

## FORMULEBLAD

**Vuistregels voor de grootte van het verschil van twee groepen**

$$2 \times 2 \text{ kruistabel } \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \text{ met } \phi = \frac{ad - bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}},$$

waarin  $a$ ,  $b$ ,  $c$  en  $d$  absolute aantallen zijn

- als  $\phi < -0,4$  of  $\phi > 0,4$ , dan zeggen we “het verschil is groot”,
- als  $-0,4 \leq \phi < -0,2$  of  $0,2 < \phi \leq 0,4$ , dan zeggen we “het verschil is matig”,
- als  $-0,2 \leq \phi \leq 0,2$ , dan zeggen we “het verschil is gering”.

Maximaal verschil in cumulatief percentage ( $\max V_{cp}$ )

(met voor beide groepen een steekproefomvang  $n > 100$ )

- als  $\max V_{cp} > 40$ , dan zeggen we “het verschil is groot”,
- als  $20 < \max V_{cp} \leq 40$ , dan zeggen we “het verschil is matig”,
- als  $\max V_{cp} \leq 20$ , dan zeggen we “het verschil is gering”.

Effectgrootte  $E = \frac{\bar{X}_1 - \bar{X}_2}{\frac{1}{2}(S_1 + S_2)}$ , met  $\bar{X}_1$  en  $\bar{X}_2$  de steekproefgemiddelden

( $\bar{X}_1 \geq \bar{X}_2$ ),  $S_1$  en  $S_2$  de steekproefstandaardafwijkingen

- als  $E > 0,8$ , dan zeggen we “het verschil is groot”,
- als  $0,4 < E \leq 0,8$ , dan zeggen we “het verschil is matig”,
- als  $E \leq 0,4$ , dan zeggen we “het verschil is gering”.

Twee boxplots vergelijken

- als de boxen<sup>1)</sup> elkaar niet overlappen, dan zeggen we “het verschil is groot”,
- als de boxen elkaar wel overlappen en een mediaan van een boxplot buiten de box van de andere boxplot ligt, dan zeggen we “het verschil is matig”,
- in alle andere gevallen zeggen we “het verschil is gering”.

noot 1 De ‘box’ is het interval vanaf het eerste kwartiel tot en met het derde kwartiel.

**Betrouwbaarheidsintervallen**

Het 95%-betrouwbaarheidsinterval voor de populatieproportie is  
 $p \pm 2 \cdot \sqrt{\frac{p(1-p)}{n}}$ , met  $p$  de steekproefproportie en  $n$  de steekproefomvang.

Het 95%-betrouwbaarheidsinterval voor het populatiegemiddelde is  
 $\bar{X} \pm 2 \cdot \frac{S}{\sqrt{n}}$ , met  $\bar{X}$  het steekproefgemiddelde,  $n$  de steekproefomvang en  
 $S$  de steekproefstandaardafwijking.