Derivation of some of the orginal paper results

Supporting material for Standardized means difference effect size measures for planned comparisons, trend analysis and other applications of contrast analysis

mc & mp

07, 12, 2018

Aim

Here we derive some of the most important formula in the paper. In this document, equations referred to with REF. # are equations of the original paper; the number # is the reference number of the equation in the original paper.

Contrasts analysis background

A contrast is a linear combination of means whose coefficients sum up to zero, meant to estimate a particular comparison of means and test it against zero. We refer to the contrast set of coefficients as $\boldsymbol{c} = \{c_i\}$, and to the expected set of means as $\boldsymbol{\mu} = \{\mu_i\}$. The contrast coefficients (weights) are chosen such that $\sum_i c_i = 0$, with $i = \{1, .., k\}$ where k is the number of means being compared. The contrast expected value is $c\mu = \sum_i (c_i \cdot \mu_i)$. As an example, consider a simple design with two groups: the comparison of the two groups means can be carried out with a simple contrast with $\boldsymbol{c} = \{1, -1\}$, in which the contrast value is simply the expected difference between means, $c\mu = c_1\mu_1 + c_2\mu_2 = \mu_1 - \mu_2$.

A contrast defined across k means of independent groups of size n can be tested employing either an independent samples t-test or an F-test. The t-test expected value, with $k(n-1)$ degrees of freedom, is (Steiger 2004, 172 EQ 43 and p.173, EQ 46):

$$E(t_{k(n-1)}) = \frac{\sum (c_i \cdot \mu_i)}{\sigma \cdot \sqrt{\frac{\sum c_i^2}{n}}} \tag{REF. 1}$$

The F-test associated with a contrast is simply $F_{1,k(n-1)} = t^2_{k(n-1)}$.

Cohen's δ measures for contrasts

Cohen (1988) defines several indices of effect size for the comparison of two means. In the context of two-groups designs, he defines:

$$\delta = \frac{\mu_1 - \mu_2}{\sigma} \tag{REF. 2}$$

When the same logic is applied to a contrast comparison, it naturally generalizes to (cf. Steiger 2004)

$$\delta_0 = \frac{\sum (c_i \cdot \mu_i)}{\sigma} \tag{REF. 3}$$

Steiger (2004), pag. 173, EQ 46, defines

$$E_s = \frac{\sum c_i \mu_i}{\sigma} \tag{1}$$

which is δ_0.

Normalized contrast effect size measure

$$\delta_z = \frac{\sum (c_i \cdot \mu_i)}{\sigma \sqrt{\sum c_i^2}} \tag{REF. 4}$$

If we set $z = 1/\sqrt{\sum c_i^2}$, we obtain:

$$\delta_z = z \cdot \delta_0 \tag{REF. 5}$$

Result in the paper

$$f = \frac{1}{\sqrt{k}} \cdot \delta_z \tag{REF. 8}$$

Derivation

Cohen (1988), p. 274 (Maxwell and Delaney 2000, EQ. 89) defines f as the ratio between the expected standard deviation of the means and the error standard deviation:

$$f = \frac{\sigma_\mu}{\sigma}$$

in our case, we are looking for:

$$f = \frac{\sigma_\psi}{\sigma}$$

where ψ is any contrast of interest. Faul et al. (2007), p. 29 defines

$$\hat{\sigma_\psi} = \sqrt{\frac{(\sum c_i \hat{\mu_i})^2}{N \sum \frac{c_i^2}{n}}}$$

since $N = kn$, we have

$$\sigma_\psi = \sqrt{\frac{(\sum c_i \mu_i)^2}{k \sum c_i^2}}$$

thus,

$$\sigma_\psi = \frac{|\sum c_i \mu_i|}{\sqrt{k \sum c_i^2}}$$

and

$$f = \frac{\sigma_\psi}{\sigma} = \frac{|\sum c_i \mu_i|}{\sqrt{k \sum c_i^2}} = \frac{1}{\sqrt{k}} \cdot \delta_z$$

■

However, it is interesting to show the same result without dealing with sample estimates. Consider a regression with $k-1$ contrasts, one is the contrast of interest ψ, and the others, ϕ_i, are contrasts orthogonal to each other and to the contrast of interest.
The total expected variance of the y variable will be:

$$\sigma_t^2 = \sigma_\psi^2 + \sum \sigma_{\psi_i}^2 + \sigma^2$$

because of the orthogonality of the contrasts. The population b coefficient associated with ψ is:

$$b_\psi = \frac{\sum c_i \mu_i / k}{\sum c_i^2 / k} = \frac{\sum c_i \mu_i}{\sum c_i^2}$$

and the variance examplained by the contrast ψ is given by:

$$\sigma_\psi^2 = b_\psi^2 \cdot \sigma_c^2 = \frac{(\sum c_i \mu_i)^2}{(\sum c_i^2)^2} \cdot \frac{\sum c_i^2}{k} = \frac{(\sum c_i \mu_i)^2}{\sum c_i^2} \cdot \frac{1}{k}$$

Taking the square root yields:

$$\sigma_\psi = \frac{1}{\sqrt{k}} \cdot \frac{|\sum c_i \mu_i|}{\sqrt{\sum c_i^2}}$$

because we were looking for the $f = \sigma_\psi / \sigma$ ratio, we obtain

$$f = \frac{\sigma_\psi}{\sigma} = \frac{1}{\sqrt{k}} \cdot \frac{\sum c_i \mu_i}{\sigma \sqrt{\sum c_i^2}} = \frac{1}{\sqrt{k}} \cdot |\delta_z|$$

■

Paper result

$$\eta^2 = \frac{\delta_z^2}{\delta_z^2 + k} \tag{REF. 9}$$

Derivation

$$\eta^2 = \frac{\sigma_\psi^2}{\sigma_\psi^2 + \sigma^2}$$

we have shonw that

$$\sigma_\psi^2 = \frac{1}{k} \cdot \frac{(\sum c_i \mu_i)^2}{\sum c_i^2}$$

thus

$$\eta^2 = \frac{\frac{1}{k} \cdot \frac{(\sum c_i \mu_i)^2}{\sum c_i^2}}{\frac{1}{k} \cdot \frac{(\sum c_i \mu_i)^2}{\sum c_i^2} + \sigma^2} =$$

$$= \frac{\frac{1}{k} \cdot \frac{(\sum c_i \mu_i)^2}{\sum c_i^2}}{\frac{(\sum c_i \mu_i)^2 + k\sigma^2 \sum c_i^2}{k \sum c_i^2}} =$$

$$= \frac{(\sum c_i \mu_i)^2}{(\sum c_i \mu_i)^2 + k\sigma^2 \sum c_i^2} =$$

let's devide numerator and denominator by $\sigma^2 \sum c_i^2$

$$= \frac{\frac{(\sum c_i \mu_i)^2}{\sigma^2 \sum c_i^2}}{\frac{(\sum c_i \mu_i)^2}{\sigma \sum c_i^2} + \frac{k\sigma^2 \sum c_i^2}{\sigma^2 \sum c_i^2}} =$$

$$= \frac{\delta_z^2}{\delta_z^2 + k}$$

■

Scaled effect size measure

A different method of scaling the constrat effect size measure which guarantees better interpretability and comparability can be suggested. Let's $g = \frac{2}{\sum_i |c_i|}$, where $|c_i|$ indicates the absolute value of c_i, then

$$\delta_g = g \cdot \delta_0 = \frac{2}{\sum |c_i|} \cdot \frac{\sum_i c_i \cdot \mu_i}{\sigma} \qquad \text{(REF. 10)}$$

Paper results

All the derivations for δ_z can be derived for δ_g by noticing that

$$\delta_z = \frac{z}{g}\delta_g$$

thus, one simply substitutes $\frac{z}{g}\delta_g$ is place of δ_z. As an example, we derive here:

$$\eta_p^2 = \frac{\delta_g^2}{\delta_g^2 + (\frac{g^2}{z^2} \cdot k)} \qquad \text{(REF. 16)}$$

Derivation

Recall that

$$= \frac{\delta_z^2}{\delta_z^2 + k}$$

and substitutes $\frac{z}{g}\delta_g$ is place of δ_z:

$$\eta_p^2 = \frac{\frac{z^2}{g^2}\delta_g^2}{\frac{z^2}{g^2}\delta_g^2 + k} =$$

$$= \frac{z^2\delta_g^2}{z^2\delta_g^2 + g^2k}$$

we now divide numerator and denominator by z^2 and we obtain:

$$= \frac{\delta_g^2}{\delta_g^2 + \frac{g^2}{z^2}k}$$

■

Comparability of δ_g

Paper result

In the paper we assert that "if one group has $\mu_1 = d$ and all other groups share the same mean $\mu_{2..k} = 0$, and the comparison is tested with the contrast $\boldsymbol{c} = \{(k-1), -1_2, -1_3, .., -1_k\}$, the contast value is $c\mu = (k-1) \cdot d$. The g-method yields, after simple algebra, $g = 1/(k-1)$, and thus $\delta_{gk} = \delta_{02}$, independently of the number of groups k."

Derivation

Let $\boldsymbol{\mu} = \{d, 0, 0, .., 0\}$ and $\boldsymbol{c} = \{(k-1), -1, -1, .., -1\}$, such that $\sum c_i = (k-1) + (k-1) \cdot (-1) = 0$. The sum of the absolute values of c_i is $\sum |c_i| = (k-1) + (k-1) \cdot |(-1)| = 2 \cdot (k-1)$ and thus $g = 2/2(\cdot k-1) = 1/(k-1)$. Then,

$$d_{gk} = \frac{1}{k-1} \cdot (k-1)d = d$$

, for any value of k.

■

Sample Estimation

$$d_z = \frac{\sum (c_i \cdot m_i)}{s_p \sqrt{\sum c_i^2}} \tag{2}$$

and

$$d_g = 2 \cdot \frac{\sum (c_i \cdot m_i)}{s_p \sum |c_i|} \tag{3}$$

Paper result

$$d_z = \frac{t_{k(n-1)}}{\sqrt{n}} \tag{REF. 20}$$

Derivation

Recall that

$$t_{k(n-1)} = \sqrt{n} \cdot d_z \tag{REF. 7}$$

thus

$$d_z = \frac{t_{k(n-1)}}{\sqrt{n}}$$

■

Paper result

$$\delta_z = \sqrt{k \cdot \frac{n-1}{n}} \cdot \frac{r_{contrast}}{\sqrt{1 - r_{contrast}^2}} \tag{REF. 22}$$

Derivation

Following (Rosenthal, Rosnow, and Rubin 2000), pag. 38, EQ 3.2

$$F_{contrast} = \frac{MS_{contrast}}{MS_{within}}$$

and

$$r_{contrast} = \sqrt{\frac{F_{contrast}}{F_{contrast} + k(n-1)}}$$

thus

$$r_{contrast} = \sqrt{\frac{t^2_{contrast}}{t^2_{contrast} + k(n-1)}}$$

therefore

$$r_{contrast} = \sqrt{\frac{n \cdot d_z^2}{n \cdot d_z^2 + k(n-1)}}$$

by dividing num and den by n, we obtain:

$$r_{contrast} = \sqrt{\frac{d_z^2}{d_z^2 + \frac{k(n-1)}{n}}}$$

which means:

$$r^2_{contrast} = \frac{d_z^2}{d_z^2 + \frac{k(n-1)}{n}}$$

$$r^2_{contrast}(d_z^2 + \frac{k(n-1)}{n}) = d_z^2$$

$$r^2_{contrast} d_z^2 + r^2_{contrast} \frac{k(n-1)}{n} = d_z^2$$

$$r^2_{contrast} \frac{k(n-1)}{n} = d_z^2 - r^2_{contrast} d_z^2$$

$$r^2_{contrast} \frac{k(n-1)}{n} = d_z^2(1 - r^2_{contrast})$$

$$\frac{k(n-1)}{n} \frac{\cdot r^2_{contrast}}{1 - r^2_{contrast}} = d_z^2$$

$$d_z = \sqrt{\frac{k(n-1)}{n}} \cdot \frac{r_{contrast}}{\sqrt{1 - r^2_{contrast}}}$$

■

Power Analysis

Let λ_t be the noncentrality parameter of the $t_{k(n-1)}$ distribution associated with the t-test for d_z.

Result

$$\lambda_t = \sqrt{n} \cdot d_z \qquad \text{(REF. 27)}$$

Following Liu (2013), page 87, EQ 4.25,

$$\lambda_t = \frac{\sum c_i \mu_i}{\sqrt{\sigma^2(\sum \frac{c_i^2}{n_i})}}$$

which simplifies for equal cell size to:

$$\lambda_t = \sqrt{n} \cdot \frac{\sum c_i \mu_i}{\sqrt{\sigma^2 \cdot \sum c_i^2}}$$

and thus

$$\lambda_t = \sqrt{n} \cdot d_z$$

■

Paper result

The non-centrality parameter of the F-test is given by:

$$\hat{\lambda_F} = k \cdot n \cdot f^2 \qquad \text{(REF. 30)}$$

Derivation

Note that this is not a result of the paper, since it is taken from Steiger (2004), p. 169 EQ 10 or Cohen (1988), p. 481, Steiger and Fouladi (1997), p. 246 EQ 9.95. However, it is interesting to derive it from REF. 27.

First, notice that for contrasts with 1 numerator df, $\lambda_F = \lambda_t^2$. Therefore

$$\lambda_F = n \cdot d_z^2$$

we can multiply and divide by k, and we get

$$\lambda_F = \frac{n \cdot k}{k} \cdot d_z^2$$

since $f = d_z/\sqrt{k}$, we reach:

$$\lambda_F = n \cdot k \cdot f^2$$

■

Confidence intervals

All the formulas in this sections are derived directly from the literature. We only want to check that

$$Pr\left[\frac{2}{\sum |c_i|} \cdot \sqrt{\frac{\sum c_i^2}{n}} \cdot \hat{\lambda}_l \leq d_g \leq \frac{2}{\sum |c_i|} \cdot \sqrt{\frac{\sum c_i^2}{n}} \cdot \hat{\lambda}_l\right] = 1 - \alpha \qquad \text{(REF. 34)}$$

can be derived from

$$Pr\left[\frac{\hat{\lambda_l}}{\sqrt{n}} \leq d_z \leq \frac{\hat{\lambda_u}}{\sqrt{n}}\right] = 1 - \alpha \qquad \text{(REF. 33)}$$

by substituing d_z with d_g.

Derivation

Recall that $d_g = \frac{g}{z} d_z$. Thus, for d_g we have:

$$Pr\left[\frac{g}{z}\frac{\hat{\lambda_l}}{\sqrt{n}} \leq \frac{g}{z} d_z \leq \frac{g}{z}\frac{\hat{\lambda_u}}{\sqrt{n}}\right] = 1 - \alpha$$

$$Pr\left[g \cdot z^{-1}\frac{\hat{\lambda_l}}{\sqrt{n}} \leq d_g \leq g \cdot z^{-1}\frac{\hat{\lambda_u}}{\sqrt{n}}\right] = 1 - \alpha \tag{REF. 33}$$

$$Pr\left[\frac{2}{\sum |c_i|} \cdot z^{-1}\frac{\hat{\lambda_l}}{\sqrt{n}} \leq d_g \leq \frac{2}{\sum |c_i|} \cdot z^{-1}\frac{\hat{\lambda_u}}{\sqrt{n}}\right] = 1 - \alpha$$

$$Pr\left[\frac{2}{\sum |c_i|} \cdot \sqrt{\sum c_i^2}\frac{\hat{\lambda_l}}{\sqrt{n}} \leq d_g \leq \frac{2}{\sum |c_i|} \cdot \sqrt{\sum c_i^2}\frac{\hat{\lambda_u}}{\sqrt{n}}\right] = 1 - \alpha$$

$$Pr\left[\frac{2}{\sum |c_i|} \cdot \sqrt{\frac{\sum c_i^2}{n}} \cdot \hat{\lambda}_l \leq d_g \leq \frac{2}{\sum |c_i|} \cdot \sqrt{\frac{\sum c_i^2}{n}} \cdot \hat{\lambda}_l\right] = 1 - \alpha$$

■

References

Cohen, Jacob. 1988. "Statistical Power Analysis for the Behavioral Sciences . Hilsdale." *NJ: Lawrence Earlbaum Associates* 2.

Faul, Franz, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. "G* Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences." *Behavior Research Methods* 39 (2). Springer: 175–91.

Liu, Xiaofeng Steven. 2013. *Statistical Power Analysis for the Social and Behavioral Sciences: Basic and Advanced Techniques*. Routledge.

Maxwell, Scott E, and Harold D Delaney. 2000. *Designing Experiments and Analyzing Data: A Model Comparison Perspective*. NJ: Mahwah, Lawrence Erlbaum Associates.

Rosenthal, Robert, Ralph L Rosnow, and Donald B Rubin. 2000. *Contrasts and Effect Sizes in Behavioral Research: A Correlational Approach*. Cambridge University Press.

Steiger, James H. 2004. "Beyond the F Test: Effect Size Confidence Intervals and Tests of Close Fit in the Analysis of Variance and Contrast Analysis." *Psychological Methods* 9 (2). American Psychological Association: 164.

Steiger, James H, and Rachel T Fouladi. 1997. "Noncentrality Interval Estimation and the Evaluation of Statistical Models." In *What If There Were No Significance Tests?: Classic Edition*, edited by Lisa Lavoie Ed Harlow, Stanley A Mulaik, and James H Steiger. Lawrence Erlbaum Associates Publishers.