

**Sampling:**

Rnorm: sample from normal distribution

Rnorm(n, mean = , sd = , )

Density: dnorm()

Distribution: pnorm()

Quantile function: qnorm()

qqplot(): quantile plot of two ordered samples

qqnorm(): quantile plot versus normal distribution

qqline(): straight line going through the two quartiles

**Normality test: check if sample has normal distribution**

Manually: qqnorm(); qqline()

**Test:** Shapiro-Wilk Normality test shapiro.test()

Null Hypothesis: sample is normally distributed

**One Sample Student's T-test: checks if sample has same average as the population**

Population variance may be unknown, assume it is the sample variance

Assumes normal distribution

mean(Sample)

sd(Sample)

Test statistic:  $t_0 = (\text{mean}(\text{Sample}) - \text{meanPopulation}) / (\text{sd}(\text{Sample}) / \sqrt{n})$

**Test:** t.test(Sample, meanPopulation, conf.level = x, alternative = 'greater' or 'smaller' if needed)

Null Hypothesis: sample has same mean as population

**One Sample Wilcoxon Test: check if sample has same average as the population**

Does not assume normal distribution, is a non-parametric test

Assumes independent and identically distributed sample, assumes continuous distribution

**Test:** wilcox.test(Sample, meanPopulation)

Null Hypothesis: sample has same mean as population

**Two Sample T-Test: compare means of two populations**

Assumes normal distribution for both samples

**Test:** t.test(Sample1, Sample2, paired = TRUE / FALSE)

Null Hypothesis: Means of the two samples are the same

Assume equal variances:

**Test:** t.test(Sample1, Sample2, var.equal = TRUE)

Null Hypothesis: Means of the two samples are the same

**Fisher's F Test: Compare variances from difference samples**

Assumes normal distribution for both samples

**Test:** var.test(Sample1, Sample2)

Null Hypothesis: variance of both samples is equal

**Wilcoxon's Test: non-parametric comparison of variances of two samples**

**Test:** `wilcox.test(Sample1, Sample2)`

Null Hypothesis: variance of both samples is equal

**Chi-Squared Test: compare distribution of a characteristic between samples**

Requires all frequencies in the table are bigger than 5.

**Test:** `chisq.test(contingencyTable)`

Null Hypothesis: distribution of the characteristic is the same (independent)

**Fisher's Exact Test: compare distribution of a characteristic in 2x2 tables**

**Test:** `fisher.test(contingencyTable)`

Null Hypothesis: distribution of the characteristic is the same (independent)

**One Sample Proportion Test: check if the true proportion of a characteristic within a sample is same as accepted population proportion**

Assume normal distribution, iid

**Test:** `prop.test(nObservationsWithChar, sampleSize, acceptedProportion)`

Null Hypothesis: true proportion is the population accepted proportion

**One Sample Binomial Proportion Test: check if the true proportion of a characteristic within a sample is same as accepted population proportion**

Assume binomial distribution, iid

**Test:** `binom.test(nObservationsWithChar, sampleSize, acceptedProportion)`

Null Hypothesis: true proportion is the population accepted proportion

**Two Sample Proportion Test: compare proportion of a characteristic between two samples**

Assume normal distribution

**Test:** `prop.test(contingencyTable)`

**Analysis of Variance (ANOVA): check if a treatment has an effect on an outcome**

`model = aov(outcome~treatments)`

`summary(aov(outcome~treatments))`

Null Hypothesis: None of the means is significantly different from the others

**Check ANOVA assumptions:**

`par(mfrow = c(2,2))`

`plot(model)` (1,1): fitted line is horizontal and points equal abv and below. (1,2): points fall on line