

Kommentarer till Sven Ove Hanssons svar på kommentarer till artikel *Signifikansmissbruket i pseudovetenskap* i nr 3 av Qvintensen 2010

Olika terminologi

Hansson menar att man vid hypotesprövning har att göra med två hypoteser, den vetenskapligt uppställda hypotesen och nollhypotesen, och att man i gängse språkbruk med den testade hypotesen avser den vetenskapligt uppställda hypotesen. Hansson har avsett denna tolkning men hans artikel handlar ju om signifikanser vid statistisk hypotesprövning och då är det väl rimligt att man håller sig till terminologin inom detta område, där det är nollhypotesen som testas. Frågan är hur Hansson tänker sig då man har ett tvåsidigt test, dvs. man vill undersöka om t.ex. en medicinsk behandling har någon effekt, positiv eller negativ. Att falsifiera att en medeffekt är skild från noll är vanligen inte helt lätt.

Tärningsexemplet

Jag har redan i mina tidigare kommentarer tagit upp möjligheten att Hansson skulle ha menat att nollhypotesen var att alla sidor var lika sannolika (det är under det antagandet han säger sig ha beräknat signifikansen). Min kritik var då att han inte bara skulle ha betraktat tre sexor utan även tre ettor, tre tvåor osv. som kritiska utfall.

Missförstånd 1: Ett uttalande om signifikans behöver inte hänvisa till en i förväg gjord förutsägelse.

Här är vi nog överens.

Missförstånd 2: Statistisk signifikans mäter sannolikheten att det man iakttar beror på slumpen.

Våra motsättningar bottenar, som Hansson tycks ha insett, i att han inte klart specificerat vad som är nollhypotesen.

Missförstånd 3: Signifikant avvikelse från nollhypotesen är ett adekvat mått på graden av vetenskaplig bevisning.

Här använder Hansson ett exempel där det gäller förmågan att avgöra könet på en person som sitter bakom ett skynke och där försökspersonen tror sig kunna förutsäga rätt med 75 % sannolikhet. Nollhypotesen tycks vara att träffprocenten är 50 % (eller mindre). Jag håller med om att det i sådana här fall, då det handlar om tvivelaktiga förmågor, är rimligt att välja en låg risk för fel av första slaget. Jag tycker dock inte det är rimligt att välja försökspersonens uppgivna träffprocent som kritisk gräns. Skulle personen ha rätt är det ju bara ungefär 50 % sannolikhet att förkasta nollhypotesen och det liknar då som Lars Ängquist påpekat mer en vadslagningssituation än en hypotesprövning. Skulle man vilja vederlägga försökspersonens påstående att om en träffsannolikhet på 75 % bör nollhypotesen vara att träffsannolikheten är minst 75 %. Väljer vi då som kritisk som kritisk gräns just 75 % blir ju sannolikheten att förkasta (signifikansnivån) ungefär 50 % då försökspersonens uppgift är korrekt och det är väl inte särskilt rimligt. Även om man är ute efter att avslöja tveksamma förmågor bör man hålla sig till principerna för hypotesprövning och inte ägna sig åt vadslagning.

I många tillämpningar av statistik bör man överväga om inte beräkning av ett konfidensintervall är mer relevant än hypotesprövning.

Jag tycker dock fortfarande inte att resonemanget kring ovanstående exempel har så mycket med formuleringen av missförståndet att göra, se mina ursprungliga kommentarer till artikeln.

11-04-11

Missförstånd 4: Ett höggradigt signifikant resultat ger ett starkt stöd åt den testade hypotesen.

Jag tror, eller åtminstone hoppas, att det för de flesta är självklart att statistisk hypotesprövning gäller statistiskt formulerade hypoteser, men om Hansson tror att ett klagörande är nödvändigt så är det väl i sin ordning.

Göran Nilsson