &= |z|e^{\mathrm{i}(\arg z)} \end{aligned}$$

its conjugate and magnitude

$$z^* = x - \mathrm{i}y$$

$$zz^* = |z|^2 = x^2 + y^2 \implies |z| = \sqrt{x^2 + y^2}$$

and its argument

$$\arg z = \begin{cases} \tan^{-1}\frac{y}{x} & x > 0 \\ \pi + \tan^{-1}\frac{y}{x} & x < 0 \wedge y \geq 0 \\ -\pi + \tan^{-1}\frac{y}{x} & x < 0 \wedge y < 0 \end{cases}$$

The properties of the argument $\arg z$ are identical to the properties of the logarithm.

If a polynomial with real coefficients has complex root α, then α^* is also a root.

The fundamental theorem of algebra states that every polynomial equation of degree n has n roots that may not be distinct.

6 Probability

Some useful results are

$$\mathrm{P}(A') = 1 - \mathrm{P}(A)$$

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$$

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

If A and B are independent events then

$$\mathrm{P}(A|B) = \mathrm{P}(A) \wedge \mathrm{P}(B|A) = \mathrm{P}(B) \wedge \mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B)$$

If A and B are mutually exclusive events then

$$\mathrm{P}(A \cap B) = 0 \wedge \mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$$

7 Distributions

For an random variable to be modelled by a binomial distribution, it must have **(a)** a fixed number of trials; **(b)** independent trials; **(c)** identical trials; **(d)** the same probability of success for each trial; and **(e)** two possible outcomes.

For an event to be modelled by a Poisson distribution, it must **(a)** occur at a constant average rate; **(b)** occur singly; and **(c)** have independent occurrences.

For an event X

$$X \sim \mathrm{N}(\mu, \sigma^2) \implies Z = \frac{X - \mu}{\sigma} \sim \mathrm{N}(0, 1)$$

$$\mathrm{P}(X \leq x) = \mathrm{P}(Z \leq \frac{x - \mu}{\sigma})$$

8 Expectation and Variance

The expected value of a random variable is the probability-weighted average of all possible values:

$$\mathrm{E}(X) = \sum x\,\mathrm{P}(X = x) = \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x$$

where $f(x)$ is the probability density function of weighing X, if it is a continuous random variable..

If X and Y are independent random variables, X_1 and X_2 are independent observations of X, and a and b are constants then

$$\begin{aligned} \mathrm{E}(a) &= a \\ \mathrm{E}(aX + bY) &= a\,\mathrm{E}(X) + b\,\mathrm{E}(Y) \\ \mathrm{Var}(a) &= 0 \\ \mathrm{Var}(aX + bY) &= a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y) \\ \mathrm{Var}(X_1 + X_2) &= \mathrm{Var}(X_1) + \mathrm{Var}(X_2) = 2\,\mathrm{Var}(X) \end{aligned}$$

9 Approximations

If X is a random discrete variable such that

$$X \sim \mathrm{B}(n, p)$$

then if n is large and p is small such that $np < 5$,

$$X \sim \mathrm{Po}(np) \text{ approximately}$$

If X is a random discrete variable such that

$$X \sim \mathrm{B}(n, p)$$

then if n is large such that $np > 5 \wedge nq > 5$,

$$X \sim \mathrm{N}(np, npq) \text{ approximately}$$

If X is a random discrete variable such that

$$X \sim \mathrm{Po}(\lambda)$$

then if $\lambda > 10$,

$$X \sim \mathrm{N}(\lambda, \lambda) \text{ approximately}$$

When a normal distribution is used to approximate a discrete distribution, continuity correction must be done.

10 Sampling

If X is a random variable with unknown or non-normal distribution where

$$\mathrm{E}(X) = \mu \text{ and } \mathrm{Var}(X) = \sigma^2$$

central limit theorem states that if sample size n is large,

$$\overline{X} \sim \mathrm{N}(\mu, \frac{\sigma^2}{n}) \text{ approximately}$$

Unbiased estimates of μ and σ^2 are sample mean $\overline{x}$ and s^2 respectively.

$$s^2 = \frac{n}{n-1}(\text{sample variance})$$

11 Hypothesis Testing

11.1 Format

Let X be the $\boxed{\text{something}}$ of a randomly chosen $\boxed{\text{thing}}$.

Let μ be the population mean $\boxed{\text{something}}$ of $\boxed{\text{things}}$.

$$\boxed{\begin{array}{ll} \textbf{One of} & \text{Given } X \sim \mathrm{N}(\mu, \sigma^2), \\ \textbf{or} & \text{Assume } X \sim \mathrm{N}(\mu, \sigma^2), \\ \textbf{or} & \text{Since } n = \boxed{n}\text{, by central limit theorem,} \end{array}}$$

$$\overline{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right) \ \boxed{\text{approximately}}$$

$\mathbf{H}_0$: $\mu = \mu_0$; $\mathbf{H}_1$: $\boxed{\mu < \mu_0 \textbf{ or } \mu > \mu_0 \textbf{ or } \mu \neq \mu_0}$

Test statistic:

$$\boxed{\begin{array}{ll} \textbf{One of } & T = \dfrac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1} \\ \textbf{or} & Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \\ \textbf{or} & Z = \dfrac{\overline{X} - \mu}{S/\sqrt{n}} \end{array}}$$

Level of significance: $\boxed{100\alpha\%}$; reject H_0 if p-value < $\boxed{\alpha}$.

Under H_0, p-value = $\boxed{\text{p-value}}$.

Since p-value = $\boxed{\text{p-value}}$ $\boxed{< \textbf{ or } >}$ α, we $\boxed{\text{do not}}$ reject H_0 and conclude that there is $\boxed{\text{in}}$sufficient evidence, at $\boxed{100\alpha\%}$ level, that $\boxed{\mathrm{H}_1}$.

11.2 Definitions

If the level of significance is $100\alpha\%$, there is a probability of α of concluding that H_1 is true when in fact, H_0 is true.

The p-value is the probability of obtaining a sample mean **(**less than or equal to **or** more than or equal to **or** as extreme or more extreme than**)** $\overline{x}$, assuming H_0 is true.

12 Sampling Methods

A population is a collection of individuals or objects from which we may collect data.

A random sample is a small representative of the population in which every member in the population has an equal probability of being selected.

We often take a sample instead of collecting data from the entire population as **(a)** it may be too costly or time consuming to collect the information from the whole population; or **(b)** the population may be infinite or too large.

12.1 Quota sampling

To take a quota sample, we **(a)** divide the population into mutually exclusive subgroups called strata, namely $\boxed{\text{group}}$ and $\boxed{\text{group}}$; and **(b)** select $\boxed{\text{number}}$ $\boxed{\text{group}}$ and $\boxed{\text{number}}$ $\boxed{\text{group}}$ to form the sample. The sample from each stratum is non-random.

Quota sampling is advantageous as **(a)** it is easy to select the sample and administer the survey; **(b)** no sampling frame needed; and **(c)** it incurs low cost.

However, it is bad in that **(a)** the sample obtained is non-random; and **(b)** it is likely to result in selection bias as interviewer may select those who are easier to interview.

12.2 Simple random sampling

To take a simple random sample, we **(a)** obtain the list of $\boxed{\text{samplees}}$ and number all $\boxed{\text{samplees}}$ from 1 to $\boxed{\text{number}}$; **(b)** use a random number generator to randomly select $\boxed{\text{number}}$ numbers; and **(c)** select the $\boxed{\text{samplees}}$ corresponding to the numbers to form the sample.

Simple random sampling is advantageous as **(a)** the data collected generally free from bias; and **(b)** analysis of data is relatively easy.

However, it is bad in that **(a)** it may be difficult to draw up the sampling frame as it is difficult to identify every member of the population; and **(b)** it may be difficult to get access to members who have been chosen for the sample

12.3 Systematic sampling

To take a systematic sample, we **(a)** obtain the list of $\boxed{\text{samplees}}$ and number all $\boxed{\text{samplees}}$ from 1 to $\boxed{\text{number}}$; **(b)** compute the sampling interval $k = \boxed{\text{calculation}}$; **(c)** randomly select an integer from 1 to $\boxed{\text{number}}$ as the start; and **(d)** select every $\boxed{k}$th member of the population thereafter until the sample of $\boxed{\text{number}}$ is obtained.

Systematic sampling is advantageous as **(a)** the data collected is generally free from bias; and **(b)** analysis of data is relatively easy.

However, it is bad in that **(a)** it may be difficult to draw up the sampling frame as it is difficult to identify every member of the population; **(b)** it may be difficult to get access to members who have been chosen for the sample; and **(c)** there may be selection bias if the members of the population is arranged in a periodic or cyclic pattern.

12.4 Stratified sampling

To take a stratified sample, we **(a)** obtain the list of $\boxed{\text{samplees}}$ and divide the population into mutually exclusive subgroups called strata, namely $\boxed{\text{group}}$ and $\boxed{\text{group}}$; **(b)** calculate the sample size that should be taken for each stratum, proportional to their size i.e. $\boxed{\text{calculation}}$ = $\boxed{\text{number}}$ for $\boxed{\text{group}}$, and $\boxed{\text{number}}$ for $\boxed{\text{group}}$; and **(c)** select the sample from each stratum using simple random sampling.

Stratified sampling is advantageous as **(a)** it is more likely to give a good representative sample of the population; and **(b)** when there are clear strata present, it usually gives more reliable estimates of the population parameters than random or systematic sampling.

However, it is bad in that **(a)** it may be difficult to draw up the sampling frame as it is difficult to identify every member of the population; **(b)** it is more difficult to conduct than random sampling; **(c)** it may be difficult to identify appropriate strata; and **(d)** strata may not be clearly defined.