**Lecture 1: Data Handling**

**Sequences**

|  |  |
| --- | --- |
| **Mean** | **Variance** |
|  |  |

**Percentiles**

* is simply the (100)th percentile when
* , and are equivalent to , and , respectively

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

**Lecture 2: Probability Distributions**

**Discrete Random Variables**

|  |  |
| --- | --- |
| **Mean** | **Variance** |
|  |  |
|  |  |

**Bernoulli Random Variables**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **PDF** | | **Notation** |
| , if “success”  , if “failure” |  | |  |
| **Mean** | | **Variance** | |
|  | |  | |

**Binomial Random Variables**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **PDF** | | **Notation** |
| Number of “successes” in a Bernoulli process of trials |  | |  |
| **Mean** | | **Variance** | |
|  | |  | |

**Poisson Random Variables**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **PDF** | | **Notation** |
| Number of certain events occurring in a time interval or region |  | |  |
| **Mean** | | **Variance** | |
|  | |  | |

|  |  |
| --- | --- |
|  | **Notation** |
| Number of certain events occurring in units of time |  |
| **Mean** | **Variance** |
|  |  |

**Continuous Random Variables**

|  |  |  |  |
| --- | --- | --- | --- |
| **Mean** | | **Variance** | |
|  | |  | |
|  | |  | |
| **Calculating the CDF of a PDF** | | | |
| If is in [, ], then | If is in [, ], then | | If is in [, ], then |

**Normal Distribution**

|  |  |  |  |
| --- | --- | --- | --- |
|  | | **Notation** | |
| …is considered to have Normal Distribution if its PDF has the form of the Normal Probability Density Function | |  | |
| **Mean** | **Variance** | | **Standard Score** |
|  |  | |  |

**Lecture 3: Sampling Distribution & Estimation**

**Mean and Standard Deviation of a Sample Mean**

* If is the mean of an SRS of size from a population with mean and standard deviation , then:

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

**Distribution of**

* If a population has distribution, then, for the sample mean, , of independent values…

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

**General Confidence Interval for**

* The 100(1 -)% CI for when is known is given by:

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

* The general conclusion is written as “We are approximately CI confident that the population mean lies between and ”

**Lecture 4: Estimation, Hypothesis & Testing**

**Selecting Sample Size**

* We can select the sample size, , that will guarantee a desired confidence level for a fixed margin of error, .

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

* Note: round up to the nearest whole.

**Test Statistic for Samples with Known Population Standard Deviations**

* To test the validity of an alternative hypothesis we standardise and obtain -score that tells us how many standard errors is from

|  |  |  |
| --- | --- | --- |
|  | : Supposed population mean denoted by | : Known population standard deviation |

**Test Statistics Samples with Unknown Population Standard Deviations**

|  |  |  |
| --- | --- | --- |
|  | : Supposed population mean denoted by | : Known sample standard deviation |

**Confidence Intervals for Samples with Unknown Population Standard Deviations**

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

**When to Use Standard Scores and Standard Errors**

* If is given and , then we should use -scores, otherwise, -scores
* If is given and , then we should use -scores, otherwise, -scores

**P-Values**

* From the tables, we find the probability of obtaining a value or more extreme than the test statistic. This probability is referred to as the -value
* If the -value is relatively small, we say that we have enough evidence to reject . i.e. there is a relatively low probability that is true, so we are willing to promote the

**Conclusions at the Level of Significance**

* “Since the -value of \_\_\_\_\_ is more / less than at the (100)% level of significance, we should reject and state that, on average, the \_\_\_\_\_ \_\_\_\_\_ more / less than ”

**Validity of Conclusions**

* For our conclusions to be valid, the chosen sample should…
  1. Be approximately normal
  2. Not be skewed
  3. Not contain any outliers