NSW GOVERNMENT

Factsheet

August 2020

Key statistical terms

This Factsheet provides an overview of key statistical terms you may come across in your work and can be used as a reference point to better understand statistics used in research and analysis.

Measures of central tendency

Mean: The mean is a number that measures the average in a set of numbers. To calculate the mean, add up all the numbers in a set and divide it by how many numbers there are. The mean is the most commonly used measure of central tendency. However, it is affected by extreme values or outliers.

Median: The median is the middle value in a set of numbers sorted in ascending order. If the number in the set is odd, the median is the middle number. If the number in the set is even, then the median is the average of the two middle numbers. The median is not affected by extreme values.

Mode: The mode is the most commonly occurring value in a set of numbers. There can be more than one mode in a dataset.

Measures of spread

Range: The range is the difference between the highest and lowest values in a set of numbers.

Variance: Variance is the average of the squared differences from the mean and measures how far a set of numbers are from their mean.

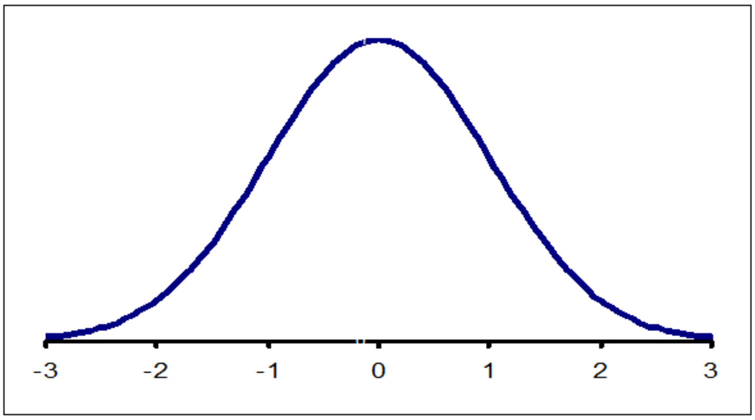
Standard Deviation: Standard Deviation (SD) also measures the spread of numbers in a data set from the mean and is the square root of the variance. It tells us how close the spread of numbers are to the mean. A small standard deviation indicates the data points tend to be close to the mean while a large standard deviation indicates the data points are spread over a wider range of values.

Percentiles: The pth percentile is the value that separates the lowest p-percent of observations from the rest of the data. For example, the 50th percentile is the value below which 50 per cent of the observations fall – this is commonly known as the median. Deciles, quintiles and quartiles are sometimes used in presenting results. These divide a set of measurements into 10, 5 and four equal parts respectively.

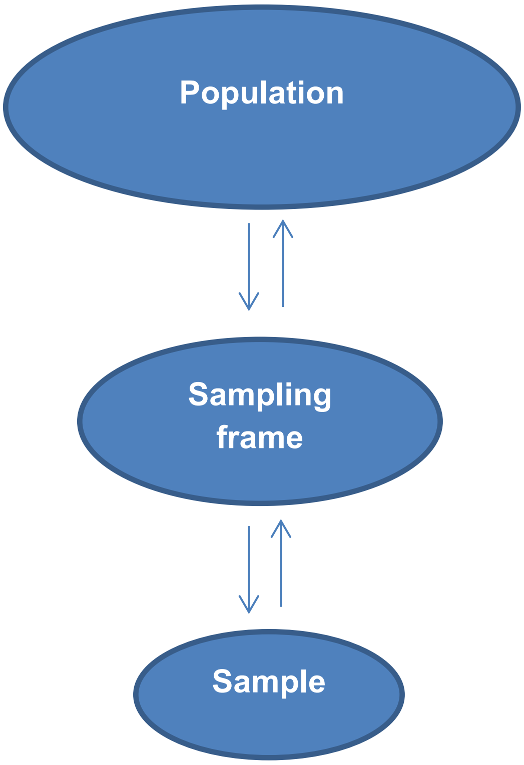
Normal distribution

The normal distribution is described by the mean and the standard deviation and can be represented by a bell-shaped curve. An example is the heights and weights of babies – most are close to average, with a few individuals much smaller or larger than average. Many tests of statistical significance assume that the study population is normally or near-normally distributed.

Figure 1: Graphical representation of a normal distribution



Sampling



Target population and scope

The target population is all relevant subjects of interest to the study. The sample is a manageable subset selected from the target population to make the study feasible.

Sample frame

A sample frame is a complete list of units (e.g. persons, households, secondary assessments) in the target population that we wish to study and from which the sample can be drawn.

Sample frames can take many formats. For a phone survey, the sample frame would be a list of phone numbers. For personal interviews with a householder, the sample frame could be a list of residential addresses or a map showing residential areas.

Sample selection

The main concern when drawing a sample is that it is representative of the larger population. We need to avoid selecting a biased sample that is higher or lower in a particular attribute than the population. There are four main types of sampling strategies that are used in quantitative research to obtain a sample that is representative of the population:

- Simple random sampling involves selecting a number of cases from the sampling frame at random.
- Systematic sampling involves randomly selecting a unit as the starting point and then selecting every nth case. For example, if the population size is 100 and the sample size required is 20 then you can select the sample by dividing 100 by 20 and selecting every 5th case.
- Stratified sampling is designed to select a more representative and accurate sample than other sampling strategies. Stratified sampling involves dividing the sampling frame into non-overlapping sub-populations called strata and taking random samples within each of these. This ensures adequate representation of all strata in the overall sample. For example, the overall sample may be divided into sub-samples of clients based on ethnicity (Anglo-Australian, Aboriginal, Torres Strait Islander and so on) to ensure that each ethnic group is accurately represented in the sample.
- Cluster sampling involves selecting the sample in stages and is commonly used when there is no sampling frame available. The population is divided into groups or clusters and a sample of these are selected to represent the population. Units within the selected clusters are then sampled using simple random sampling or systematic sampling.

Sampling and non-sampling error

Sampling error measures the difference between an estimate derived from a sample and the true value that would be obtained if a census of the population was taken. Sampling error can be measured and is affected by the size of the sample and the sample design.

Non-sampling errors can occur in data derived from both samples and censuses. They can be the result of non-response, or errors in measurement, reporting, coding and processing. They are difficult to measure and can introduce bias into the results.

Sample size

The determination of a sample size is crucial to the success of any study. Larger samples will provide more accurate results as the sampling error will be smaller. However, a survey with a large sample size will be more expensive and difficult to administer than a smaller survey.

Suitable sample sizes are normally calculated using power calculations which are available in most statistical software packages. A power calculation will recommend a certain minimum sample size based on the estimates of variability in the study population and the size of difference to be detected.

Confidence intervals

A confidence interval (calculated from a sample of data) indicates the range within which there is a certain level of confidence that the true population value of a measure lies. The confidence interval is calculated based on the estimated value and its margin of error.

The confidence level (usually expressed as a percentage) is how sure we are that the true value lies within the confidence interval. For example, a 95% confidence interval means that we are 95% sure that the true value will lie within the calculated interval.

Example: The Spotlight on Safety survey calculated a maximum 95% confidence interval of ± 2.53. This means that if 50% of the 1,500 respondents in the survey agreed to an issue then we are 95% confident that the true proportion of the population agreeing with the issue is between 47.47% and 52.53% (i.e. 50 ± 2.53). If 70% of the respondents agreed with an issue then the 95% confidence interval would be narrower at 70 ± 2.32.

Statistical significance

When comparing results it is important to understand whether any differences are real or whether they have occurred by chance. Statistical significance plays an important role in statistical hypothesis testing. Statistical significance refers to the likelihood that two different results, or a particular pattern, could have arisen by pure chance. If it is very unlikely that a particular result could have occurred by chance, then the result is said to be statistically significant.

Null hypothesis

A null hypothesis proposes that no statistically significant differences exist in a set of observations. If the null hypothesis is rejected, statistically significant differences have been observed and these differences are not due to chance.

P-value

The p-value is the way in which statistical significance is usually reported. A p-value of 0.05 indicates that there is a less than 5 per cent chance that the results are due to random chance. P-values usually need to be low (0.05 or less) for a study to make strong claims based on the results.

Effect size (*d*)

Statistical significance refers to the likelihood that the observed results could have arisen by pure chance. Even a very small deviation from the null hypothesis will be statistically significant given a large enough sample.

The effect size is a measure of how strong or important the observed results are and allows comparisons of relative sizes of effects from different studies. The effect size is independent of sample size and for two independent groups is the standardised mean difference between the two groups (known as Cohen's d):

$$Effect\ size\ (d) = \frac{mean_1 - mean_2}{pooled\ SD}$$

T-test (*t*)

A t-test is a simple statistical procedure that compares the means of two number sets and calculates whether they are statistically significantly different. The t-test uses the t statistic *(t)* and its p-value to test the null hypothesis (H_0) that there is no difference between the two samples i.e. the means are the same. A paired t-test can also be used to compare two population means where observations in one sample can be paired with observations in the other sample.

For example, research comparing the number of Risk of Significant Harm (ROSH) reports for children enrolled in a family support program compared to those not enrolled in a program may obtain the following data (figures are based on mock data for illustration only):

	Enrolled in a family support program	Not enrolled in a family support program
Mean $\bar{x}$	5.6	5.8
Standard deviation s	0.54	0.49
Sample size n	310	275

From this data the t statistic is 4.669 and the p-value is <0.001, indicating a highly significant result.

Useful resources

Statistical Language provides a glossary of basic statistical terms designed to assist people to understand statistics.

A guide for using statistics for evidence based policy, published by the Australian Bureau of Statistics (ABS), provides an overview of how statistical information can be used to make well informed policy decisions and to communicate statistical findings. They have also produced A guide for communicating statistical findings, which provides tips on writing statistical commentary and using tables and graphs to communicate statistical findings.

Produced by
Strategy, Outcomes and Evidence
FACS Insights, Analysis and Research (FACSIAR)
NSW Department of Communities and Justice
320 Liverpool Rd, Ashfield NSW 2131
www.facs.nsw.gov.au
Email: facsiar@facs.nsw.gov.au

For more information contact FACSIAR at FACSIAR@facs.nsw.gov.au