Design Analysis with Bayesian Factor

1 Effect Size

Consider n_1 observations $\mathbf{x}_1$ from a $\mathcal{N}(0,1)$ distribution, and n_2 observations $\mathbf{x}_2$ from a $\mathcal{N}(\theta,1)$ distribution. We want to evaluate the effect size θ by considering Cohen's d, Hedges' g (unbiased estimator of θ derived by correcting Cohen's d by a multiplicative factor) or Glass' Δ (using only of the standard deviation of the control group). Moreover, we want to estimate the following quantities:

- β = prob. of true positives (power);
- ϵ_S = prob. that a significant effect is estimated in the wrong direction;
- ϵ_M = average overestimation of an effect that emerges as significant;
- ϵ_N = prob. of false negatives;
- ϵ_I = prob. of an inconclusive result when the effect is significant;

Table 1 shows the possible effect sizes e. Let Z be the effect scaled by an appropriate multiplicative factor; then Z follows a noncentral t-distribution $T_{\nu,\theta N}$, with ν degrees of freedom and noncentrality parameter θN, where

$$N = \sqrt{\frac{n_1 n_2}{n_1 + n_2}}.$$

Table 1: Effect sizes.

	Effect e	Scaled effect Z	Degrees of freedom ν
Cohen	$\frac{\bar{X}_2-\bar{X}_1}{s}$	eN	n_1+n_2-2
Hedges	$\left(1-\frac{3}{4\nu-1}\right)\frac{\bar{X}_2-\bar{X}_1}{s}$	$eN\left(1-\frac{3}{4\nu-1}\right)^{-1}$	n_1+n_2-2
Glass	$\frac{\bar{X}_2-\bar{X}_1}{s_1}$	eN	n_1-1

2 Frequentist Approach

Consider the null hypothesis $H_0 : \theta = 0$, and the alternative hypothesis $H_1 : \theta \neq 0$. Under H_0, Z follows a centered t-distribution T_ν with ν degrees of freedom. The p-value corresponding to an observed value z is

$$\text{p-value} = P(Z > |z|).$$

For a given statistical significance α (e.g. $\alpha = 0.05$):

- if p-value $\geq \alpha$, then the evidence supports H_0;
- if p-value $< \alpha$, then the evidence supports H_1.

Moreover:

$$\begin{aligned} \beta &= \\ e_S &= & e_M &= \\ e_N &= 1 - \beta & e_I &= 0. \end{aligned}$$

3 Bayesian Approach

Consider the null hypothesis $H_0 : \theta = 0$, and the alternative hypothesis $H_1 : \theta \sim F$, where F is a distribution (analysis prior) on a set D containing zero. For simplicity, assume that F is continuous with density f. Then the Bayes factor is

$$B = \frac{P(z \,|\, H_1)}{P(z \,|\, H_0)} = \frac{\int_D P(z \,|\, \theta = w)\, f(w)\, \mathrm{d}w}{P(z \,|\, \theta = 0)} = \frac{\int_D t_{\eta, wN}(z)\, f(w)\, \mathrm{d}w}{t_{\eta,0}(z)}$$

where $t_{\eta,\theta N}$ denotes the pdf of Z.

For a given threshold k (e.g. $k = 3$):

- if $B < 1/k$, then the evidence supports H_0;
- if $1/k \leq B \leq k$, then the analysis is inconclusive;
- if $B > k$, then the evidence supports H_1.

Moreover:

$$\begin{aligned} \beta &= P(B > k \,|\, H_1) \\ e_S &= P(B > k,\ \mathrm{sgn}(e) \neq \mathrm{sgn}(\theta) \,|\, H_1,\ B > k) & e_M &= \frac{\mathrm{mean}\{|e| \,:\, B > k\}}{\theta} \\ e_N &= P\left(B < \frac{1}{k} \,|\, H_1\right) & e_I &= P\left(\frac{1}{k} \leq B \leq k \,|\, H_1\right). \end{aligned}$$