Programs SG.EXE, PEG.EXE and SDG.EXE

Considerer $x_i \sim B(n_i, p_i)$ where:

- **i = 1** for the **standard** treatment
- **i = 2** for the **new** treatment

The aim is to perform inferences about $d = p_2 - p_1$. The probability of an experimental (x_1, x_2) is:

$$P(x_1, x_1) = \binom{n_1}{x_1}\binom{n_2}{x_2} p_1^{x_1}(1-p_1)^{y_1}\, p_2^{x_2}(1-p_2)^{y_2}$$

Under the null hypothesis H: $d = \delta$ (where $-1<\delta<+1$), if $p_1=p$ then $p_2=p+\delta$ and so (1) is:

$$P(x_1, x_2|p, \delta)=\binom{n_1}{x_1}\binom{n_2}{x_2}\, p^{x_1}(1-p)^{n_1-x_1}(p+\delta)^{x_2}(1-p-\delta)^{n_2-x_2}$$

where p is a nuisance parameter taking the values:

$$\max(0; -\delta) \le p \le \min(1; 1-\delta)$$

For a critical region CR formed by a set of pairs (x_1, x_2), the error α of the test will be $\alpha(p|\delta)=\Sigma_{CR}\, P(x_1, x_2|p, \delta)$, and the size of the test will be $\alpha^*(\delta) = \mathrm{Max}_p\, \alpha(p|\delta)$.

There are various ways of obtaining the CR (some of which appear when executing the programs). The most traditional is that based on the order given by the *z*-pooled statistic without a continuity correction (order Z in the program). The generally most powerful test is based on *z* with a continuity correction (order ZC or ZY in the program). The expressions are:

$$Z_c^2(\delta) = \frac{\left\{|\hat{d}-\delta| - c\right\}^2}{V_1+V_2} \text{ with } \begin{cases} V_1 = \hat{p}(1-\hat{p})/n_1 \\ V_2 = (\hat{p}+\delta)(1-\hat{p}-\delta)/n_2 \end{cases}$$

where $\hat{d} = \hat{p}_2 - \hat{p}_1$, $\hat{p}_i = x_i/n_i$, $\hat{p}$ is the maximum likelihood estimator of p and it is understood that $Z_c^2(\delta)=0$ when $|\hat{d}-\delta| \le c$. When c=0 one obtains the statistic Z. When $c=1/N$ if $n_1 \ne n_2$, $c=2/N$ if $n_1=n_2$, with $N=[(n_1+1)(n_2+1)]$, one obtains the statistic ZC. When $c=n/(2n_1n_2)$ –with $n=n_1+n_2$– one obtains the statistic ZY (Yates's statistics).

The exact p-value $\alpha^* = \mathrm{Max}_{\delta\in H}\, \alpha^*(\delta)$ of the observed pair (x_1, x_2) depends on the null hypothesis H and alternative hypothesis K to be demonstrated, but if $\hat{d}$ -Delta(vero) in the program- $\in$ H, , then p-value = 1.

There are 3 programs:

Program SG.EXE

- Case **SG**: H: $d \le \delta$ vs. K: $d > \delta$ (Superiority Generalized). In particular:
  - ∗ When $\delta = 0$: is the classic case **S** of Superiority.
  - ∗ When $\delta < 0$: is the case **NI** of Non-Inferiority.
  - ∗ When $\delta > 0$: is the case **SS** of Substantial-Superiority.
- Case **IG**: H: $d \ge \delta$ vs. K: $d < \delta$ (Inferiority Generalized).

Program PEG.EXE

H: $d \le \delta_1$ or $d \ge \delta_2$ vs. K: $\delta_1 < d < \delta_2$, where $\delta_1 < \delta_2$ and $-1 < \delta_i < +1$, is the case **PEG**. In particular, when $\delta_2 = -\delta_1 = \Delta > 0$:

H: $|d| \ge \Delta$ vs. K: $|d| < \Delta$ is the case **PE** (Practice-Equality or Equivalence).

Program SDG.EXE

H: $\delta_1 \le d \le \delta_2$ vs. K: $d < \delta_1$ or $d > \delta_2$, where $\delta_1 \le \delta_2$ and $-1 < \delta_i < +1$, is the case **SDG**. In particular:

- When $\delta_1 = \delta_2 = 0$: H: $d=0$ vs. K: $d \ne 0$ is the classic case **D** of null Difference.
- When $\delta_1 = \delta_2 = \delta$: H: $d = \delta$ vs. K: $d \ne \delta$ is the case **SG2** of non-null Difference.
- When $\delta_2 = -\delta_1 = \Delta > 0$: H: $|d| \le \Delta$ vs. K: $|d| > \Delta$: is the case **SD** of Substantially-Difference.

∗∗∗

The cases of greatest practical interest for the problems of equivalence between proportions are those which have as their alternative hypotheses:

SG (NI): K: $d > -\Delta$, **SG (SS)**: K: $d > \Delta$, **PE**: K: $|d| < \Delta$, **SD**: K: $|d| > \Delta$

where $\Delta \ge 0$ in all cases ($\Delta > 0$ in PE case).

∗∗∗

For more details see:

Martín Andrés, A. and Herranz Tejedor, I. (2004). Exact unconditional non-classical tests on the difference of two proportions. *Computational Statistics and Data Analysis 45 (2)*, 373-388.

Herranz Tejedor, I. and Martín Andrés, A. (2008). A numerical comparison of several unconditional exact tests in problems of equivalence based on the difference of proportions. *Journal of Statistical Computation and Simulation 78 (11)*, 969-981.