

Svar

Därmed torde flertalet missförstånd vara utredda. Den enda åsiktsskillnaden i sak gäller situationer där man inte testar en vetenskapligt uppställd hypotes utan en persons påstående att hon eller han har en märklig förmåga som inte är erkänd av vetenskapen. (Detta behandlas under Missförstånd 3). Om man i ett sådant fall använder gängse principer för prövning av vetenskapliga hypoteser, inklusive det kritiska p -värdet 0,05, kommer cirka var tjugonde person som testas för en sådan förmåga att i efterhand kunna säga: "Min förmåga har konstaterats med statistisk signifikans i en vetenskaplig undersökning." Detta är tyvärr en form av signifikansmissbruk som förekommer i praktiken.

Problemet är att gängse testmetoder är "kalibrerade" för prövning av vetenskapliga hypoteser som har tämligen hög trovärdighet. I denna (avsedda) tillämpning har vi anledning att förvänta oss att andelen falskt positiva utfall blir avsevärt lägre än det kritiska p -värdet (som oftast är 0,05). Men om vi i stället testar en hypotes med mycket låg trovärdighet kan andelen falskt positiva förväntas hamna nära det kritiska p -värdet. Om detta värde är 0,05 får man enligt min mening en oacceptabelt hög förväntad andel falskt positiva utfall. För att undvika detta kan man antingen rakt av sänka p -värdet eller, som jag föreslog i min ursprungliga artikel, utforma testet med utgångspunkt från den träffsäkerhet som personen i fråga själv säger sig ha. Nilsson har naturligtvis rätt i att man då inte bör utforma testet så att försökspersonen skulle ha cirka 50 % chans att lyckas om hans eller hennes egen uppfattning om träffsäkerheten är riktig. Man kan be försökspersonen uppge ett resultat som han eller hon känner sig säker på att uppnå. Utifrån denna träffsäkerhetsnivå kan man sedan, som jag föreslog i artikeln, "räkna ut hur många gånger man behöver upprepa försöket för att en iakttagen effekt av den storleken ska ha rimligt låg sannolikhet att uppstå om det inte finns någon effekt alls. Man kalibrerar då lämpligen till en betydligt strängare signifikansnivå än 5 %."

Jag håller med Nilsson om att konfidensintervall ofta är mera upplysande än signifikanstest. Statistisk styrka (power) är också underutnyttjad i många sammanhang.

Sven Ove Hansson