

Strunt är strunt ($p = 0.03$)

De senaste åren har Quartilen emellanåt publicerat inlägg i den pågående debatten om svenska språkets ställning i Sverige. I septembernumret 2005 hävdade således ett antal lärde från Uppsala att svenska fungerar ”signifikant sämre” än engelska, ett inlägg som bemöttes av undertecknad i decembernumret 2005. Den undersökning som uppenbarligen ligger till grund för de lärdes påstående är Hedda Söderlundhs D-uppsats i nordiska språk vid Uppsala universitet. Även andra har åberopat denna uppsats, till exempel Språkrådets chef Olle Josephsson som i en artikel i Svenska Dagbladet beskrev den som ”fiffig” och hävdade att dess resultat var ”entydigt”. Oavsett vad man har för åsikter i språkdebatten så är Söderlundhs uppsats av intresse för statistiker eftersom den bjuder på ett veritabelt smörgåsbord av allsköns felaktigheter som bara är alltför välbekanta för generationer av statistiklärare.

Söderlundhs syfte var att påvisa skillnaden i förståelse av svenska och engelska texter. Hon lät därför ett antal studenter läsa en samhällsvetenskaplig text, hälften av studenterna på svenska, den andra hälften på engelska. Hon ställde sedan fyra frågor (på svenska) och graderade svaren för att kunna mäta förståelsedjupet. För var och en av dessa frågor genomfördes sedan ett statistiskt test (Mann-Whitney) för att se om skillnaden mellan grupperna var signifikant. Resultatet var att endast en av dessa frågor gav en signifikant skillnad till svenskans fördel, två frågor gav skillnader till svenskans fördel som inte var signifikanta och på den fjärde frågan klarade sig till och med den engelska gruppen något bättre (dock ej signifikant). Konfronterad med dessa data vidtog Söderlundh raskt två åtgärder. Den fråga där den engelska gruppen svarat bättre slopades med argumentet att den varit ”felkonstruerad” och de övriga tre frågorna slogs ihop. Detta resulterade i en i slutänden signifikant skillnad till svenskans fördel ($p = 0.03$) tack vare att den enda signifikanta skillnad som observerats var stor nog att kompensera för de andra två. Fiffigt? Sannerligen. Entydigt? Knappast.

Söderlundh klämmer också in den statistikokunnige statistikanvändarens favoritfras ”trend i hypotesens riktning” om icke-signifikanta skillnader och om nu inte detta skulle vara nog för att ge gamla statistiklärare apoplektiska anfall så konstaterar hon dessutom glatt att det p -värde på 1% hon fick på en av sina frågor betyder att ”resultatet i testgruppen med nittionio procents sannolikhet ger uttryck för ett faktiskt förhållande”. För att inskräpa sin

feltolkning av p -värden lägger hon också till att resultatet kan med "99% sannolikhet generaliseras till att gälla en större population".

Men alldeles oavsett dessa flagranta brister så illustrerar Söderlundhs studie ett måhända djupare problem: användande av statistik för att ge självklara svar på ointressanta frågor. För visst är det självklart att den svenska gruppen skall klara sig bättre och allt ett statistiskt test kan åstadkomma är att förhoppningsvis bekräfta denna självklarhet med det heliga p -värdet. I sin roande artikel "The Earth Is Round ($p < .05$)" diskuterar Jacob Cohen den övertro på hypotesprövning som härskar inom psykologin och i en annan artikel beskriver Nigel Yoccoz ett liknande förhållande inom de biologiska vetenskaperna. För att slänga in ett mer välkänt namn så berättar Olle Häggström i *Qvartilen* 3/2004 om en obstinat pedagog som vägrade hålla med om de mest uppenbara självklarheter om de inte åtföljdes av ett p -värde. Denna p -värdesbesatthet är något som vi nog borde tillägna mer tid i statistikundervisningen men måhända är vi alltför yrkesstolta för att stå och föreläsa om något för att sedan avråda från att det används.

Med undersökningar som Söderlundhs blir det extra problematiskt eftersom de finner sin väg ut i media och används i samhällsdebatten. Den statistikokunnige sätter förmodligen instinktivt likhetstecken mellan "signifikant" och "stor" men som *Qvartilens* läsekrets vet kan ju en statistiskt signifikant skillnad vara så liten att den är ointressant ur praktisk synvinkel. Statistiska titaner som Sir David Cox och John Tukey varnade tidigt för överanvändandet av hypotestest på bekostnad av de betydligt mer informativa konfidensintervallen och i Söderlundhs undersökning är förstas det intressanta inte *om* det föreligger en skillnad utan *hur stor* den är. Och om man skärskådar Söderlundhs studie så är det markant hur liten skillnaden mellan de två grupperna är, en observation som är av betydligt större intresse än vad p -värdet är. I slutändan handlar det ju om att fatta beslut, till exempel huruvida resurser skall spenderas på ökad översättning av samhällsvetenskaplig kurslitteratur, och något sådant vore en synnerligen förhastad slutsats att dra av Söderlundhs studie som, ironiskt nog, snarare pekar i motsatt riktning.

Avslutningsvis vill jag påpeka att jag inte alls tror att Hedda Söderlundh haft för avsikt att vilseleda. Med tanke på hur ogenerat hon redovisar sina metoder misstänker jag att hon är lyckligt omedveten om sina misstag och tyvärr bara följer de gängse normerna inom sin vetenskap. Om Söderlundh har fortsatt till doktorandstudier på den inslagna vägen vill jag dock varmt rekommendera att hon tar kontakt med någon av Uppsalas utomordentliga

statistiker innan hon utför vidare statistiska undersökningar.

Referenser

Jacob Cohen (1994), The Earth Is Round ($p < .05$). *American Psychologist* 49, 997-1003

Olle Häggström (2004), Statistikern och den undermåliga vetenskapen, *Qvar-tilen* 3/2004

Olle Josephsson (2005), Djup förståelse kräver svenska texter. *Svenska Dagbladet*, http://www.svd.se/dynamiskt/kultur/did_9130405.asp

Per-Åke Lindblom (2005), Vänstern och språkfrågan – varför måste svenska språket försvaras?. *Proletären*, 2005-10-13.

Hedda Söderlundh (2004), Svårare på engelska. D-uppsats, Uppsala Universitet, <http://www.regeringen.se/content/1/c6/03/74/78/95862bf9.pdf>

Nigel Yoccoz (1991), The Use, Overuse, and Misuse of Significance Tests in Evolutionary Biology and Ecology. *Bulletin of the Ecological Society of America* 32: 106-111