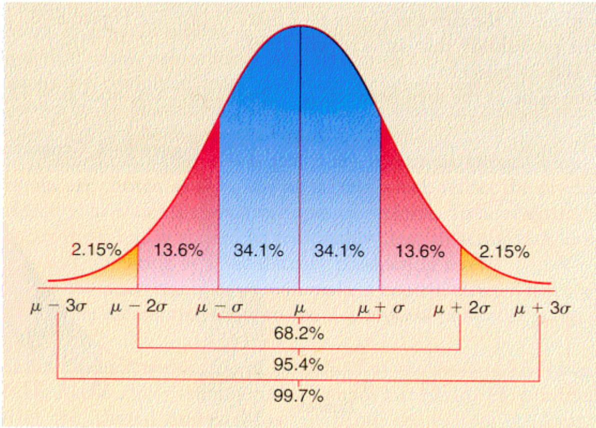
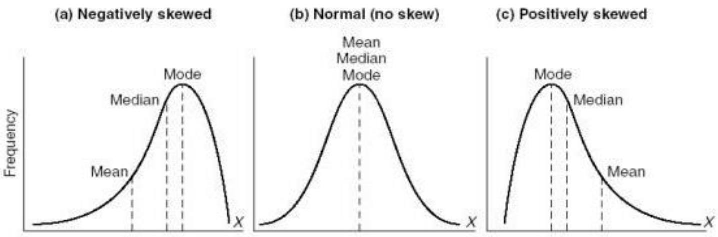
QUANTITATIVE METHODS:

STATISTICAL THEORY:

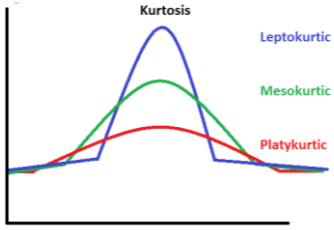
- **Inferential statistics:** using the data to extrapolate to the broader population and predict/learn something about situations and/or events
- **DESCRIPTIVE STATISTICS:** describes the data set in meaningful ways for us to represent and understand the data.
- **Measures of Central Tendency:**
  - **Mean:** arithmetic average
  - **Median:** middle score of data when arranged in ascending numerical order
  - **Mode:** most frequent/popular score
- **Spread:**
  - **Range:** distance between maximum and minimum scores (100%-0%)
  - **IQR (Inter-quartile range):** middle 50% of data set (75%-25%)
  - **Standard Deviation:** measures the average deviation of values from the mean
    - 68% within 1 standard deviation
    - 95% within 2 standard deviations
    - 99.7% within 3 standard deviations



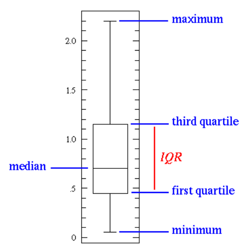
- **Tabulating data:**
  - **Frequency tables:** useful for counting the number of instances of each possible value
  - **Cross tabulation:** may be useful for the breakdown of one variable by another.
- **Skew:** asymmetry of the distribution. Very common and causes issues in data analysis
  - **Positive skew:** majority of data is left skewed
  - **Negative skew:** majority of data is right skewed



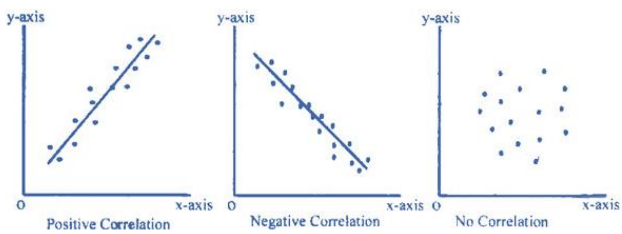
- **Kurtosis:** describes pointiness. It is not often a major effect on graphs and data analysis.



- **Boxplots:** data described with 5 number summary. (Min, 25%, Median, 75%, Max)
  - **Outliers:** can be worked out and excluded from data for better analysis. Can be risky as it changes data (outlier could be genuine).
    - **Lower boundary:** $First\ Quartile - 1.5 \times IQR$
    - **Upper Boundary:** $Third\ Quartile - 1.5 \times IQR$



- **Scatterplots:** used to visualise relationship between two variables.
  - **Positive relationship:** when X variable increases, Y variable also increases
  - **Negative relationship:** when X variable increases, Y variable decreases.

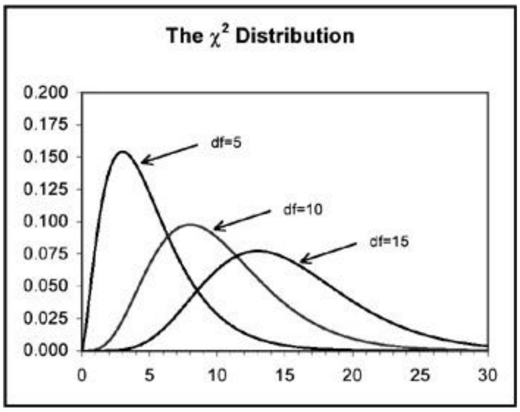


- **Correlation (Pearson's r):** quantifies strength and direction of relationship. (Value between -1 and +1). Assumes relationship is linear.
  - **+ indicates:** positive relationship
  - **– indicates:** negative relationship
  - **Magnitude:** indicates strength
- **PROBABILITY:** relative likelihood that a certain event will or will not occur relative to other events. Expressed as a value between 0 (won't happen) to 1 (will happen) or percentages.
- **Probability distributions:** represent the underlying likelihood of given event outcome.
- **Normal distribution:** symmetrical bell shaped curve. Mean, median and mode are represented at the same place and graph generally shows 3 standard deviations from the mean.
- **Standard normal distribution:** a normal distribution with mean of 0 and standard deviation of 1.
- **Standard Scores (Z scores):** used to compare the relative position of observations within different distributions (e.g. weight of oranges and apples).

- **Formula:** $Z = \frac{X-\mu}{\sigma}$

- **Samples and Populations:** population is ALL members of given group. Sample is a portion of members in the given group. We use samples because it is impossible to measure all members in a population. Therefore using the sample, we estimate what the population might be like. Due to this we have standard error.
- **Sampling Error:** occurs anytime we take a sample of a population. It is the degree to which sample statistics actually differ from population parameters. Use Null Hypothesis Significance Testing
- **Standard Error:** Measure of the degree to which sample statistic deviates from the mean of the sampling distribution.
  - **Formula (Standard Error of Mean): SEM =** $\frac{\sigma}{\sqrt{N}}$
- **Sampling distributions:** generated when we take numerous random samples from the given population.
  - Tend to be near normally distributed
  - Mean of sampling distribution is close to mean of the population distribution.
  - Greater sample size, less variance
- **Confidence Intervals:** provide us with a range of scores within which we can be confident the population parameters lie. Usually estimate 95% confidence = Mean ± 1.96 × Standard Error. Tells us that 95% of the scores lie within 1.96 SD of the true population value.
  - Larger sample size makes the range of Confidence interval smaller.
  - Variability of what is being studied affects range, more precise = narrower CI
  - Degree of confidence required. (Can use 99% Confidence Interval).
- **NULL HYPOTHESIS SIGNIFICANCE TESTING:** helps us determine whether the sample statistics are meaningful.
  - **Formulate a hypothesis:** two competing hypotheses. The Null and Alternative Hypotheses. The null hypothesis predicts that there is no significance.
  - **Measure the variables involved and examine the relationship between them:** this is where we can support or disprove the null hypothesis and consequently adopt or reject the alternative hypothesis.
  - **Calculate the probability of obtaining such a relationship if there was no relationship within the population:** this means we estimate the sampling distribution that would exist if the null hypothesis were true, and then compare the score we obtained with this theoretical distribution. Because there is a probability associated with all the scores under a given sampling distribution, we can calculate the probability associated with obtaining the score that we have actually measured.
  - **If this calculated probability is small enough, it suggests the pattern of findings is unlikely to have arisen by chance and so we can assume that it reflects a genuine relationship within the population:** If the associated probability is low enough, we reject the null hypothesis. The probability associated with a statistical test is called a p value expressed as a decimal ranging between 0 and 1. Usually the decision rule is at: $p < 0.05$ or $p < 0.01$.
- **Interpreting p values:** they are a tool to help us determine whether what we have observed in the samples is a meaningful reflection of the true population values. They represent the probability of an event occurring by chance if the null hypothesis were true.
- **Type 1 errors:** when we reject the null hypothesis although it is really true. (denoted by α)
- **Type 2 errors:** when we fail to reject the null hypothesis when it is really false. (denoted by β)

CATEGORICAL DATA:

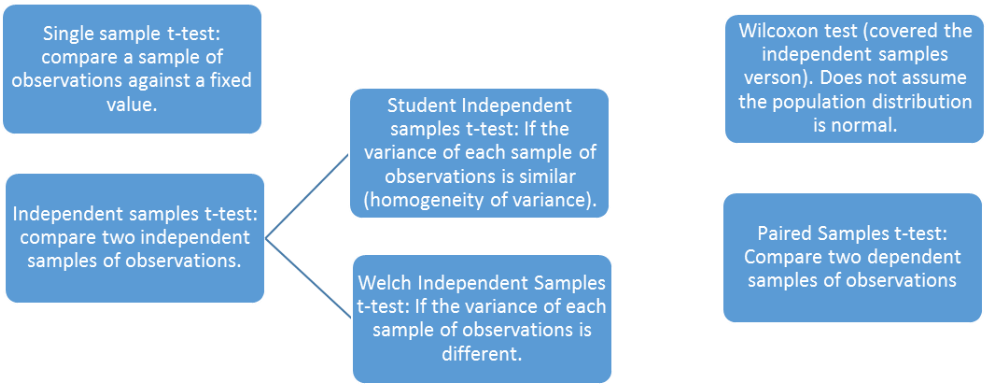


- **Categorical variable:** a variable that has a finite number of possible values, most of the time, are counted.
- **Continuous Variable:** a variable with an infinite number of values. Are usually treated mathematically (averages, sums etc.)
- **Hypothesis testing with categorical data:**
  - **Independence Test:** A hypothesis test which is applied to independent variables
  - **Chi-Squared Test/ Goodness of fit test:** the observed counts are what they are, inferentially, we can use the observed counts to test a hypothesis concerning the underlying probabilities of occurrence.
- **Chi Squared Distribution:** developed as an easy way to calculate the approximation to the G test. Mean is k (the degrees of freedom). We reject the null hypothesis at the 0.05 level of significance.
  - **Degrees of freedom:** the chi-squared distribution depends on k (no. of values sampled from normal distribution). In a goodness of fit test, this is called degrees of freedom.

- Therefore: degrees of freedom = number of categories/ peaks in the data.

COMPARING TWO MEANS (T-TESTS):

- **Z Scores (as previously described):** $Z = \frac{X-\mu}{\sigma}$ Used when raw scores are difficult to interpret without knowing the scale. Therefore, used as a means of comparing two qualitatively different things.



- **Single Samples T-Test:** $(\bar{X} - \mu)/(\hat{\sigma}/\sqrt{n})$ used to find out whether the mean of the samples is true to the population mean.
  - T statistic can be used to test the null hypothesis. The difference between the sample mean and the population mean should be close to 0 if the null hypothesis is true.
  - The statistic is associated with a sample mean rather than a single observation.
  - The sampling distribution is a t-distribution, standard score associated with a single mean rather than a single score.
  - The sampling distribution of our t statistic is a t distribution with N-1 degrees of freedom (number of quantities – number of constraints). As N gets larger, the estimate of the standard error becomes more precise. So the t distribution becomes more similar to a normal.
- **Student Independent Samples T-Test:** used to compare two independent samples of observations if the variance of each sample of observations is similar. Assumes Homogeneity of Variance (generally unrealistic).
- **Welch Independent Samples T-Test:** compares two independent samples of observations if the variance of each sample of observation is different.
  - Effective degrees of freedom is adjusted to be a bit smaller, depending on how unequal the sample standard deviations are.
  - Uses both groups' standard deviations separately.
- **Student or Welch??** Welch tends to be better, but the student test is the one most people are familiar with.
- **Paired Samples T-Test:** compares two dependent samples of observations. Used when we want to see changes within individuals rather than groups. We want to know whether each person's scores went up/ down.
  - **How it works:** run a sample t-test of the difference of scores. Null hypothesis: the mean difference is 0, Alternative Hypothesis: mean difference is not 0.
  - One option is to create a difference variable and run a one sample t test.
- **One sided tests: example:**
  - Null hypothesis: mean is less than or equal to 67.5
  - Alternative hypothesis: mean is greater than 67.5
- **Measuring effect size (Cohen's d):** $(\bar{X}_1 - \bar{X}_2) / \sigma$ is the difference in means divided by the standard deviation
  - If participants are showing the same trend, there is likely to be a significant effect. Increasing the sample size will increase the likelihood of a significant effect. But the effect size itself will not be affected by sample size (does not take into account, sample size N, in the formula).
  - P value will determine how reliable this is.
  - Interpretation:
    - D = 0.2 is a small effect size
    - D = 0.5 is a medium effect size
    - D = 0.8 is a large effect size
- **Assumptions of T tests:** population distribution is normal. Data are independent of one another, except in those respects that the test specifies. Student T-Test assumes that the two groups are of the same variance.

- **Quantile-Quantile plots (QQ):** scatterplots of the observed data against the theoretical quantiles of the normal distribution. If the data is normally distributed, you'd expect the quantile to be identical, giving you a nice straight line. If you don't have a straight line, it isn't a normal distribution.
- **Shapiro-Wilk Test:** the test statistic is W. Values of W less than one imply deviations from normality.
- **When the data isn't normal?** Non parametric tests avoid making assumptions about the shape of the distribution. Although, usually not as powerful as the corresponding parametric tests.
- **Wilcoxon Test:** does not assume the population distribution is normal

ANALYSIS OF VARIANCE (ANOVA):

- **Analysis of variance:** applies to many different situations, especially when you want to see if different groups have different means. Can be used for one grouping variable, or many.
- **One way ANOVA:** used for groups defined by a single variable.
- **How does ANOVA work?**
  - Suppose we have G groups, use μ to refer to the grand mean. Use μ_g to refer to the population mean for group G.
  - The total variability of the outcome variable is literally the sum of between groups. This can be used to test the null hypothesis
  - Null hypothesis: population means for all groups are the same. If the null hypothesis is true, this would mean that the true population means for all the groups are identical. Null hypothesis predicts within group variability rather than between group variability.
  - Alternative hypothesis: population means are not all the same.
- **Summary:**
  - If the null hypothesis is true, we expect that the between groups variability will be small relative to the within groups variability
  - If SS(b) turns out to be large compared with SS(w) we have evidence that the null hypothesis is wrong, and we may reject it.
  - This is the basis for our test statistic, which is a ratio.
- **Intuition to a Test Statistic:** analysis of variance does not use the raw SS values. Instead it uses a corrected estimate based on the degrees of freedom.
  - **Between:** df = G – 1. There are G separated groups means. There is only one grand mean from which they deviate.
  - **Within:** df = N – G. There are N distinct observations. There are G groups of means from which these observations deviate.
- **Mean square:**
  - Within group MS: if all the groups have the same population variance, then MS is an unbiased estimator.
  - Between groups MS: estimates the σ^2 + the effect size. If there is no effect, then MS(b) should be nearly identical to MS(w). IF there is an effect, MS(b) > MS(w).
- **The F Statistic: MS(b) / MS(w).** If the null hypothesis is true, the sampling distribution of the F statistic is an F distribution. F will be close to 1. If it is false, F > 1
- **Effect Size:** the eta squared statistic is conventional, calculated by dividing the SS(b) by the SS(w). Interpreted as the total variance attribute to the grouping variable.
- **Post-hoc Test:** refers to any test that you can run after another test like an ANOVA which has no clear hypothesis or plan about differences you might be interested in.
- **Multiple test correction:** an adjustment that you make to your p value to try and keep the type 1 error in check
  - **Family wise error rate:** the probability of getting at least one type error across multiple tests within one experiment. If you want to keep the error rate at 5%, adjust the p value for the number of tests you do. The simplest way to do this is to multiply all your original p values by the number of tests.
- **Assumptions of ANOVA:** population distributions are normal (check using QQ and Shapiro-Wilk test or residuals), Homogeneity of Variance; that the variances are equal access groups (use levene's test for equality, or welsh and one way ANOVA if violated).
- **Levene's Test:** checks that different groups have the same standard deviation. (ANOVA comparing standard deviation instead of raw scores).