STAT110/115 Tutoring Materials – 01 Terms

Disambiguation

|  |  |  |  |
| --- | --- | --- | --- |
| Term | Definition | Difficulity | Importance |
| µ | **population** mean | 1 | 5 |
| σ | **population** standard deviation(sd) | 1 | 5 |
| π | **population** proportion | 1 | 5 |
| Lower case Roman letters | represent the observed or realised value.  Lower case Roman letters  Pr(X = x) means 'the probability that the random  variable X takes the value x'. | 1 | 4 |
| ¯  x | a statistic and an estimate (for STAT110 paper only) |  |  |
|  |  |  |  |
|  |  |  |  |

Definition

|  |  |  |  |
| --- | --- | --- | --- |
| inference | the formal name given to learning from data using  statistical tools | 1 | 1 |
| Random variables are  described by | Random variables are described by **probability**  **distributions** | 2 | 3 |
| Observed values of  random variables are | Observed values of random variables are data. |  |  |
| Statistics | A statistic is a numerical summary of data. |  |  |
| estimate | a special kind of statistic used as an  intelligent guess for a parameter.  estimate  Often estimates are denoted by adding a  circumflex: ˆµ is an estimate of the parameter µ |  |  |
| Parameter | The numerical measure of the quantity of interest in  the population.  Parameter  Parameters are generally unknown, but can be  hypothetical |  |  |
| Difference between  random variales and  obserbed/realised value | random variables: unknown quantity varies  unpredictable  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.  Discrete - can be put in one-to-one correspondence  with the counting numbers.  Types of variables  Categorical - restricted to one of a set of  categories. For example ‘Heads’ or ‘Tails’.  type 1  0 - 1 binary  A/B/O/AB more than two  type 2  A/B/O/AB nominal  pass/fail ordinal |  |  |
| the difference between  rates and ratio | Rates are like ratios for quantities with different  units. |  |  |
| Right censored | the true value is known to be larger than a  recorded value  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  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  for example, we know the date of infection with HPV  is after a negative test and before a positive test 2 years later |  |  |
| Contingency tables | Contingency tables are often used to record and  analyse the relationship between two or more categorical  variables |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Reason

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Example

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |