Business Statistics - MS6107E

Arjun Anil Kumar

April 8, 2025

Outline

Statistical Thinking [1]

- Statistics is the Science of the Data.
- Statistics helps to describe and understand **variability**.
- By variability, we mean that successive observations of a system or phenomenon do not produce exactly the same result.
- For example : The observed variability in gasoline mileage depends on many factors, such as the type of driving that has occurred most recently (city versus highway), the changes in condition of the vehicle over time, the brand and/or octane number of the gasoline used. These factors represent potential sources of **variability** in the system.

Random Experiment

Definition

An experiment that can result in different outcomes, even though it is repeated in the same manner every time, is called a **random** experiment.

Sample Space & Event

Definition

The set of all possible outcomes of a random experiment is called the sample space of the experiment. The sample space is denoted as S.

Definition

A sample space is discrete if it consists of a finite or countable infinite set of outcomes. A sample space is continuous if it contains an interval (either finite or infinite) of real numbers.

Definition

An event is a subset of the sample space of a random experiment.

Basic Set Operations

Definition

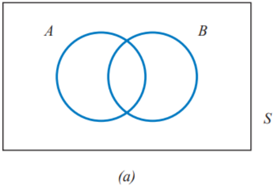
- The union of two events is the event that consists of all outcomes that are contained in either of the two events. We denote the union as $E_1 \cup E_2$.
- The intersection of two events is the event that consists of all outcomes that are contained in both of the two events. We denote the intersection as $E_1 \cap E_2$.
- The complement of an event in a sample space is the set of outcomes in the sample space that are not in the event. We denote the complement of the event E as E^c.

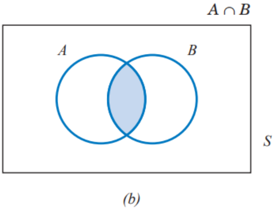
Mutually Exclusive

Definition

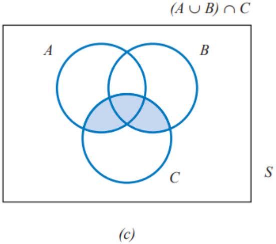
Two events, denoted as E_1 and E_2, such that are said to be mutually exclusive if $E_1 \cap E_2 = \emptyset$

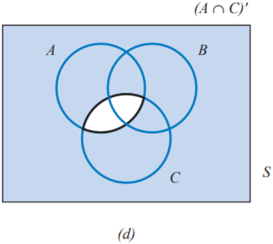
Venn Diagrams





Sample space S with events A and B





Counting Techniques

Definition

Assume an operation can be described as a sequence of k steps, and the number of ways of completing step 1 is n_1, and the number of ways of completing step 2 is n_2 for each way of completing step 1, and the number of ways of completing step 3 is n_3 for each way of completing step 2, and so forth. The total number of ways of completing the operation is $n_1.n_2..n_k$.

Counting Techniques

Permutation

A permutation of the elements is an ordered sequence of the elements.

Definition

The number of permutations of n different elements is n! where n! = n * (n-1) * (n-2) * (n-3) .. 2 * 1

Counting Techniques

Combination

Definition

The number of combinations, subsets of size r that can be selected from a set of n elements, is denoted as $\binom{n}{r} = \frac{n!}{(r!)(n-r)!}$

Counting Techniques

Permutation of Similar Objects

Definition

The number of permutations of $n = n_1 + n_2 + ..n_r$ objects of which n_1 are of one type, n_2 are of a second type and n_r are of an r^{th} type is $\frac{n!}{n_1!n_2!..n_r!}$.

Interpretations and Axioms of Probability

- Probability is used to quantify the likelihood, or chance, that an outcome of a random experiment will occur.
- The likelihood of an outcome is quantified by assigning a number from the interval [0, 1] to the outcome (or a percentage from 0 to 100%.
- Higher numbers indicate that the outcome is more likely than lower numbers.
- The probability of an outcome is interpreted as the limiting value of the proportion of times the outcome occurs in n repetitions of the random experiment as n increases beyond all bounds - Relative Frequency Interpretation of Probability.
- Whenever a sample space consists of N possible outcomes that are equally likely, the probability of each outcome is $\frac{1}{N}$.

Axioms of Probability

Definition

Probability is a number that is assigned to each member of a collection of events from a random experiment that satisfies the following properties. If S is the sample space and E is any event in a random experiment,

1. $P(S) = 1$
2. $0 \leq P(E) \leq 1$
3. For two events E_1 and E_2 with $E_1 \cap E_2 = \emptyset$, $P(E_1 \cup E_2) = P(E_1) + P(E_2)$

Addition Rules

Definition

$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

Definition

$P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3) - P(E_1 \cap E_2) - P(E_2 \cap E_3) - P(E_1 \cap E_3) + P(E_1 \cap E_2 \cap E_3)$

Definition

A collection of events, E_1, E_2, E_k, is said to be mutually exclusive if for all pairs, then $P(E_1 \cup E_2 \cup ..E_k) = P(E_1) + P(E_2) + ...P(E_k)$

Conditional Probability

Definition

The probability of an event B under the knowledge that the outcome will be in event A is denoted as P($\frac{B}{A}$).

Definition

The conditional probability of an event B given an event A, denoted as P($\frac{B}{A}$), is

$$P(\frac{B}{A}) = \frac{P(A \cap B)}{P(A)} \tag{1}$$

where $P(A) > 0$

Multiplication Rule

$$P(A) * P(\frac{B}{A}) = P(A \cap B) \quad (2)$$

Total Probability Rule - Two Events

$$B = (A \cap B) \cup (A^c \cap B)$$

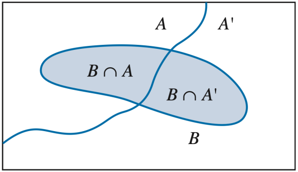


Figure: Partitioning an event into two mutually exclusive events

$$P(B) = P((A \cap B) \cup (A^c \cap B)) = P(B/A)P(A) + P(B/A^c)P(A^c)$$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B/A).P(A)}{P(B)} \quad (3)$$

Concept of Independence

Definition

In some cases, the conditional probability of P(B/A) might equal P(B). In this special case, knowledge that the outcome of the experiment is in event A does not affect the probability that the outcome is in event B.

Two events A and B are independent when one of the three conditions are satisfied.

$P(A/B) = P(A)$
$P(B/A) = P(B)$
$P(A \cap B) = P(A).P(B)$

Bayes Theorem

For $P(B) > 0$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B/A).P(A)}{P(B)} \tag{4}$$

P(A/B) is the conditional probability
P(A/B) is also called posterior probability
P(A) is the prior probability
P(B) is the marginal probability
P(B/A) is the likelihood

Random Variables

Definition

A **random variable** is a function that assigns a real number to each outcome in the sample space of a random experiment.

Definition

A **discrete** random variable is a random variable with a finite (or countably infinite) range. A **continuous** random variable is a random variable with an interval (either finite or infinite) of real numbers for its range.

Probability Distributions and Probability Mass Functions

Definition

The probability distribution of a random variable X is a description of the probabilities associated with the possible values of X.

Definition

For a discrete random variable X with possible values x_1, x_2, .. x_n , a probability mass function is a function such that
$f(x_i) >= 0$
$\sum_1^n f(x_i) = 1$
$f(x_i) = P(X = x_i)$

Cumulative Distributive Function

Definition

The cumulative distribution function of a discrete random variable X, denoted as F(x), is
$F(x) = P(X <= x) = \sum_{x_i <= x} f(x_i)$
$F(x) <= 1$
If $(x < y)$, then $F(x) <= F(y)$

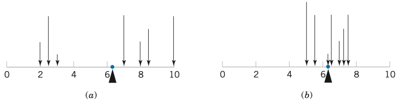
Mean and Variance of Discrete Random Variable

Definition

The mean or expected value of the discrete random variable X, denoted as or E[X], is
$E[X] = \sum_x xf(x)$

Definition

The variance of X denoted as σ^2 or V(X) is
$V[X] = E[(X - \mu)^2] = E[X^2] - \mu^2$



Expected Value of a Function of a Random Variable

Definition

If X is a discrete random variable with probability mass function f(x),
$E[h(X)] = \sum_x h(x)f(x)$

Discrete Distributions

Discrete Uniform Distribution

Definition

A random variable X has a discrete uniform distribution if each of the n values in its range, say, x_1, x_2, .. , x_n, has equal probability. Then, $f(x_i) = \frac{1}{n}$

Discrete Distribution

Binomial Distribution

Examples of Binomial Random Variables

- Flip a coin 10 times. Let X be the number of heads obtained.
- A multiple-choice test contains 10 questions, each with four choices, and you guess at each question. Let X be the number of questions answered correctly.
- In the next 20 births at a hospital, let X be the number of female births.
- Of all patients suffering a particular illness, 35% experience improvement from a particular medication. In the next 100 patients administered the medication, let X be the number of patients who experience improvement.

Discrete Distribution

Binomial Distribution

Definition

A random experiment consists of n Bernoulli trials such that
The trials are independent.Each trial results in only two possible outcomes, labeled as “success” and “failure”.The probability of a success in each trial, denoted as p, remains constant.
The random variable X that equals the number of trials that result in a success has a binomial random variable with parameters $0 <= p <= 1$ and n = 1, 2,.The probability mass function of X is

$$f(x) = \binom{n}{x} p^x.(1-p)^{n-x}$$

where x = 0,1,2,3,..n

Discrete Distribution

Binomial Distribution

- $\binom{n}{x}$ equals the total number of different sequences of trials that contain x successes and n - x failures.
- The total number of different sequences that contain x successes and n - x failures times the probability of each sequence equals P(X=x).
- The sum of probability mass function of binomial distribution is 1.
- For a fixed n, the distribution becomes more symmetric as p increases from 0 to 0.5 or decreases from 1 to 0.5.
- For a fixed p, the distribution becomes more symmetric as n increases.

Discrete Distribution

Binomial Distribution

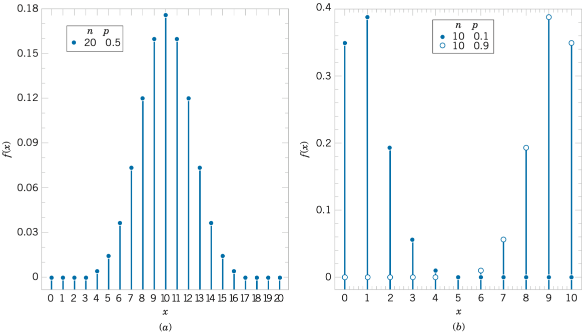


Figure: Binomial Distribution for selected p and n

Discrete Distribution

Binomial Distribution

Definition

If X is a binomial random variable with parameters p and n,

$$\mu = E[X] = np$$

$$\sigma^2 = V[X] = np.(1-p)$$

Discrete Distributions

Geometric Distribution

- Geometric Distribution is very closely linked with Binomial.
- Again, assume a series of Bernoulli trials (independent trials with constant probability p of a success on each trial).
- However, instead of a fixed number of trials, trials are conducted until a success is obtained.
- Let the random variable X denote the number of trials until the first success.

Definition

Then X is a geometric random variable with parameter $0 < p < 1$ and
$f(x) = (1-p)^{x-1}.p$ where x = 1,2,....

Discrete Distributions

Geometric Distribution

Definition

If X is a geometric variable with parameter p, then

$$\mu = \frac{1}{p}$$

$$\sigma^2 = V(x) = \frac{1-p}{p^2}$$

Discrete Distributions

Negative Binomial Distribution

Definition

In a series of Bernoulli trials (independent trials with constant probability p of a success), let the random variable X denote the number of trials until r successes occur. Then X is a negative binomial random variable with parameters $0 < p < 1$ and r = 1, 2, 3, .. and

$$f(x) = \binom{x-1}{r-1}(1-p)^{x-r}.(p)^r$$

Discrete Distributions

Negative Binomial Distribution

Definition

If X is a negative binomial random variable with parameters p and r,

$$\mu = E[X] = r/p$$

$$\sigma^2 = V(x) = \frac{r.(1-p)}{p^2}$$

Discrete Distribution

Hypergeometric Distribution

Definition

A set of N objects contains K objects classified as successes and N - K objects classified as failures. A sample of size n objects is selected randomly (without replacement) from the N objects, where K $\leq$ N and n $\leq$ N. Let the random variable X denote the number of successes in the sample. Then X is a hyper-geometric random variable and

$$f(x) = \frac{\binom{K}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

$x = max(0, n + K - N)$ to $min(K, n)$

Discrete Distribution

Hypergeometric Distribution

Definition

If X is a hypergeometric random variable with parameters N, K, and n, then

$$\mu = E[X] = np$$

$$\sigma^2 = V(x) = n.p.(1-p)\left[\frac{N-n}{N-1}\right]$$

where p $= \frac{K}{N}$

Discrete Distribution

Hypergeometric Distribution

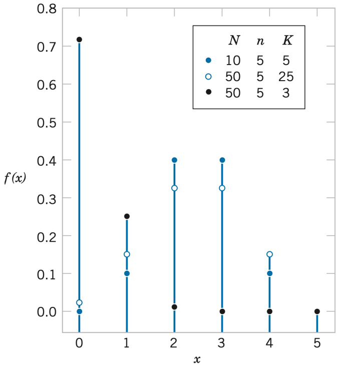


Figure: Hypergeometric Distributions for different N,K and n

Discrete Distribution

Poisson Distribution

Definition

As the number of trials n approaches ∞ with the mean remaining constant, a binomial distribution approaches a Poisson distribution.

$f(x) = \lim_{N\to\infty} \binom{n}{x} p^x.(1-p)^{n-x} = \frac{(e^{-\lambda}).\lambda^x}{x!}$
x = 0,1,2,...

Definition

The random variable X that equals the number of events in a Poisson process is a Poisson random variable with parameter $\lambda >$ 0, and the probability mass function of X is $f(x) = \frac{(e^{-\lambda}).\lambda^x}{x!}$
x = 0,1,2,...

Poisson Distribution

		Probability of X successes						
n	p	$X = 0$	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$	$X = 6$
4	0.25	0.316	0.422	0.211	0.047	0.004		
5	0.20	0.328	0.410	0.205	0.051	0.006	0.000	
10	0.10	0.349	0.387	0.194	0.058	0.011	0.001	0.000
20	0.05	0.359	0.377	0.189	0.060	0.013	0.002	0.000
100	0.01	0.366	0.370	0.185	0.061	0.014	0.003	0.001
1000	0.001	0.368	0.368	0.184	0.061	0.015	0.003	0.001
10000	0.0001	0.368	0.368	0.184	0.061	0.015	0.003	0.001

Figure: Binomial to Poisson as n increases & p decreases and np = 1

Discrete Distribution

Poisson Distribution

Flaws occur at random along the length of a road. Let X denote the random variable that counts the number of flaws in a length of T meters of road and suppose that the average number of flaws in T meters is λ.

- We assume that the long road of T km is divided into strips of length (ΔT) and hence the number of strips is $n = \frac{T}{\Delta T}$.
- We also assume that two flaws cannot occur within the strip that we have created.
- The event in each subinterval is independent of other subintervals.
- The average number of flaws in T meters $\lambda = n.p$
- This implies the chances of strip having a flaw p is $\frac{\lambda}{n} = \frac{\lambda \Delta T}{T}$

Discrete Distributions

Poisson Distribution

Definition

If X is a Poisson random variable with parameters λ,

$$\mu = E[X] = \lambda$$

$$\sigma^2 = V(x) = \lambda$$

Discrete Distributions

Poisson Distribution

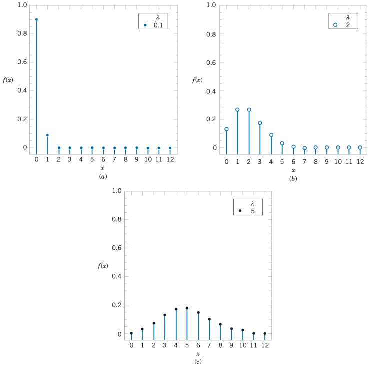


Figure: Parameter .1,2,5 - Poisson Distribution

Probability Density Function

Definition

For a continuous random variable X, a probability density function is a function such that

$f(x) >= 0$

$\int_{-\infty}^{\infty} f(x)dx = 1$

$P(a \leq X \leq b) = \int_{a}^{b} f(x)dx =>$ Area under f(x) from a to b

Probability Density Function

A probability density function f(x) can be used to describe the probability distribution of a continuous random variable X.

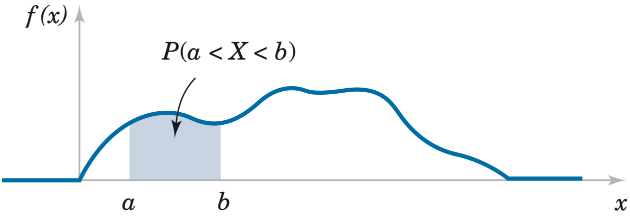


Figure: Area under PDF represents Probability

Histogram & PDF

Definition

A histogram is an approximation to a probability density function. For each interval of the histogram, the area of the bar equals the relative frequency (proportion) of the measurements in the interval.

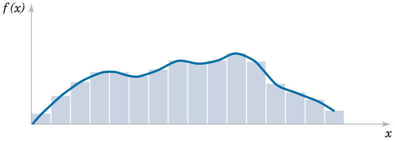


Figure: Histogram from PDF

Cumulative Distributive Function

Definition

The cumulative distribution function of a continuous random variable X is $F(x) = P(X <= x) = \int_{-\infty}^{x} f(u)du$ for $-\infty \leq x \leq \infty$

Mean and Variance of Continuous Random Variable

Definition

Suppose X is a continuous random variable with probability density function f(x). The mean or expected value of X, denoted as μ or E(X), is $\mu = E[X] = \int_{-\infty}^{\infty} x.f(x)dx$

Definition

The variance of X is denoted by V(X) or σ^2
$\sigma^2 = V(X) = \int_{-\infty}^{\infty} (x-\mu)^2.f(x)dx$

$$\sigma^2 = E[X^2] - E[X]^2$$

The square root of variance (σ^2) is the standard deviation (σ).

Expected Value of a Function of a Random Variable

Definition

If X is a continuous random variable with probability density function f(x),
$E[h(X)] = \int_{-\infty}^{\infty} h(x)f(x)$

Continuous Uniform Distribution

Definition

A continuous random variable X with probability density function $f(x) = \frac{1}{b-a}$, where a $\leq$ x $\leq$ b is a continuous uniform random variable

Definition

If X is a continuous uniform random variable over a $\leq$ x $\leq$ b
$\mu = \frac{a+b}{2}$
$\sigma^2 = \frac{(b-a)^2}{2}$

Normal Distribution

- Whenever a random experiment is replicated, the random variable that equals the average (or total) result over the replicates tends to have a normal distribution as the number of replicates becomes large.
- De Moivre presented this fundamental result, known as the central limit theorem, in 1733.
- Although De Moivre was later credited with the derivation, a normal distribution is also referred to as a Gaussian distribution.

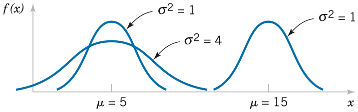


Figure: Normal Distribution

Normal Distribution

Definition

A random variable X with probability density function
$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

is a normal random variable with parameters μ, where $\infty \leq \mu \leq \infty$ and $\sigma > 0$. The mean E[X] is μ and variance V[X] is σ^2.
The normal distribution is denoted as $\mathcal{N}(\mu, \sigma^2)$

Normal Distribution - Symmetry

Definition

A normal distribution is a symmetric distribution because it is symmetric around the mean μ $P(X > \mu) = P(X < \mu)$

Definition

Mean = Mode = Median for a normal distribution

Definition

The area under a normal probability density function beyond 3σ from the mean is quite small.

Normal Distribution - Interesting Properties

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = .6827$$
$$P(\mu - 2.\sigma \leq X \leq \mu + 2.\sigma) = .9545$$
$$P(\mu - 3.\sigma \leq X \leq \mu + 3.\sigma) = .9973$$

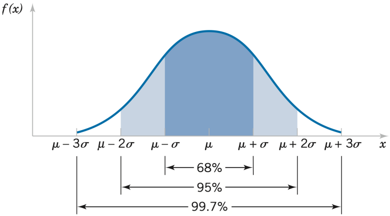


Figure: 6 Sigma Property - Normal Distribution

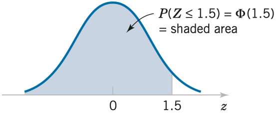
Standard Normal Variable

Definition

A normal random variable with $\mu = 0$ and $\sigma = 1$ is called a standard normal variable and is denoted as Z.

Definition

The cumulative distribution function of a standard normal random variable is denoted as $\phi(z) = P(Z \leq z)$



Standard Normal Table

z	0.00	0.01	0.02	0.03
0	0.50000	0.50399	0.50398	0.51197
$\vdots$		$\vdots$		
1.5	0.93319	0.93448	0.93574	0.93699

Figure: Standard Normal Table $\phi(1.5) = .93319$

Standardizing a Normal Random Variable

Definition

If X is a normal random variable with E(X) $= \mu$ and V(X) $= \sigma^2$, the random variable

$$Z = \frac{X - \mu}{\sigma} \tag{5}$$

is a normal random variable with E(Z) $= 0$ and V(Z) $= 1$.

Definition

A standard normal random variable Z with probability density function $f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(x\right)^2\right)$

Standardizing Normal Variable

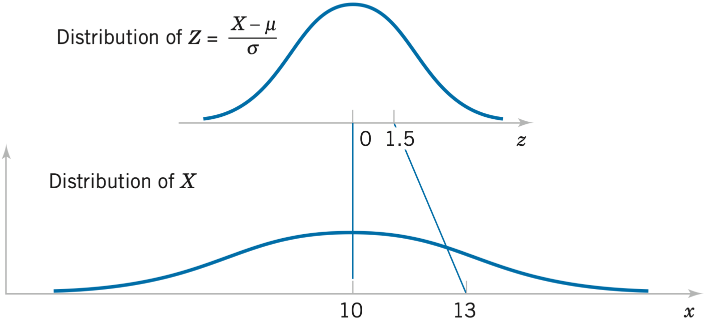


Figure: Standardizing a Normal Variable

Standardizing to Calculate Probability of Normal Random Variable

Definition

Suppose X is a normal random variable with mean μ and variance σ^2

$$P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P\left(Z \leq \frac{x-\mu}{\sigma}\right) \quad (6)$$

Normal Approximation to Binomial Distribution

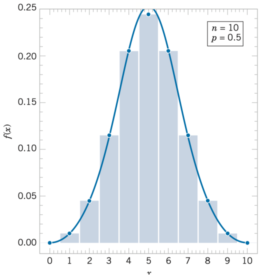


Figure: Normal Approximation to Binomial Random Variable

A binomial random variable has a symmetric distribution when $np > 5 andn.(1-p) > 5$.

Normal Approximation to Binomial Distribution

Definition

If X is a binomial random variable and $np > 5 and n.(1-p) > 5$, then the symmetric binomial random variable can be approximated by a standard normal variable Z, where Z $= \frac{X-np}{\sqrt{np.(1-p)}}$

Normal Approximation to Poisson Distribution

Definition

If X is a Poisson Random Variable with E[X] $= \lambda$ and V[X] $= \lambda$ and $\lambda > 5$, then the poisson distributed random variable X can be approximated as a standard normal variable Z, where Z $= \frac{X-\lambda}{\lambda}$.

Exponential Distribution

Flaws occur at random along the length of a road. Let N denote the random variable that counts the number of flaws in a length of T meters of road. While N follows the Discrete Poisson distribution, the distance to the first flaw follows an Continuous Exponential distribution.

Definition

Let the random variable X denote the length from any starting point on the road until a flaw is detected. If λ represent the mean number of flaws per unit distance,

$$P(X > x) = P(N = 0) = \frac{e^{-\lambda.x}.(\lambda.x)^0}{0!} = e^{-\lambda.x}$$

Exponential Distribution

Definition

Let the random variable X denote the length from any starting point on the road until a flaw is detected. If λ represent the mean number of flaws per unit distance,
$F(x) = P(X \leq x) = 1 - e^{(-\lambda.x)}$, $x >= 0$
$f(x) = \lambda.e^{(-\lambda.x)}$, where $x >= 0$

Note that the starting point for measuring X doesn't matter because the probability of the number of flaws in an interval of a Poisson process depends only on the length of the interval, not on the location.

Exponential Distribution

Definition

The random variable X that equals the distance between successive events of a Poisson process with mean number of events $\lambda > 0$ per unit interval is an exponential random variable with parameter λ. The probability density function of X is

$$f(x) = \lambda.e^{(-\lambda.x)} \tag{7}$$

where $0 \leq x < \infty$

Exponential Distribution

Mean and Variance

Definition

If X is a RV that follows Exponential Distribution

$$E[X] = \frac{1}{\lambda} \tag{8}$$

$$V[X] = \frac{1}{\lambda^2} \tag{9}$$

Exponential Distribution

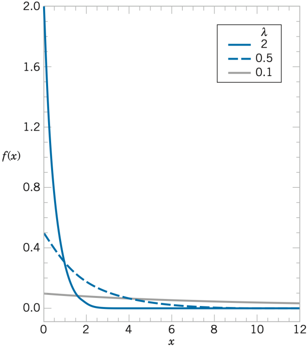


Figure: Exponential Distribution for different λ

Distributions - Summary

Discrete	Continuous
Uniform	Uniform
Bernouli	Normal
Binomial	
Poisson	Exponential
Geometric	Gamma
Hypergeometric	
Negative Binomial	

Table: Distributions

Descriptive Statistics

$data <- c(1,1,1,1,4,4,4,4,8,8,8,8,12,12,12,12)$
summary(data)

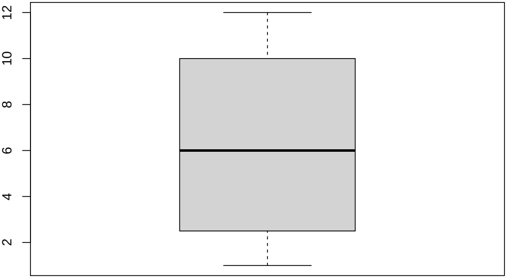
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.00	3.25	6.00	6.25	9.00	12.00

Figure: Descriptive Statistics

Descriptive Statistics

- Descriptive Statistics summarizes the data set.
- The central tendency of any data distribution can be estimated using the **mean, mode and median**.
- The variation of any data distribution around the mean can be estimated using the **variance** or **inter-quartile range (IQR)** $= (Q_{.75} - Q_{.25})$.
-
    - Mean - Average of a data set.
    - Median - The data point that divides the data into equal halves.
    - Maximum/Minimum - The maximum/minimum of a data set.
    - Quantile - Q_λ seperates the data set into two halves where the left half would contain around $100.\lambda\%$ of the data points and the right half would contain around $100.(1-\lambda)\%$ of the data points. $(0 \leq \lambda \leq 1)$

Boxplot



Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.00	3.25	6.00	6.25	9.00	12.00

Figure: Boxplot

Boxplot

- Box-Whisker plot - Box + Whiskers
- Boxplot - Whiskers are drawn at $Q_{.75}$ + 1.5*IQR and $Q_{.25}$ - 1.5*IQR, where IQR is the inter-quartile range ($Q_{.75} - Q_{.25}$).
- Higher IQR indicates higher variance in the data
- Boxplot does not have mean in the plot
- The 5 main points of a box plot are whisker 1, whisker 2, median $Q_{.5}$ and the two quantiles ($Q_{.25}$&$Q_{.75}$).

Boxplot and Outliers

$c(100, 110, 110, 110, 120, 120, 130, 140, 140, 150, 170, 220)$
$Q_{.75}$ = 142.5, $Q_{.25}$ = 110, [L,U] = [61.25,191.25]

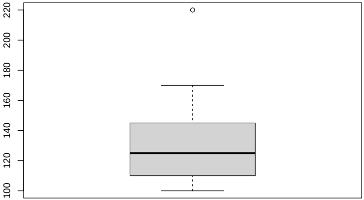


Figure: Interquartile Range $Q_{.75} - Q_{.25} = 32.5$

Symmetric Distributions

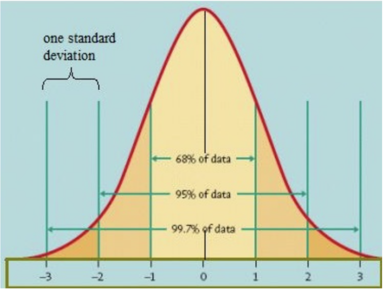


Figure: Symmetric Distribution

A **symmetric** distribution is a type of distribution where the left side of the distribution mirrors the right side.

Symmetric Distributions

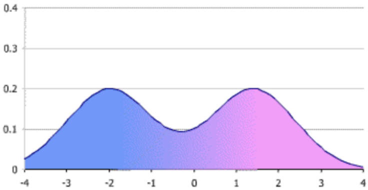
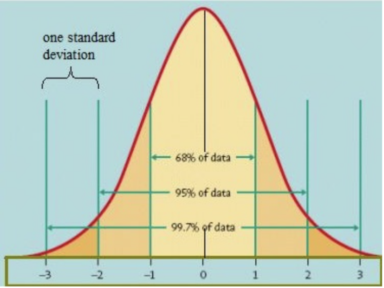
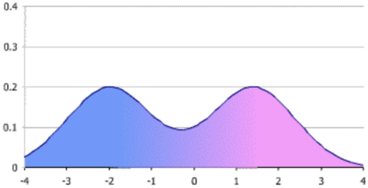


Figure: Symmetric Distributions can be Bi-Modal too!

Symmetric Distributions - Mean, Median, Mode





Symmetric Distributions

Definition

A probability distribution is said to be symmetric if and only if there exists a value x_0, such that $f(x_0 + \delta) = f(x_0 - \delta)$, for $\forall\ \mathbb{R}$, where where f is the probability density function if the distribution is continuous or the probability mass function if the distribution is discrete.

Symmetric Distributions can be either discrete or continuous.

Symmetric Distributions

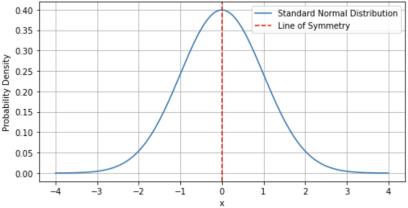
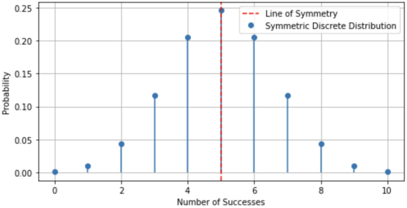
Continuous	Discrete
	

Table: Symmetric Distributions

Skewness

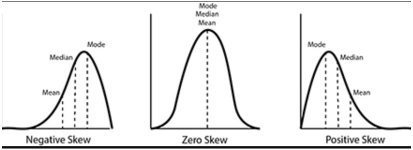


Figure: Skewness

Definition

Skewness measures the **degree** and **direction** of **asymmetry**. A symmetric distribution such as a normal distribution has a skewness of 0, and a distribution that is skewed to the left, e.g. when the mean is less than the median, has a negative skewness and vice-versa.

Skewness

Definition

Skewness can be measured using estimates such as Pearson's Skewness Coefficients S_1,S_2

$$S_1 = \frac{Mean - Mode}{\sigma} \tag{10}$$

$$S_2 = 3.\left(\frac{Mean - Median}{\sigma}\right) \tag{11}$$

If a distribution is symmetric, the skewness is zero

For a moderately skewed distribution $S_1 = S_2$

Bar Plot

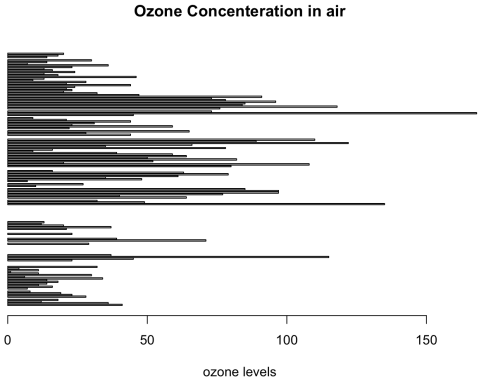


Figure: Bar Plot - Ozone Concentration in Air

Histogram

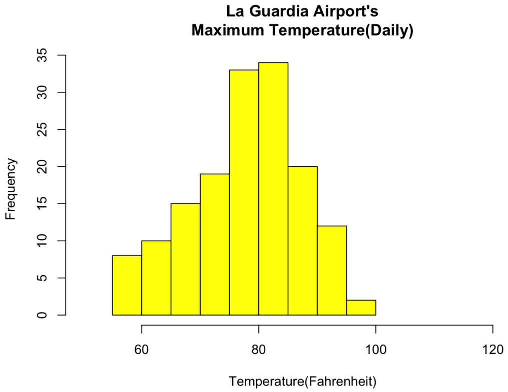
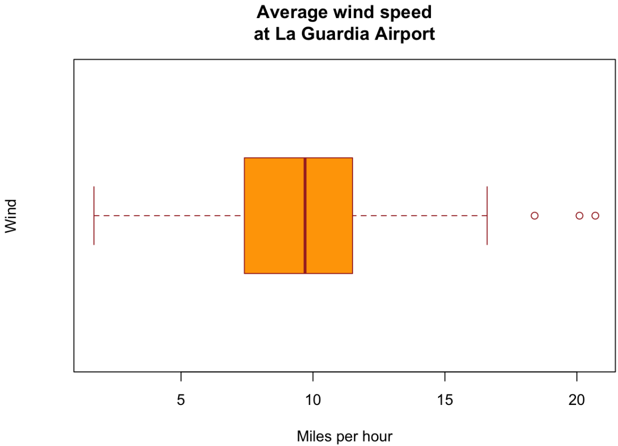
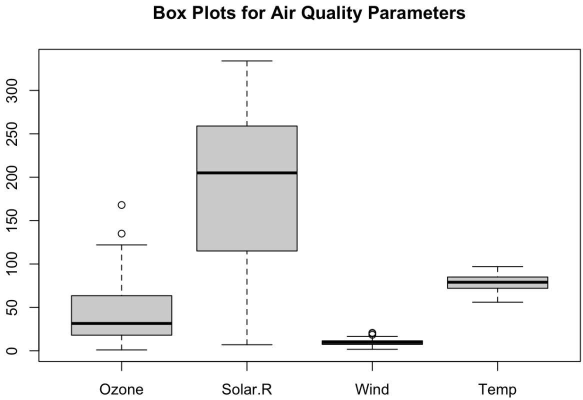


Figure: Histogram - Airport Maximum Temperature

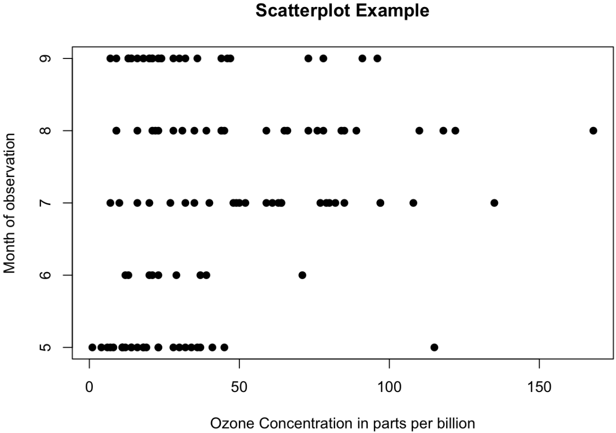
Boxplot


Average wind speed
at La Guardia Airport
Wind
5
10
15
20
Miles per hour

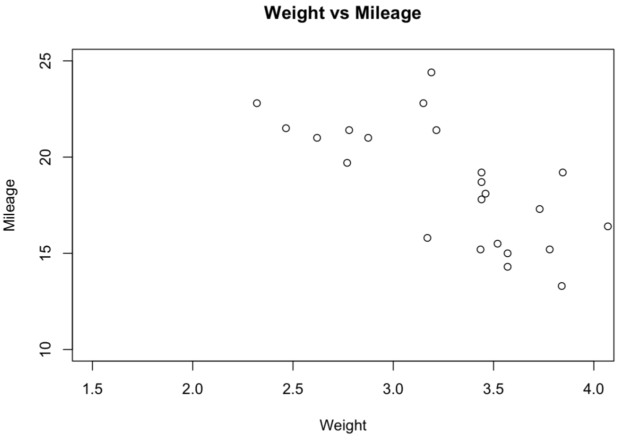
Multiple Boxplot



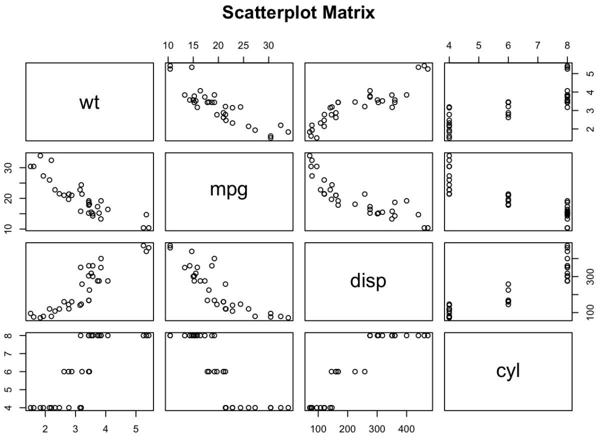
Scatter Plot



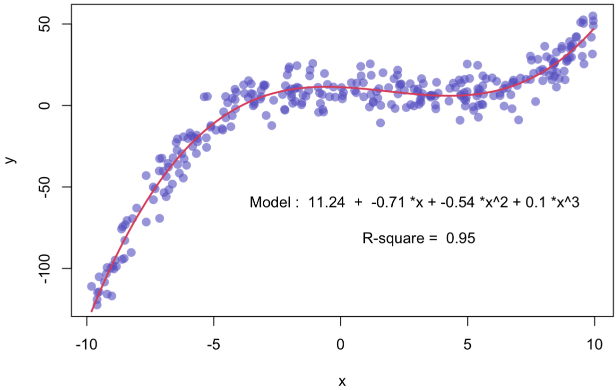
Scatter Plot



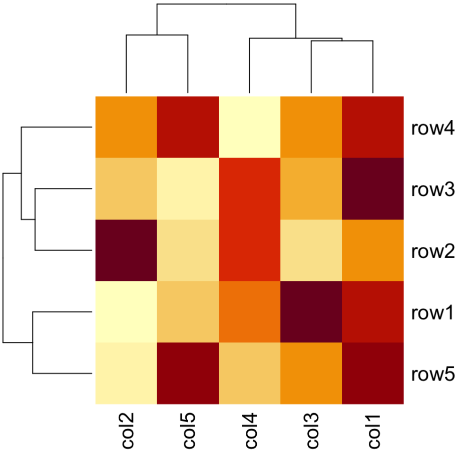
Scatter Plot Matrix



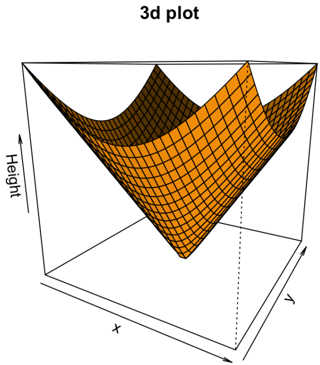
Polynomial Scatter Plot



Heat Map



3-D Plot



Univariate Graphs

Simple Bar Chart

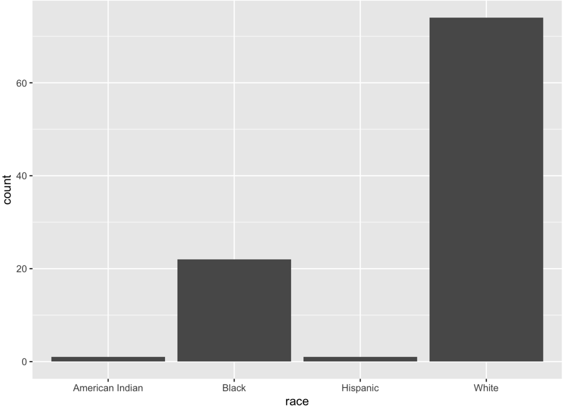


Figure: Barplot - Variable - Race

Univariate Graphs

Distribution - Categorical

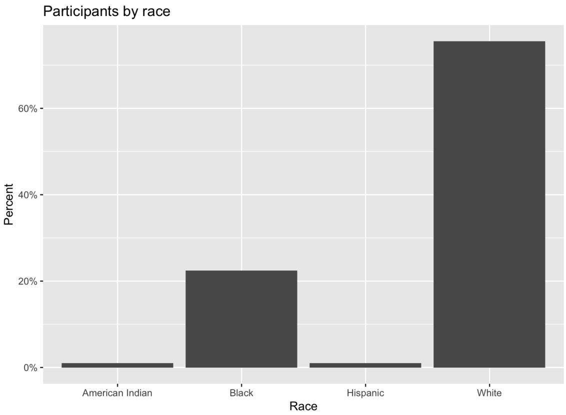


Figure: Distribution - Variable - Race

Univariate Graphs

Pie Chart - Categorical

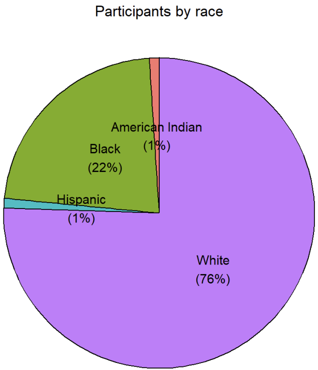


Figure: Pie Chart - Variable - Race

Univariate Graphs

Tree Map - Categorical

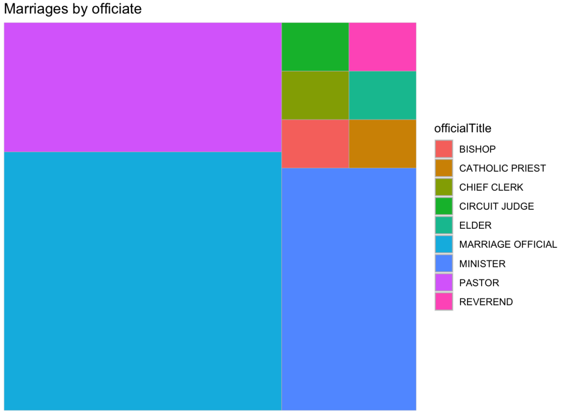


Figure: Tree Map - Marriages by Officiate

Univariate Graphs

Waffle Chart - Categorical

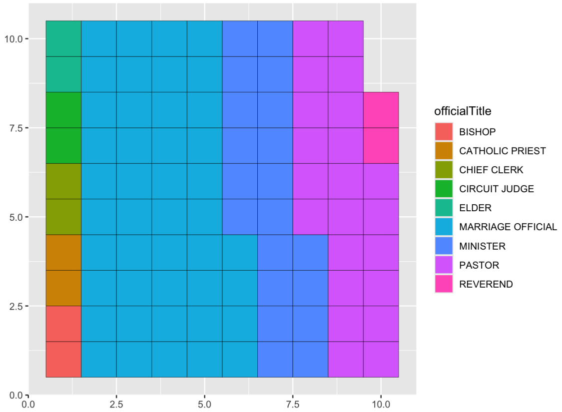


Figure: Waffle Chart - Marriages by Officiate

Univariate Graphs

Histogram - Continuous

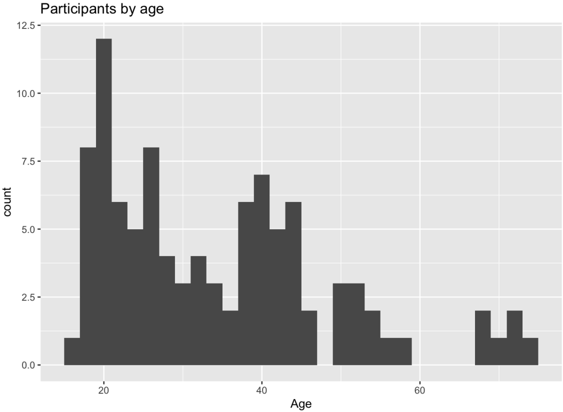


Figure: Histogram - Age

Univariate Graphs

Kernel Density - Continuous

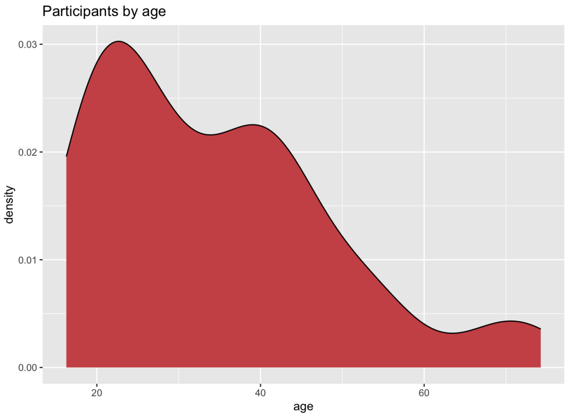


Figure: Kernel Density - Age

Normal Distribution - 1

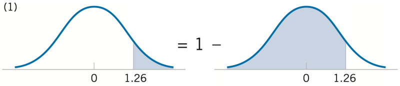


Figure: $P(Z > 1.26)$

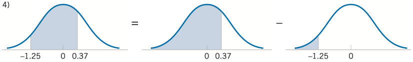


Figure: $P(1.25 < Z < 0.372)$

Normal Distribution - 2

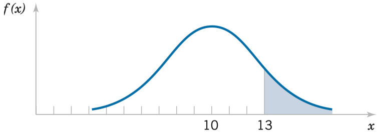


Figure: Find P(X>13) : $\mu = 10$ and $\sigma^2 = 4$

Symmetric Distributions

Case 1 - "A newly started IIM advertisement says "Mean salary is 18 lakh per annum". A candidate joins the institute and the faculty advisor tells him that the median salary is 6 lakhs. "Is there a cause of grave concern?"

Case 2 - "A premier national engineering college advertisement says "Mean salary is 18 lakh per annum". A candidate joins the institute and the faculty advisor tells him that the median salary is 25 lakh. "Is there a cause of grave concern?"

References I

Douglas C Montgomery and George C Runger.
Applied statistics and probability for engineers.
John wiley & sons, 2010.