How to choose your test?

Probability and Statistical Inference

Bojan Božić

TU Dublin

Winter 2020

Descriptive Statistics and Visuals

Dependent on measurement type and shape of distribution:

- For continuous, normally distributed variables:
  - Count, mean, standard deviation (minimum, maximum).
  - Histograms, dot plots, box plots, scatter plots.
- For continuous, skewed variables:
  - Count, median, inter-quartile range (minimum, maximum).
  - Histograms, dot plots, box plots, scatter plots.
- For categorical variables:
  - Frequency counts, percentages.
  - One-way tables, two-way tables.
  - Bar charts, pie charts.

Exploring Data

Descriptive Statistics			Visuals	
Categorical	**Continuous Central Tendency**	**Continuous Variation**	**Categorical**	**Continuous**
Frequency	Mean (non-skewed)	Standard Deviation (non-skewed)	Bar chart	Histogram (with density curve)
Percentage (row, column or total)	Median (skewed)	Inter-quartile range (skewed)	Clustered bar charts (two categorical variables)	Box plot (can be plotted against a categorical variable)
			Pie chart	Scatter plot two continuous variables

Deciding Normality

Visual

Inspect:

- Histogram with density curve.
- QQPlot.

Skew

Calculate standardised skew:

- Approaching normality if $[-2, +2]$.

Kurtosis

Calculate standardised kurtosis:

- Approaching normality if [-2,+2].

Z Scores

Create standardised scores for variable (Z scores):

- Approaching normality if 95% between:
  - $[-3.29, +3.29]$ for sample size > 80.
  - $[-2.5, +2.5]$ for sample size ≤ 80.

Deciding Normality

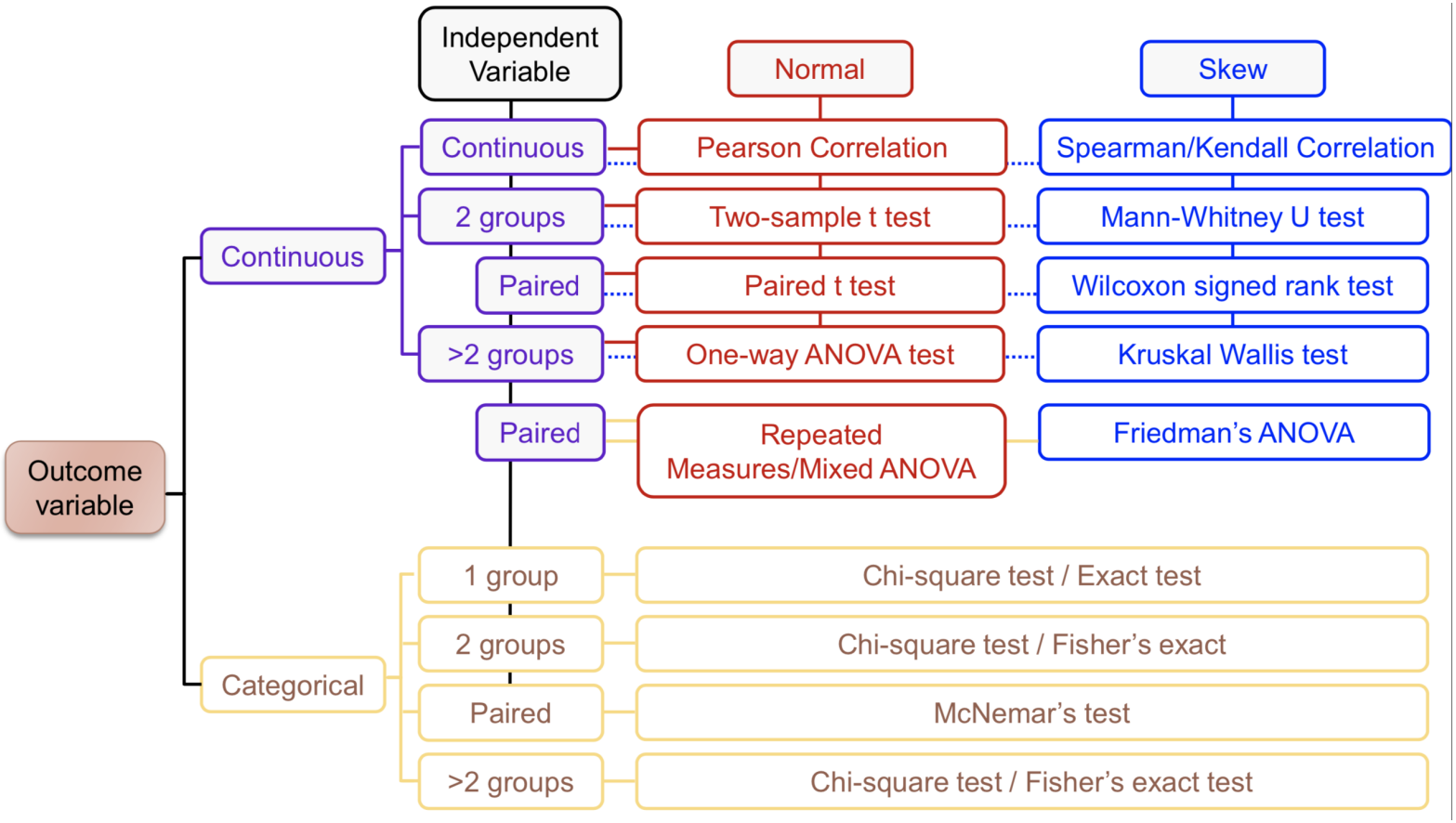
- There are statistical tests for normality.
  - Shapiro-Wilks for small samples.
  - Kolmogorov-Smirnov for large samples (> 50).
- We have not covered these yet as they will usually return a result to indicate non-normality.
  - This does not mean you can't proceed with parametric tests.
- You will complete a normality assessment in your CA as outlined during the module.
  - You cannot rely solely on normality tests, however.

Which test?

What is your measurement type?

- For the outcome variable?
  - Continuous (normal, skew) or categorical?
  - If more than one outcome, are they paired or related?
- For your independent variable?
  - Continuous or Categorical (1 group, 2 groups, more than 2 groups).
  - For 2 or more than 2 groups: Independent (Unrelated) / Paired (Related).

How to choose your test?



Independent variable/ Number of groups	Outcome(Dependent) variable		
	Continuous and Normally distributed (Parametric)	**Continuous and skewed / Ordinal (Non-parametric)**	**Binary (2 categories)**
Continuous	Pearson correlation `cor.test method='Pearson'`	Spearman/Kendall correlation `cor.test method='Spearman'/method='Kendall`	

2 independent groups	T test
	`t.test var.equal=TRUE/FALSE`
	`Pre-check Levene`
		Mann-Whitney U test
		`coin::wilcox.test`
			Chi-square test / Fisher's Exact
			`gmodels::Crosstable`
			`fisher=TRUE`
			`chisq=TRUE`

Paired (related) sample (2 time points)	Paired t test
	`t.test paired=TRUE/FALSE`
	`var.equal=TRUE/FALSE`
		Wilcoxon signed rank test
		`coin::wilcox.test`
		`Paired=TRUE`
			McNemar's test
			`mcnemar.test`

>2 independent groups	One-way ANOVA test
	`aov posthoc: Tukey/Games Howell`
	`posthoc`
	`Pre-check Bartlett`
	
	`Userfriendlyscience::onewaypost`
	`hoc=tukey/games-howell`
		Kruskal-Wallis test
		`kruskal.test`
		`Posthoc test: FSA::dunnTest`
			Chi-square test / Fisher's Exact Test
			`gmodels::Crosstable`
			`fisher=TRUE`
			`chisq=TRUE`

>2 related samples (>2 time points)	Repeated measures ANOVA `Aov`	Friedman's Test `Friedman.test`	

Parametric Difference Tests - Pre-Check

T-test

Levene's Test:

- Non-significant result variances homogenous in groups (var.equal=TRUE).
- Significant result variances heterogeneous (var.equal=FALSE).

ANOVA

Bartlett's Test:

- Non-significant result variances homogeneous in groups (Tukey post-hoc).
- Significant result variances heterogeneous (Games Howell post-hoc).

Test	Effect
Correlation	Pearson, Spearman and Kendall Cohen's d .10 Small; .30 Moderate; .50 Large
Difference (t-test)	Cohen's d .10 Small; .30 Moderate; .50 Large Eta 0.01 = small, 0.06 = moderate, 0.14 =large
Difference (Mann-Whitney/Wilcoxon Signed Rank)	Rosentahl's r 0.1 small 0.5 moderate 0.8 large
Difference (ANOVA)	Eta 0.01 = small, 0.06 = moderate, 0.14 =large
Difference (Kruskal-Wallis)	Eta 0.01 = small, 0.06 = moderate, 0.14 =large
Difference (Chi-squared)	Phi (2 x 2) 0.1 small,;0.3 medium and 0.5 large. Cramer's V

df	*small*	*medium*	*large*
1	.10	.30	.50
2	.07	.21	.35
3	.06	.17	.29
4	.05	.15	.25
5	.04	.13	.22