# terms - STAT110 Otago

## Students also viewed



## Terms in this set (23)

| | |
|---|---|
| μ | population mean |
| σ | population standard deviation |
| π | population proportion |
| Inference | the formal name given to learning from data using statistical tools. |
| Lower case Roman letters | represent the observed or realised value.<br><br>Pr(X = x) means 'the probability that the random variable X takes the value x'. |
| Random variables are described by | Random variables are described by **probability distributions.** |
| Observed values of random variables are | Observed values of random variables are **data.** |
| Statistics | A statistic is **a numerical summary of data**. |
| estimate | An estimate is a special kind of statistic used as an intelligent guess for a parameter.<br><br>Often **estimates are denoted by adding a circumflex: ˆμ** is an estimate of the parameter μ |

| | |
|---|---|
| $\bar{x}$ | is a statistic and an estimate (for STAT110 paper only) |
| statistical model | a mathematical description of the way the data are generated. |
| Parameter | The numerical measure of the quantity of interest in the population.<br><br>**Parameters are generally unknown, but can be hypothetical** |
| Difference between random variales and obserbed/realised value | **random variables: unknown quantity** varies unpredictable<br>**obserbed/realised value:** got the **actual quantity** of the unknown quantity |
| Types of variables | **Continuous** - can be expressed on a continuous scale in which every value is possible.<br><br>**Discrete** - can be put in one-to-one correspondence with the counting numbers.<br><br>**Categorical** - restricted to one of a set of categories. For example 'Heads' or 'Tails'.<br>*type 1*<br>0 - 1 **binary**<br>A/B/O/AB **more than two**<br>type 2<br>A/B/O/AB **nominal**<br>pass/fail **ordinal** |
| Ratio | fraction given by one quantity over another. Both quantities have the same units. |
| Proportion | fraction of one quantity when compared to the whole. |
| Rates<br>(the difference between rates and ratio) | **Rates are like ratios for quantities with different units.** |
| score | e.g., 'a great deal' / 'somewhat' / 'not much' / 'not at all' |

| | |
|---|---|
| **Types of sensored data** | **Right censored**<br>**Left censored**<br>**Interval-censored**<br><br>Censored data are categorised by two variables |
| Right censored | the **true value is known to be larger than a recorded value**<br><br>for example, we know that someone lived until at least 31 Dec 2017. *50+* |
| Left censored | the **true value is known to be smaller than a recorded value**<br><br>for example, we know that a measurement is less than a known limit of detection. *10-* |
| Interval-censored | the **true value is known to lie between two values**<br><br>for example, we know the date of infection with HPV is after a<br>negative test and before a positive test 2 years later |
| Contingency tables | Contingency tables are often used to record and analyse the<br>relationship between two or more categorical variables |