Inferential Statistics with Variance Analyses

DV = dependent variable (your measure)
IV = independent variable (your factor)
RM = repeated measures (within)
sig. = significant ($p < .05$)
H1 = alternative hypothesis
H0 = null hypothesis
Ⓐ = there is an option for Bayes' Statistics :)
● = regressions require visual inspection of homoscedasticity (Q-Q Plot)

How many DV in hypothesis? → 1 / ≥2

1 DV: Which kind of DV? → continuous

- How many IV? → 1
  - Which kind of IV? → discrete
    - If discrete IV, how many levels? → 2 (B)
      - Within-Subject IV? → no: yes: independent t-test (Ⓐ); no: Mann-Whitney-U-Test — Effect Size: Cohen's d / r (rank-biserial)
      - Within-Subject IV? → yes: paired t-test (Ⓐ); no: Wilcoxon signed-rank test — Effect Size: Cohen's d / r (rank-biserial)
    - If discrete IV, how many levels? → >2
      - Within-Subject IV? → no: yes: one-way ANOVA (Ⓐ); no: Kruskal-Wallis one-way ANOVA — Effect Size: Cohen's f, η^2, partial η^2 / rank ϵ^2
      - Within-Subject IV? → yes: one-way RM ANOVA (Ⓐ); no: Friedman test — Effect Size: Cohen's f, η^2, partial η^2 / Kendall's W
  - Which kind of IV? → continuous
    - Parametric data of DV? → yes: Pearson correlation / linear regression (Ⓐ); no: Spearman rank correlation / Kendall's tau — Effect Size: Pearson's r / R^2, std. β / Spearman's ρ, Kendall's τ
- How many IV? → ≥2
  - Which kind of IV? → discrete
    - Within-Subject IV? → no: yes: multifactorial ANOVA (Ⓐ); no: aligned rank transform ANOVA — Effect Size: f^2, partial η^2 / R^2, std. β / partial η^2
    - Within-Subject IV? → yes: multifactorial RM ANOVA (Ⓐ); no: aligned rank transform RM ANOVA — Effect Size: partial η^2
    - Within-Subject IV? → both: yes: multifactorial mixed-design ANOVA (Ⓐ); no: aligned rank transform mixed-design ANOVA — Effect Size: partial η^2
  - Which kind of IV? → continuous
    - Parametric data of DV? → yes: multiple regression (Ⓐ) — Effect Size: R^2, Cohen's f^2, std. β
  - Which kind of IV? → both
    - Parametric data of DV? → yes: multiple regression / ANCOVA (Ⓐ) — Effect Size: R^2, f^2, std. β / partial η^2

How many DV in hypothesis? → ≥2

- Which kind of DV? → discrete
  - How many IV? → several
    - Parametric data of DV? → yes: Pearson X^2 test or likelihood ratio test (Ⓐ) — Effect Size: Cramer's V, Cohen's W
- Which kind of DV? → continuous
  - How many IV? → 1
    - Which kind of IV? → discrete
      - Parametric data of DV? → yes: MANOVA — Effect Size: partial η^2
  - How many IV? → ≥2
    - Which kind of IV? → discrete
      - Parametric data of DV? → yes: multifactorial MANOVA — Effect Size: partial η^2
    - Which kind of IV? → both
      - Parametric data of DV? → yes: MANCOVA — Effect Size: partial η^2

widely used MANOVA test is *Wilks* λ
using *Pillai's Trace*, *Hotelling's Trace*, *Roy's Largest Root* mostly needs justification

Any Effect sig.?
- no
- yes → go to Ⓐ

Tests for Equivalence

If not sig.:
- Two One-Sided Tests of Equivalence (TOST)
- Mann-Whitney Test for Equivalence (MAWI)
- TOST Paired Samples
- MAWI Paired Samples

Post Hoc Analysis

Sphericity assumed?
- yes
- no → Greenhouse-Geisser

No sig. effect required! → Tukey HSD

or

Main Effect sig.?
- no
- yes → go to Ⓑ

Pairwise Comparisons only with > 2 IV levels

Interaction Effect sig.?
- no
- yes → Check Hypothesis! (Effect Type) → Pairwise Interactions ("Differences of Differences")

Sphericity assumed?
- yes
- no → Greenhouse-Geisser

Main Effect sig.?
- no
- yes → X² Post Hoc

sig.?
- no
- yes → Correlation Matrix

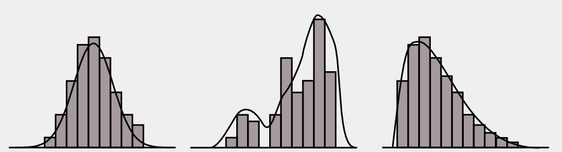
Parametric/Non-Parametric Data

discrete	Nominal (e.g., gender, ZIP codes, animals, strings, types, categories, ...)
continuous	Ordinal (e.g., marks in school, integers, ranks, single Likert items, counts, ...)
continuous	Interval (arbitrary zero: e.g, floats, distance, weight, IQ, temperature in C, ...)
continuous	Ratio (fixed zero: e.g., age, volume, temperature in K, scores, bandwidth,...)

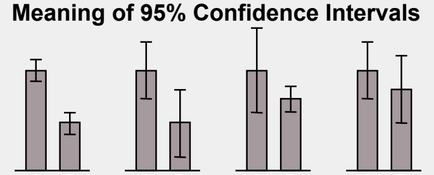
Parametric?
- no
- Test for Normal Distribution (on each condition): Shapiro-Wilk (≤ 50 samples), Kolmogorov-Smirnov
  - one or more sig.?
    - yes → no
    - no → yes

Normality tests can be sensitive to sample size.
Graphical tests (histogram, Q-Q plot) are often more convincing.

Normal Distribution



Meaning of 95% Confidence Intervals



Interpretation of Effect Sizes

	negligible	small	medium	large
Cohen's d	0.20	0.50	0.80	
r, Cohen's/Kendall's W	0.10	0.30	0.50	
Cohen's f	0.10	0.25	0.40	
Cohen's f^2	0.02	0.15	0.35	
η^2	0.02	0.13	0.26	
rank ϵ^2, partial η^2	0.01	0.06	0.14	

Kendall (1939), Cohen (1988)

Some effect sizes can be negative.
This does not affect the interpretation.

Interpretation of Correlations / Fits

	negligible	weak	moderate	strong	perfect
Pearson's r, ρ, τ, R^2, std. β	.20	.50	.80	=1	
Cramer's V (df=1)	.10	.30	.50	=1	
Cramer's V (df=2)	.07	.21	.35	=1	
Cramer's V (df=3)	.06	.17	.29	=1	

Kendall (1939), Cohen (1988), Gravetter & Wallnau (2004)

Effect Types