

Test and effect size details

Indrajeet Patil

2020-07-19

Source: vignettes/stats_details.Rmd

This vignette provides a go-to summary for which test is carried out for each function included in the package and what effect size it returns. Additionally, there are also recommendations on how to interpret those effect sizes.

Summary of statistical tests and effect sizes

Here is a summary table of all the statistical tests currently supported across various functions:

| Functions | Type | Test | Effect size | 95% CI available? |
|--|----------------|---|--|-------------------|
| <code>expr_anova_parametric</code> (2 groups) | Parametric | Student's and Welch's t -test | Cohen's d , Hedge's g | ✓✓ |
| <code>expr_anova_parametric</code> (> 2 groups) | Parametric | Fisher's and Welch's one-way ANOVA | $\eta^2, \eta^2_p, \omega^2, \omega^2_p, \eta^2, \eta^2_p, \omega^2, \omega^2_p$ | ✓✓ |
| <code>expr_anova_nonparametric</code> (2 groups) | Non-parametric | Mann-Whitney U -test | r | ✓✓ |
| <code>expr_anova_nonparametric</code> (> 2 groups) | Non-parametric | Kruskal-Wallis Rank Sum Test | ϵ^2, ϵ^2 | ✓✓ |
| <code>expr_anova_robust</code> (2 groups) | Robust | Yuen's test for trimmed means | ξ, ξ | ✓✓ |
| <code>expr_anova_robust</code> (> 2 groups) | Robust | Heteroscedastic one-way ANOVA for trimmed means | ξ, ξ | ✓✓ |
| <code>expr_anova_parametric</code> (2 groups) | Parametric | Student's t -test | Cohen's d , Hedge's g | ✓✓ |
| <code>expr_anova_parametric</code> (> 2 groups) | Parametric | Fisher's one-way repeated measures ANOVA | $\eta^2_p, \omega^2, \eta^2_p, \omega^2$ | ✓✓ |
| <code>expr_anova_nonparametric</code> (2 groups) | Non-parametric | Wilcoxon signed-rank test | r | ✓✓ |
| <code>expr_anova_nonparametric</code> (> 2 groups) | Non-parametric | Friedman rank sum test | $W, Kendall, W, Kendall$ | ✓✓ |

| Functions | Type | Test | Effect size | 95% CI available? |
|--|----------------|---|---------------------------|-------------------|
| <code>expr_anova_robust</code> (2 groups) | Robust | Yuen's test on trimmed means for dependent samples | $\xi\xi$ | ✓✓ |
| <code>expr_anova_robust</code> (> 2 groups) | Robust | Heteroscedastic one-way repeated measures ANOVA for trimmed means | ×× | ×× |
| <code>expr_contingency_tab</code> (unpaired) | Parametric | Pearson's χ^2 test | Cramér's V | ✓✓ |
| <code>expr_contingency_tab</code> (paired) | Parametric | McNemar's test | Cohen's g | ✓✓ |
| <code>expr_contingency_tab</code> | Parametric | One-sample proportion test | Cramér's V | ✓✓ |
| <code>expr_corr_test</code> | Parametric | Pearson's r | r | ✓✓ |
| <code>expr_corr_test</code> | Non-parametric | Spearman's ρ | $\rho\rho$ | ✓✓ |
| <code>expr_corr_test</code> | Robust | Percentage bend correlation | r | ✓✓ |
| <code>expr_t_onesample</code> | Parametric | One-sample t -test | Cohen's d , Hedge's g | ✓✓ |
| <code>expr_t_onesample</code> | Non-parametric | One-sample Wilcoxon signed rank test | r | ✓✓ |
| <code>expr_t_onesample</code> | Robust | One-sample percentile bootstrap | robust estimator | ✓✓ |
| <code>expr_meta_parametric</code> | Parametric | Meta-analysis via random-effects models | $\beta\beta$ | ✓✓ |
| <code>expr_meta_robust</code> | Robust | Meta-analysis via robust random-effects models | $\beta\beta$ | ✓✓ |

Note that the following recommendations on how to interpret the effect sizes are just suggestions and there is nothing universal about them. The interpretation of **any** effect size measures is always going to be relative to the discipline, the specific data, and the aims of the analyst. Here the guidelines are given for *small*, *medium*, and *large* effects and references should shed more information on the baseline discipline with respect to which these guidelines were recommended. This is important because what might be considered a small effect in psychology might be large for some other field like public health.

(Additionally, you will also see which function is used internally to compute the effect size and their confidence intervals.)

One-sample tests

parametric

Test: One-sample *t*-test

Effect size: Cohen's *d*, Hedge's *g*

Function: [effectsize::cohens_d](#)

| Effect size | Small | Medium | Large | Range |
|------------------|--------------|-----------------|-------------|------------|
| Cohen's <i>d</i> | $0 - < 0.20$ | $0.20 - < 0.50$ | ≥ 0.80 | [-Inf,Inf] |
| Hedge's <i>g</i> | $0 - < 0.20$ | $0.20 - < 0.50$ | ≥ 0.80 | [-Inf,Inf] |

non-parametric

Test: One-sample Wilcoxon Signed-rank Test

Effect size: $rr (= Z/(\sqrt{N_{obs}})Z/(N_{obs}))$

Function: [rcompanion::wilcoxonOneSampleR](#)

| Effect size | Small | Medium | Large | Range |
|-------------|-----------------|-----------------|-------------|-------|
| <i>r</i> | $0.10 - < 0.30$ | $0.30 - < 0.50$ | ≥ 0.50 | [0,1] |

robust

Test: One-sample percentile bootstrap test

Effect size: robust location measure

Function: [WRS2::onesampb](#)

Two-sample tests within-subjects design parametric

Test: Student's dependent samples t -test

Effect size: Cohen's d , Hedge's g

Function: [effectsize::cohens_d](#)

| Effect size | Small | Medium | Large | Range |
|-------------|-------|--------|-------|-------|
| Cohen's d | 0.20 | 0.50 | 0.80 | [0,1] |
| Hedge's g | 0.20 | 0.50 | 0.80 | [0,1] |

non-parametric

Test: Wilcoxon signed-rank test

Effect size: $r_r (= Z/(\sqrt{N_{\text{pairs}}})Z/(N_{\text{pairs}}))$

Function: [rcompanion::wilcoxonPairedR](#)

| Effect size | Small | Medium | Large | Range |
|-------------|---------------|---------------|-------------|-------|
| r | 0.10 – < 0.30 | 0.30 – < 0.50 | ≥ 0.50 | [0,1] |

robust

Test: Yuen's dependent sample trimmed means t -test

Effect size: robust (trimmed-Winsorized) standardized difference similar to Cohen's d

Function: [WRS2::dep.effect](#)

| Effect size | Small | Medium | Large | Range |
|----------------------|---------------|---------------|-------------|-------|
| $\delta_{R\delta R}$ | 0.10 – < 0.30 | 0.30 – < 0.50 | ≥ 0.50 | [0,1] |

Reference: - <https://CRAN.R-project.org/package=WRS2/vignettes/WRS2.pdf> - <https://journals.sagepub.com/doi/10.1177/0013164406288161>

between-subjects design

parametric

Test: Student's and Welch's independent samples t -test

Effect size: Cohen's d , Hedge's g

Function: [effectsize::cohens_d](#)

| Effect size | Small | Medium | Large | Range |
|-------------|-------|--------|-------|------------|
| Cohen's d | 0.20 | 0.50 | 0.80 | [-Inf,Inf] |
| Hedge's g | 0.20 | 0.50 | 0.80 | [-Inf,Inf] |

non-parametric

Test: Two-sample Mann–Whitney U Test

Effect size: $r_r (= Z/(\sqrt{N_{\text{obs}}})Z/(N_{\text{obs}}))$

Function: [rcompanion::wilcoxonR](#)

| Effect size | Small | Medium | Large | Range |
|-------------|---------------|---------------|-------------|-------|
| r | 0.10 – < 0.30 | 0.30 – < 0.50 | ≥ 0.50 | [0,1] |

Reference: https://rcompanion.org/handbook/F_04.html

robust

Test: Yuen's independent sample trimmed means t -test

Effect size: Explanatory measure of effect size ($\xi\xi$)

Function: [WRS2::yuen.effect.ci](#)

| Effect size | Small | Medium | Large | Range |
|-------------|---------------|---------------|-------------|-------|
| $\xi\xi$ | 0.10 – < 0.30 | 0.30 – < 0.50 | ≥ 0.50 | [0,1] |

Reference: <https://CRAN.R-project.org/package=WRS2/vignettes/WRS2.pdf>

One-way ANOVAs

within-subjects design

parametric

Test: Fisher's repeated measures one-way ANOVA

Effect size: η^2_{p} , ω^2

Function: `effectsize::eta_squared` and `effectsize::omega_squared`

| Effect size | Small | Medium | Large | Range |
|---------------------|---------------|---------------|-------------|-------|
| ω^2 | 0.01 – < 0.06 | 0.06 – < 0.14 | ≥ 0.14 | [0,1] |
| η^2_{p} | 0.01 – < 0.06 | 0.06 – < 0.14 | ≥ 0.14 | [0,1] |

Reference:

- <http://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize>
- <http://www.psy.gla.ac.uk/~steve/best/effect.html>

non-parametric

Test: Friedman's rank sum test

Effect size: Kendall's W

Function: `rcompanion::kendallW`

In the following table, k is the number of treatments, groups, or things being rated.

| k | Small | Medium | Large | Range |
|---------|--------|---------------|-------------|-------|
| $k = 3$ | < 0.10 | 0.10 – < 0.30 | ≥ 0.30 | [0,1] |
| $k = 5$ | < 0.10 | 0.10 – < 0.25 | ≥ 0.25 | [0,1] |
| $k = 7$ | < 0.10 | 0.10 – < 0.20 | ≥ 0.20 | [0,1] |
| $k = 9$ | < 0.10 | 0.10 – < 0.20 | ≥ 0.20 | [0,1] |

robust

Test: Heteroscedastic one-way repeated measures ANOVA for trimmed means

Effect size: Not available

between-subjects design

parametric

Test: Fisher's or Welch's one-way ANOVA

Effect size: η^2 , η^2_{pnp2} , ω^2 , ω^2_{pwp2}

Function: [effectsize::eta_squared](#) and [effectsize::omega_squared](#)

| Effect size | Small | Medium | Large | Range |
|--------------------------|---------------|---------------|-------------|-------|
| η^2 | 0.01 – < 0.06 | 0.06 – < 0.14 | ≥ 0.14 | [0,1] |
| ω^2 | 0.01 – < 0.06 | 0.06 – < 0.14 | ≥ 0.14 | [0,1] |
| η^2_{pnp2} | 0.01 – < 0.06 | 0.06 – < 0.14 | ≥ 0.14 | [0,1] |
| ω^2_{pwp2} | 0.01 – < 0.06 | 0.06 – < 0.14 | ≥ 0.14 | [0,1] |

Reference:

- <http://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize>
- <http://www.psy.gla.ac.uk/~steve/best/effect.html>

non-parametric

Test: Kruskal–Wallis test

Effect size: ϵ^2

Function: [rcompanion::epsilonSquared](#)

| Effect size | Small | Medium | Large | Range |
|--------------|---------------|---------------|-------------|-------|
| ϵ^2 | 0.01 – < 0.08 | 0.08 – < 0.26 | ≥ 0.26 | [0,1] |

Reference: https://rcompanion.org/handbook/F_08.html

robust

Test: Heteroscedastic one-way ANOVA for trimmed means

Effect size: Explanatory measure of effect size (ξ)

Function: [WRS2::t1way](#)

| Effect size | Small | Medium | Large | Range |
|-------------|---------------|---------------|-------------|-------|
| ξ | 0.10 – < 0.30 | 0.30 – < 0.50 | ≥ 0.50 | [0,1] |

Reference: <https://CRAN.R-project.org/package=WRS2/vignettes/WRS2.pdf>

Contingency table analyses

association test - unpaired

Test: Pearson's χ^2 -squared test

Effect size: Cramér's V

Function: [rcompanion::cramerV](#)

In the following table, k is the minimum number of categories in either rows or columns.

| k | Small | Medium | Large | Range |
|---------|-----------------|-----------------|-------------|-------|
| $k = 2$ | $0.10 - < 0.30$ | $0.30 - < 0.50$ | ≥ 0.50 | [0,1] |
| $k = 3$ | $0.07 - < 0.20$ | $0.20 - < 0.35$ | ≥ 0.35 | [0,1] |
| $k = 4$ | $0.06 - < 0.17$ | $0.17 - < 0.29$ | ≥ 0.29 | [0,1] |

Reference: https://rcompanion.org/handbook/H_10.html

association test - paired

Test: McNemar's test

Effect size: Cohen's g

Function: [rcompanion::cohenG](#)

| Effect size | Small | Medium | Large | Range |
|-------------|-----------------|-----------------|-------------|-------|
| Cohen's g | $0.05 - < 0.15$ | $0.15 - < 0.25$ | ≥ 0.25 | [0,1] |

Reference: https://rcompanion.org/handbook/H_05.html

goodness-of-fit test

Test: Pearson's χ^2 -squared goodness-of-fit test

Effect size: Cramér's V

Function: `rcompanion::cramerVFit`

In the following table, k is the number of categories.

| k | Small | Medium | Large | Range |
|----------|-------------------|-------------------|--------------|-------|
| $k = 2$ | $0.100 - < 0.300$ | $0.300 - < 0.500$ | ≥ 0.500 | [0,1] |
| $k = 3$ | $0.071 - < 0.212$ | $0.212 - < 0.354$ | ≥ 0.354 | [0,1] |
| $k = 4$ | $0.058 - < 0.173$ | $0.173 - < 0.289$ | ≥ 0.289 | [0,1] |
| $k = 5$ | $0.050 - < 0.150$ | $0.150 - < 0.250$ | ≥ 0.250 | [0,1] |
| $k = 6$ | $0.045 - < 0.134$ | $0.134 - < 0.224$ | ≥ 0.224 | [0,1] |
| $k = 7$ | $0.043 - < 0.130$ | $0.130 - < 0.217$ | ≥ 0.217 | [0,1] |
| $k = 8$ | $0.042 - < 0.127$ | $0.127 - < 0.212$ | ≥ 0.212 | [0,1] |
| $k = 9$ | $0.042 - < 0.125$ | $0.125 - < 0.209$ | ≥ 0.209 | [0,1] |
| $k = 10$ | $0.041 - < 0.124$ | $0.124 - < 0.207$ | ≥ 0.207 | [0,1] |

Reference: https://rcompanion.org/handbook/H_03.html

Correlation analyses

parametric

Test: Pearson product-moment correlation coefficient

Effect size: Pearson's correlation coefficient (r)

Function: [correlation::correlation](#)

| Effect size | Small | Medium | Large | Range |
|---------------|-----------------|-----------------|-------------|----------|
| Pearson's r | $0.10 - < 0.30$ | $0.30 - < 0.50$ | ≥ 0.50 | $[-1,1]$ |

non-parametric

Test: Spearman's rank correlation coefficient

Effect size: Spearman's rank correlation coefficient ($\rho\rho$)

Function: [correlation::correlation](#)

| Effect size | Small | Medium | Large | Range |
|-----------------------|-----------------|-----------------|-------------|----------|
| Spearman's $\rho\rho$ | $0.10 - < 0.30$ | $0.30 - < 0.50$ | ≥ 0.50 | $[-1,1]$ |

robust

Test: Percentage bend correlation coefficient

Effect size: Percentage bend correlation coefficient ($\rho\rho b\rho\rho b$)

Function: [correlation::correlation](#)

| Effect size | Small | Medium | Large | Range |
|------------------------|-----------------|-----------------|-------------|----------|
| $\rho\rho b\rho\rho b$ | $0.10 - < 0.30$ | $0.30 - < 0.50$ | ≥ 0.50 | $[-1,1]$ |

Meta-analysis parametric

Test: Parametric random-effects meta-analysis

Effect size: Regression estimate ($\beta\beta$)

Function: [metafor::rma](#)

robust

Test: Random-effects meta-analysis using a mixture of normals for the random effect

Effect size: Regression estimate ($\beta\beta$)

Function: [metaplus::metaplus](#)

Bayesian

Test: Bayesian random-effects meta-analysis

Effect size: Regression estimate ($\beta\beta$)

Function: [metaBMA::meta_random](#)

Suggestions

If you find any bugs or have any suggestions/remarks, please file an issue on GitHub: <https://github.com/IndrajeetPatil/ggstatsplot/issues>

Session Information

For details, see- https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/session_info.html