

# Fällor och fel i odontologisk statistik

Några synpunkter och exempel

## Hilding Björn

En av förutsättningarna för att kunna tolka och dra slutsatser av uppgifter från exempelvis epidemiologiska undersökningar är en del basala kunskaper i statistisk bearbetning. Behovet av sådana har stigit och kommer att stiga i takt med det ökade intresset för epidemiologiska studier av tandhälsoutvecklingen såväl inom barn- och ungdomsklientelet som inom hela den vuxna befolkningen i vårt land. Våra möjligheter att förstå och värdera betydelsen av de resultat som presenteras i vetenskapliga artiklar är dessutom betydligt större om vi besitter en del kunskaper inom det statistiska ämnesområdet. I artikeln belyses några av de fallgropar, som vi kan hamna i om vi på egen hand ger oss in på områden som vi inte till fullo behärskar.

Hilding Björn är professor emeritus vid Odontologiska fakulteten, Lunds universitet.

*Motto: Statistical analysis is no substitute for common sense.*

De flesta fel som man träffar på är gemensamma för all biologisk statistik. Ett av de svåraste – och vanligaste – bygger på den vidskepliga föreställningen att det "signifikanta" resultatet av ett statistiskt test återspeglar en oomkullkastlig sanning. För det första utgör de numeriska mätvärden eller skattningar som statistiken grundas på av siffror som i själva verket endast är mer eller mindre förenklade avbilder av komplicerade biologiska företeelser och förlopp. Dessa data återger alltså endast en del av

verkligheten. Redan detta faktum manar till en försiktig attityd gentemot statistiska resultat.

En dator med ett bra program är ett utmärkt hjälpmedel. Men ingen dator kan dra slutsatser. Statistik är också ett utmärkt hjälpmedel, men endast ett hjälpmedel. För meningsfulla slutsatser krävs det något ytterligare: erfarenhet, omdöme och framför allt sunt förnuft. En numera vanlig form av oförnuft är inbillningen att man utan egentliga förkunskaper i statistik kan klara den statistiska bearbetningen bara man har tillgång till en dator och en statistisk "kokbok".

### Population och stickprov

Att studera en totalpopulation vore i många fall alltför dyrbart i tid, arbete och pengar. De flesta större undersökningar utförs därför på data samlade från en miniatyrupplaga av populationen, ett stickprov (sample). För att medge meningsfulla slutsatser måste stickprovet i princip dras på ett sådant sätt, att

*varje enskilda individ (= försöksperson, patient, djur, cylinderampull eller vad det nu är man studerar) i populationen har samma chans att komma med i stickprovet, och att individerna dras oberoende av varandra.*

Detta resulterar i ett oberoende slumpmässigt urval. Härigenom undviks inverkan av ovidkommande faktorer (bias). Endast slumpen bestämmer sam-

.....  
*Datorn är ingen statistisk "kokbok"*  
.....

mansättningen av urvalet. *Oberoende slumpmässigt urval* gör det möjligt, att dra slutsatser rörande populationen på basis av data samlade från stickprovet.

Nu ett exempel ur en allvarligt menad forskningsrapport i en "erkänd" odontologisk tidskrift: "Eight adult patients were chosen at random. They exhibited advanced destructive periodontitis ...". Ingenting finnes sagt om populationen – troligen patienterna vid det "Dental Center" varifrån arbetet emanerar. Och vad betyder "at random"? Patienterna på ett sådant ställe brukar inte bestå enbart av individer med avancerad parodontit utan av människor med friskt parodontium och sådana med parodontit av varierande grad. Om det vore sant att det verkligen rört sig om ett slumpmässigt urval, så skulle varje tänkande forskare dra öronen åt sig. Det hade i så fall rört sig om ett högst olyckligt urval som inte varit tjänligt för en seriös studie. I själva verket har det med all sannolikhet inte alls rört sig om ett slumpmässigt urval utan om ett urval av individer med just avancerad parodontit.

"Slumpmässigheten" har med all sannolikhet bestått i att man inte har följt någon viss plan vid urvalet. Den ofta anlitade metoden att sas gå ut i väntrummet och på måfå välja ut ett antal försökspersoner är raka motsatsen till vad man inom statistiken kallar slumpmässighet. Inom statistiken betyder slumpmässigt urval ett strikt reglerat förfarande: urval med hjälp av en slumpaltastabell, en slumpaltsgenerator i en dator eller något annat hjälpmedel som utesluter bias.

### Oberoende

Det är inom odontologien mycket vanligt att man åsidosätter regeln om oberoende, speciellt inom cariologisk och parodontologisk forskning genom att man i sina statistiska beräkningar använder antal *tänder* trots att stickprovsheten är *individer*. Det uppstår härigenom en orealistisk inflation av antalet enheter på vilka beräkningarna görs. Man kalkylerar mao med upp till 32 tänder gånger antalet indi-

Urvalsenheten måste bibehållas

vider i stället för att räkna med antalet individer. Ännu värre blir denna inflation om man räknar med tandytor. Om man tex har en sample på 20 individer med 32 möjliga tänder och i genomsnitt 5 ytor pr tand, blir  $n = 20 \times 32 \times 5 = 3\,200$  i stället för  $n = 20$ , som det borde vara. Eller när man som i talrika parodontologiska undersökningar har behandlat delar av parodontiet runt en och samma tand som oberoende enheter! Då kan det – som i en artikel – hända att 22 patienter blir till 1 155 "sites". Dessa "sites" är en ur statistisk synpunkt högst tvivelaktig uppfinning. Hundar och apor är dyrbara försöksdjur. Frestelsen ligger nära till hands att utnyttja sådana "sites" för att nedbringa antalet försöksdjur och ändå ha ett – som man tror – tillräckligt stort  $n$ . Så kan man nöja sig med 3 djur eller ibland tom endast ett enda.

### Bortfall

Vid de flesta stickprovsundersökningar är det omöjligt att nå varje individ i det avsedda stickprovet – det uppkommer mao ett bortfall. Orsaken kan vara: vägran att delta, död, sjukdom, hög ålder, byte av bostadsort etc. Det är nödvändigt att göra en analys av bortfallets storlek och orsaker. Vissa orsaker har rimligen intet samband med tandstatus eller vad det nu är man undersöker och påverkar därför inte resultatet medan andra kan ha ett starkt samband med den variabel eller de variabler som man vill undersöka. Man har tex skäl att misstänka, att personer med dåligt tandstatus och tandlösa individer eller äldre, som ofta har svårt att röra sig, är mindre benägna att delta i en tandhälsoundersökning.

#### Tumregel:

"Förkasta på förhand varje undersökning som inte klart och tydligt anger:

1. hur stickprovet dragits;
2. hur stort bortfallet är;
3. vilken populationen är.

Först om svar på dessa frågor finns, är en kritisk bedömning av undersökningen möjlig."

(Vejde 1962: Hur man räknar statistik.)

Statistisk slumpmässighet är strikt reglerat



## Parametrisk versus icke-parametrisk statistik

Man brukar dela in de statistiska metoderna i *parametriska* (tex aritmetiskt medium, varians och standardavvikelse, "Student's" t-test, korrelation och regression, multipel korrelation, variansanalys mfl, mfl) samt *icke-parametriska* resp *fördelningsfria* (tex median och percentiler, teckentest, rangsummatest enl Wilcoxon, U-test enl Mann-Whitney, rangkorrelation enligt Spearman, variansanalys enl Friedman,  $\chi^2$ , Fisher's exakta test o.a. Det råder inte sällan förvirring rörande dessa begrepp.

Parametrisk statistik bygger på parametrarna hos någon viss teoretisk frekvensfördelning tex aritmetiska mediet och standarddeviationen hos normalfördelningen. Fördelningsfri är den statistik som inte förutsätter någon viss fördelningsform. Termerna "icke-parametrisk" och "fördelningsfri" är inte helt synonyma. Det finns åtminstone ett vanligt statistiskt test som är såväl parametriskt som fördelningsfritt, nämligen teckentestet, som bygger på binomialfördelningen. Å andra sidan finns det icke-parametriska tests, som inte är helt fördelningsfria, nämligen testen enligt Wilcoxon och Mann-Whitney för differensen mellan medianer. Dessa kräver nämligen att fördelningar som skall analyseras har samma form (i praktiken approximativt samma form).

Ett parametriskt test förutsätter att variablerna är mätta med intervall- eller kvotskalor. Föreligger ordinal- eller nominalskalor måste icke-parametrisk statistik tillgripas.

(Ibland träffar man på vanföreställningen att icke-parametrisk statistik skulle kompensera för brist på slumpmässighet i urvalet – vilket naturligtvis är rent nonsens).

## Statistisk analys

Grunden för en god statistisk analys läggs redan vid planeringen av en undersökning. Alltför ofta går det så till, att man genomför en undersökning, samlar material och sedan börjar fundera på statistisk analys. Eventuellt presenterar man materialet för en statistiker och ber honom göra analysen. Man skall då ej sällan finna att undersökningen lagts upp eller materialet insamlats på ett sådant sätt att det inte går att lösa de uppställda problemen eller att man inte får ut så mycket som varit möjligt med ett

## Rådgör med statistiker

rationellare tillvägagångssätt. Bästa profylaxen mot en sådan utgång är givetvis att kontakta en kompetent statistiker redan i samband med planeringen.

Nästa punkt där statistikern kan vara till hjälp är vid valet av analysmetod(er). Det är inte alltid som de gängse standardmetoderna t-test, korrelation, regression och  $\chi^2$  är tillfyllest eller de lämpligaste. Slutligen kan en statistiker hjälpa till med att tolka resultaten av analysen.

## Metoder

Valet av metod(er) beror på

*Mätskala:* nominal-, ordinal-, intervallsskala eller kvotskala?

*Frekvensfördelning:* Normalfördelning eller ej?  
Symmetrisk eller ej?

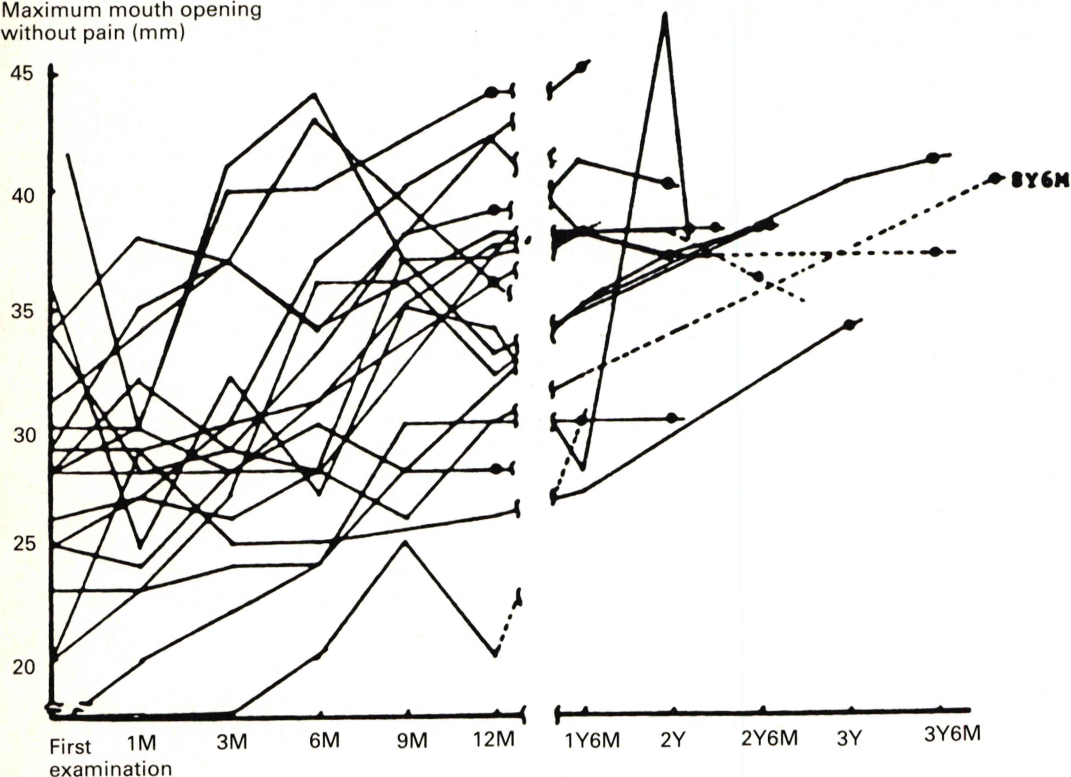
*Varianser:* likstora eller ej?

Varje statistiskt test bygger på givna teoretiska förutsättningar. Så tex förutsätter Student's t-test att värdena i de variabler som skall jämföras är normalfördelade; att varianserna är av samma storleksordning och att mätskalan är minst en intervallsskala. De material som man möter i odontologisk forskning avviker nästan alltid mer eller mindre, ej sällan excessivt från samtliga dessa förutsättningar. Man är allmänt överens om att t-testet är ett "robust" test: Måttliga avvikelser ger så små fel att de kan försummas. Man nöjer sig med att förutsättningarna är – som man säger – "approximativt uppfyllda". Men det finns naturligtvis gränser för vad man kan tillåta. Hur stora avvikelser som kan tolereras finns det dock inga regler för. Det blir en omdömessak som kräver stor erfarenhet.

Dessvärre är avvikelserna såväl från normalfördelningsformen som från variansernas homogenitet ofta så stora att de kan leda till avsevärda fel. Det är därför oundgängligen nödvändigt att man granskar förutsättningarna. Det gäller alla testmetoder, parametriska som icke-parametriska. Man skall då ibland finna, att det inte går att hitta någon metod vars förutsättningar uppfylls av de samples som man vill analysera.

Figur 1. Diagrammet illustrerar förändringar över tiden i maximal gapning hos fall med osteoarthritis.

Maximum mouth opening  
without pain (mm)



En gammal regel är att testmetod och signifikansnivå skall väljas *innan* någon analys utförs, detta för att resultaten inte skall påverka valet av den metod man slutligen presenterar, mao för att man inte skall frestas att välja den metod som ger ett (medvetet eller omedvetet) önskat resultat. Denna regel kan ifrågasättas. Det visar sig ibland vara nödvändigt, att anlita flera metoder för att ur de samlade resultaten försöka nå en rimlig slutsats.

### Deskriptiv statistik

Man träffar inte sällan på triviala fel som kan göra framställningen onjutbar för en läsare. Hit hör krångliga diagram, *figur 1*.

Ett nu omhuldat mode i våra tidskrifter bygger på föreställningen att diagram är lättare att läsa och mera informativa än tabeller. Sanningen är inte sällan den motsatta! En av fördelarna med tabeller är dessutom att de kan förmedla exakta data.

Värdet av tabeller saboteras dock inte så sällan. Ofast sker det i form av tabeller som meddelar %-tal utan att meddela de antal som utgör basen för beräkning av %-värdena, *tabell 1*. Siffrorna i sådana tabeller är meningslösa och kan dessutom ibland utnyttjas för att dölja materialets otillräcklighet. 88% låter mera imponerande än 7 fall av 8. Man bör

Tabell 1. Per cent Take off Grafts with Time

|                 | Days |    |     |    |    |    |
|-----------------|------|----|-----|----|----|----|
|                 | 5    | 8  | 11  | 14 | 30 | 60 |
| Porcine skin    | 82   | 42 | 0.2 | 0  | 0  | 0  |
| Autogenous skin | 96   | 85 | 75  | 73 | 70 | 70 |
| Palatal mucosa  | 93   | 80 | 70  | 69 | 65 | 65 |



när man laborerar med % ha i minnet att procent betyder: per hundra.

Inte sällan saknas också väsentliga upplysningar av annan natur i publicerade tabeller (tabell 2). Den enda upplysning som ges är att mätvärdena är i mm. I övrigt inte ett ord om siffrornas natur! Det första värdet i varje kolumn måste vara ett medelvärde, men vad betyder det andra? Är det standarddeviation eller medeltalets medelfel (standard error)? En tabell bör förses med all information som är nödvändig för att man skall kunna förstå den utan att behöva leta efter eventuella upplysningar i texten.

Tabell 2. Mean Change in Vertical Dimension (mm) of

| Graft   | Wounds with time |         |         |
|---------|------------------|---------|---------|
|         | 5                | 8       | 11      |
| Control | 10±0             | 8.7±0.7 | 7.3±0.7 |
| Porcine | 10±0             | 9.4±0.8 | 8.2±0.7 |
| Palate  | 10±0             | 9.7±0.7 | 8.5±0.7 |
| Skin    | 10±0             | 9.1±0.6 | 8.4±0.7 |

### Parade och oberoende variabler

Om man räknar antalet tänder på höger och vänster sida av munnen hos ett antal personer så får man en serie med variabelvärden som hänger samman två och två: *parade värden*\*. Vill man i stället jämföra antalet kvarvarande tänder hos individer tillhörande två olika åldersgrupper så får man två serier av *oberoende* variabelvärden, en för vardera gruppen. Den statistiska tekniken är olika för parade och oberoende samples, något som odontologiska författare inte alltid haft klart för sig.

### Teckentest

Rangordningsstatistik ("order statistics") är baserad på ordningsföljden (rangordningen) mellan observerade data och inte på dessas numeriska värden och kan ofta användas där villkoren för parametrisk

\* Intra-individuell jämförelse är inte alltid möjlig. Man kan då försöka använda individer som parvis liknar varandra i så hög grad som möjligt. Hur man kan gå till väga beskrivs utförligt i statistiska handböcker. Det finns många varierande möjligheter till sådan inter-individuell parbildning. Kriterierna måste dock hållas stränga.

statistik inte är uppfyllda, tex kravet på minst intervallskala. Teckentestet är ett sådant test för parade variabler. Det är enkelt att tillämpa på en rad olika problem – även på sådana där villkoren för parametriska tests är uppfyllda.

Ett antal begrepp beträffande vilkas innebörd åtskilliga författare tycks vara osäkra kan åskådliggöras med ett exempel på teckentest. Antag att vi vill undersöka effekten av parodontologisk behandling genom att fastställa huruvida det finns någon skillnad i röntgenologiskt mätt benhöjd hos en grupp på 25 personer efter ett års observationstid. När vi jämför benhöjderna hos individ efter individ finner vi att 18 personer har ökad, 4 minskad benhöjd och att 3 visar status quo. Hos dessa 3 har vi mao uppmätt exakt samma bennivå, dvs det föreligger vad som brukar kallas "bundna värden" ("ties"). När vi tar differenserna mellan observationerna vid kontrollbesöket och dem som gjorts före behandlingen får vi alltså 18 plus- och 4 minusdifferenser samt 5 noll-differenser. De senare bortser vi från eftersom variabelvärden vars differens är noll inte kan bidra till lösningen av problemet. Totalt 22 teckenförsedda differenser.

|    |   |    |
|----|---|----|
| +  | - | Σ  |
| 18 | 4 | 22 |

Det förefaller alltså som om man hade vunnit benhöjd genom behandlingen. Men den här fördelningen av differenser kan ju tänkas vara en av slumpens skördar? Den statistiska analysen går ut på att söka avgöra vilken sannolikheten är för att skillnaden *inte* beror på en slump. Låt oss försöka avgöra frågan.

### Noll-hypotes

Vi uppställer hypotesen att skillnaden mellan benhöjden vid kontrollmätningen och före behandlingen är noll.  $A - B = 0$ .

Om det inte föreläge någon skillnad före och efter behandlingen så hade vi kunnat förvänta oss lika många plus- som minustecken, dvs 12 plus- och 12 minustecken. Detta är de *förväntade* frekvenserna i motsats till de *observerade*.

Vi måste nu bestämma oss för den grad av sannolikhet – mao den *sannolikhetsnivå* – som vi kan nöja oss med. Vi väljer den i biologiska samman-



Resultatet av ett signifikanstest slår inte fast  
"sanningen"

hang konventionella 5%-nivån. Vi accepterar därmed en 5%-ig risk att vår slutsats blir felaktig. Det är en viktig princip – som det syndas mycket mot – att det inte är resultatet som skall bestämma signifikansnivån. Det skall forskaren göra – *innan* resultatet av den statistiska analysen föreligger.

Sannolikheten för att nollhypotesen är sann finns tabellerad i en tabell som bygger på binomialfördelningen. I det här fallet kunde hypotesen skrivas  $\Pr(A) = \Pr(B) = 1:2$ . (Pr står för probability, sannolikhet). I denna tabell finner vi  $P \leq 0.02$ . Sannolikheten för att nollhypotesen skall vara sann är alltså lägre än den valda 5%-nivån. Vi kan därför inte upprätthålla nollhypotesen utan måste förkasta den till förmån för den alternativa hypotesen  $A \neq B$ , dvs den genomsnittliga benhöjden i denna sample är sannolikt inte densamma *efter* som *före* behandling. Eftersom plustecknen överväger har benhöjden alltså med minst 95%:s sannolikhet ökat under observationstiden.

Detta upprepande av ordet "sannolikhet" har sina skäl. Det är viktigt att förstå att resultatet av ett signifikantest inte slår fast "sanningen". Ett statistiskt test kan inte bevisa att en hypotes är vare sig 100% sann eller 100% falsk. Det kan endast visa graden av sannolikhet för det ena eller andra. I det här exemplet fann vi 95% sannolikhet för att det föreligger en ökad benhöjd. Det bör kanske upprepas att vi löper en 5% risk att vår slutsats är falsk. Detta är ett av skälen till att man inte skall tro att man har löst något problem med en enda undersökning – i synnerhet en med så pass litet material. Odontologiska forskare verkar ibland att vara något lättsinniga på den här punkten. Kontrollundersökningar, helst flera, är av nöden om man vill komma *närmare* sanningen. Men vi måste som Fälldin vara nöjda även om vi inte når ända fram.

Sannolikheten för ett signifikant resultat ökar med ökande storlek på samplen. Det är emellertid viktigt att förstå att även en liten differens mellan medelvärden kan bli statistiskt signifikant om anta-

let fall bara är tillräckligt stort. En sådan kanske trestjärnigt signifikant differens kan mycket väl sakna varje spår av *klinisk* eller praktisk signifikans.

### Enkel- eller dubbelsidigt test?

Det enkelsidiga testet prövar säs endast en av fördelningens svansar, *figur 2* överst. Nollhypotesen  $A < B$ . I detta exempel har sannolikhet för att nollhypotesen skall vara sann antagits vara  $p = 0.025$ . Det dubbelsidiga testet jämför båda fördelningens svansar, *figur 2* nederst. Sannolikheten att 0-hypotesen är sann är i detta exempel  $p = 2.5\% + 2.5\% = 5\%$ .

Valet mellan enkel- och dubbelsidigt test skall göras *innan* någon analys utförs och skall baseras på tidigare kunskap och/eller rationella överväganden. Det gäller alltså att mot bakgrund av tidigare erfarenheter – egna eller andras – avgöra huruvida en eventuell differens mellan medelvärdena – det må vara aritmetiska medier eller medianer – rimligen kan förväntas gå endast i *en* riktning. I fall av tvekan är det dubbelsidiga testet att föredra.

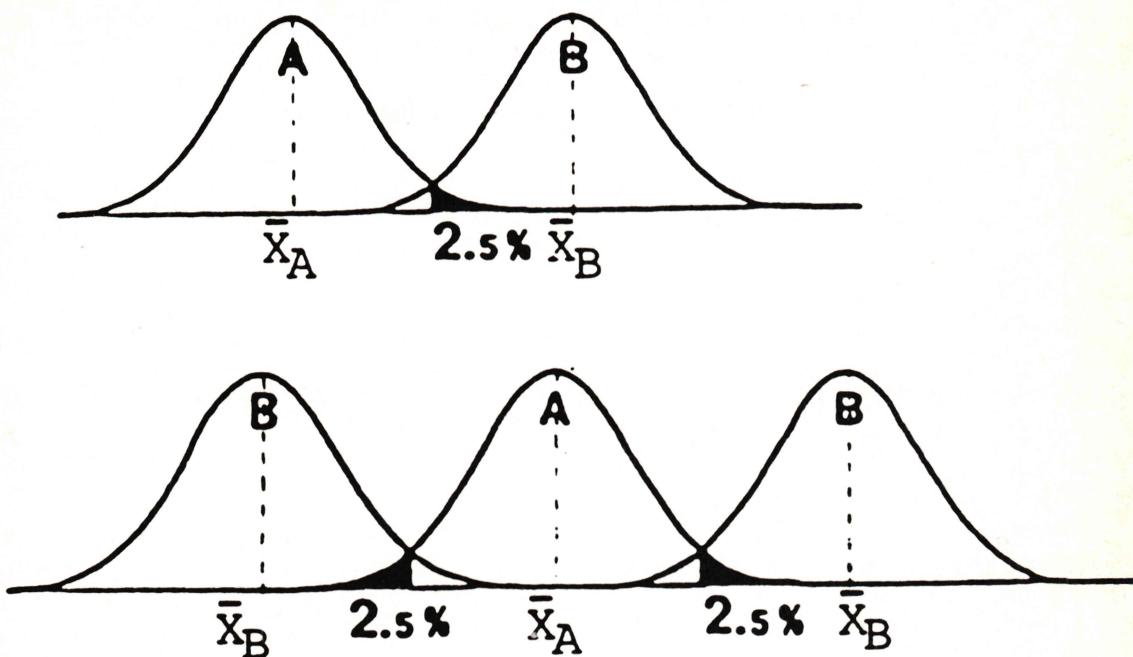
Varför är det nu så viktigt att gå igenom det här så i detalj? Skälet är att det enkelsidiga testet är känsligare, skarpare diskriminerande än det dubbelsidiga. Om man inte använder det enkelsidiga testet där det är indicerat löper man risken att missa en faktiskt föreliggande signifikans. Å andra sidan – om man använder det där det inte borde tillämpas riskerar man att acceptera en differens såsom signifikant även om den i själva verket inte är det.

I en vetenskaplig rapport bör det alltid anges huruvida enkel- eller dubbelsidigt test har utförts och dessutom skälen för det gjorda valet. Då kan läsaren själv bedöma och eventuellt vidta enligt en hans åsikt nödvändig korrektion av slutsatserna.

### Statistisk hantering av pH-värden

Som alla vet är pH den negativa 10-logaritmen av vätejonkoncentrationen, härnadan symboliserad med  $H^+$ :  $pH = -\log H^+$ ;  $H^+ = 10^{-pH}$ . Den logaritmiska skala i vilken pH mätes är en skala i avsaknad av fasta intervall mellan skalstrecken, en form av ordinalskala. Trots detta behandlas den och andra logaritmiska skalor i odontologisk litteratur i regel som vore det frågan om intervallskalor. pH-fördel-

Figur 2. Signifikanstest av skillnaden mellan två medelvärden,  $\bar{X}_A$  och  $\bar{X}_B$ .  
Enkelsidigt (överst) resp dubbelsidigt test (nederst).



ningar måste dock just på grund av skalans natur analyseras med icke-parametriska metoder.

Ett exempel med *oberoende* pH-variabler:

Sample A 6.7 5.1 5.4 7.6 6.9  
Sample B 4.7 3.1 3.4 4.3

Student's t-test kan inte användas på grund av ordinalskalan. Ett tänkbart – och ofta använt sätt – sätt att komma förbi den svårigheten vore att transformera pH till  $H^+$ . Då visar sig ett annat hinder nämligen en höggradigt signifikant skillnad mellan varianserna, något som framgår av ett F-test (variansanalys enl Fisher i sin enklaste form):

| Sample A          |                                     | Sample B              |                        |
|-------------------|-------------------------------------|-----------------------|------------------------|
| pH                | $H^+$                               | pH                    | $H^+$                  |
| 6.7               | $1.955 \times 10^{-7}$              | 4.7                   | $1.995 \times 10^{-5}$ |
| 5.1               | $7.943 \times 10^{-6}$              | 3.1                   | $7.043 \times 10^{-4}$ |
| etc               |                                     |                       |                        |
| Medelv            | $6.34 \quad 2.455 \times 10^{-6}$   | 3.88                  | $3.156 \times 10^{-4}$ |
| Varians           | $1.113 \quad 1.222 \times 10^{-11}$ | 0.563                 | $1.313 \times 10^{-7}$ |
| $F_{(pH)} = 1.98$ | N.S.                                | $F_{(H^+)} = 10747.0$ | ***                    |

Med t-tests baserade på pH-resp  $H^+$ -värden skulle man få helt olika resultat – inget av dem riktigt. t-

testet förutsätter bla lika stora varianser. Wilcoxon's rangsummatest har inte dessa restriktioner. Däremot fordras (approximativt) likformiga fördelningar. Rangsummatestet kan tillämpas på pH eller  $H^+$  med exakt samma resultat. Vem som helst kan med hjälp av en fick-kalkylator övertyga sig om att rangordningen mellan variabelvärdena blir densamma i båda fallen.

Gäller undersökningen *parade variabler* ställer sig saken helt annorlunda. Tag t ex differensen mellan pH-värdena 6.7 och 6.5 som är 0.2 och mellan 5.1 och 4.9 som också är 0.2. Alltså samma numeriska värde och följaktligen samma rangvärde. Transformera nu dessa pH-värden till  $H^+$  och beräkna differenserna.

| Resultat               |                        |                         |
|------------------------|------------------------|-------------------------|
| $H^+$                  | $H^+$                  | Differens               |
| $1.995 \times 10^{-7}$ | $3.162 \times 10^{-7}$ | $-1.167 \times 10^{-7}$ |
| $7.943 \times 10^{-6}$ | $1.259 \times 10^{-5}$ | $-4.646 \times 10^{-6}$ |

Alltså olikstora differenser och olika rangvärden! Dessa banala matematiska fakta tycks i stort sett



.....

*Analysera skillnader i pH-värden och bakterietal med icke-parametriska metoder*

.....

vara okända bland odontologiska forskare som sysslar med pH-värden.

Konsekvensen av det sagda blir att analys av parade pH-värden måste föregås av transformation till vätejonkoncentrationer.

Genom transformationen erhålles en intervallskala men frekvensfördelningarna blir i stället ofta excessivt onormala (sneda och toppiga). Analysen måste ske med icke-parametriska tests, förslagsvis U-test,  $\chi^2$ -test eller Kolmogorov-Smirnov's test.

### Bakterietal

Bestämning av bakterietal, tex lactobacill- eller mutantstal resulterar som regel i excessiva snedfördelningar, som utesluter användning av parametrisk statistik. På grund av rädsla för eller obekantskap med icke-parametrisk statistik har man i många arbeten gått in för logaritmisk transformation. Man har därvid oftast bortsett från eller glömt bort att den logaritmiska skalan omöjliggör parametrisk statistik och använt t-test, regressionsanalys o a parametriska metoder med de felslut som detta kan resultera i. Bakterietal bör liksom pH-värden analyseras med icke-parametriska metoder.

Som det förhoppningsvis har framgått härovan finns det många möjligheter till misstag vid statistisk bearbetning av odontologiska (och övriga medicinska) material. Många är av den art att de med litet eftertanke kan undvikas, andra är mera intrikata. Statistisk analys förutsätter kunskaper och ständig övning. Lika litet som en tandläkare utan specialutbildning kan ta fram en narkosapparat en eller två gånger om året och ge en effektiv och riskfri lustgasanalogesi, lika litet kan han/hon någon gång om året ta fram en dator och göra en korrekt statistisk analys.

Ett ord, varningens ord, eller två är säkerligen också på sin plats till läsaren av odontologiska (och medicinska) tidskrifter.

"Men tro ej allt vad skeppare förtälja  
om underbara ting som de erfarit  
och om de vises sten som de ha funnit."

(Tegner: Tal vid magisterpromotionen  
i Lund 1821).

"Read it at least three times;  
first to see what it is all about;  
secondly, to see what it says;  
and thirdly, in an attitude of friendly hostility."  
(Edward H. Angle 1907)

### REFERENSER

- Alanen, Pentti*: Parametric versus nonparametric tests in caries studies based on DMF index. Proc. Finn. Dent. Soc. 81:6:1985.
- Bradley, James V*: Distribution-free statistical tests. Prentice-Hall 1968. Kapitel 2: Distribution-free versus classical tests.
- Hooke, Robert*: Vetenskapliga försök. Prisma, Sthlm 1966 (originalets titel: Introduction to scientific inference).
- Reichmann, W. J.*: Rätt och fel i statistik. W & W. Sthlm 1969 (originalets titel: Use and Abuse of Statistics).
- Siegel, Sidney*: Nonparametric statistics for the behavioral sciences. McGraw-Hill, New York, Toronto, London. Kogakusha Company, Tokyo 1956.
- Smith, Martyn H.*: Parametric vs. nonparametric. J. of Periodontal Research 17: 514, 1982.
- Zeisel, Hans*: Tolkning av siffror. W & W. Sthlm 1969 (originalets titel: Say it with figures).