Simulative Approach to Study the Operating Characteristics of a Testing Procedure

2 Ks size & power

From a frequentist perspective, once we jot down a system of statistical hypotheses like

$$H_0: \theta \in \Theta_0 \quad \text{vs} \quad H_1: \theta \notin \Theta_0$$

(null) (alternative)

and we design a suitable test statistic (from data), namely T_n, so that we Reject H_0 if $T_n > c_\alpha$ for some well tuned critical level c_α we are interested in studying:

SIZE: $\alpha = \Pr(\text{False Discovery}) = \Pr(\text{Reject } H_0 \mid H_0 \text{ TRUE})$

Type I error

POWER: $1-\beta = \Pr(\text{Correct Discovery}) = \Pr(\text{Reject } H_0 \mid H_0 \text{ FALSE})$

IF we are unable to study these quantities analytically, we can set up a suitable simulation study.

$\alpha \approx$ proportion of rejections we get by applying our test on M dataset generated from a model "COMPATIBLE" WITH H_0

$1-\beta \approx$ proportion of rejections we get by applying our test on M dataset generated from a model "NOT COMPATIBLE" WITH H_0

the "closer" to H_0 the more difficult to get a rejection...