Statistical Methods for Big Data

Exercise Sheet 2

1. What exactly does a p-value tell you?

The p-value gives you the probability of obtaining a test statistic at least as large as the one calculated from the data, assuming that the null hypothesis is true.

1. What is a type II error?

A type II error occurs when the null hypothesis is false and we fail to reject it (i.e. a false negative, failing to detect a relationship when one does exist.)

1. As the sample size increases does the probability of making a type II error increase or decrease?

Decreases

1. Define the bias of an estimator. Explain what it means for an estimator to be *unbiased*.

The bias of an estimator is the difference between an estimator's expected value and the true value of the parameter being estimated.

Bias() = E[]-θ

An estimator is said to be an unbiased estimate of a given population parameter when the mean of the sampling distribution of that estimator can be shown to be equal to the parameter being estimated.

1. See script CLT\_die.R