

Arbeidsnotater

S T A T I S T I S K S E N T R A L B Y R Å

Dronningensgt. 16, Dep, Oslo 1. Tlf.* (02) 41 38 20

10 77/42

15. november 1977

SAMANLIKNING AV POPULASJONAR BASERT PÅ IKKJE-PARAMETRISKE METODAR

Av
Rolf Aaberge^{x)}

INNHALD

	Side
1. Innleiing	1
2. Samanlikning av to populasjonar når modellen er lineær	3
2a. Samanlikning av Student t-test og Wilcoxon testen når observasjonane er genererte frå gammafordelingen	5
2b. Wilcoxon testen nytta på data frå Ferieundersøkinga 1974	7
3. Kruskal-Wallis testen, ein ikkje-parametrisk metode for samanlikning av fleire populasjonar	8
4. Eksempel på bruk av grafiske metodar for å vurdere modellar	9
5. Samanlikning av to populasjonar når modellen er ikkje-lineær	12
5a. Estimering i generelle modellar	12
5b. Konfidensband for $\Delta(x)$	13
5c. Eksempel på samanlikning av reiseaktiviteten mellom to aldersgrupper når modellen er ikkje-lineær	14
Appendiks 1. Testing av modellar	17
Appendiks 2. Det normale sannsynplottet	19

^{x)} Halvor Strømme har utført programmeringsarbeidet. John Dagsvik, Tor Haldorsen og Petter Laake har gitt nyttige kommentarar til manuskriptet.

SAMANDRAG

Føremålet med dette notatet er å visa verdien av å nytta ikkje-parametriske modellar ved samanlikning av populasjonar. Desse modellane krev mindre føresetnader om fordelinga til observasjonane enn dei parametriske modellane. Bruk av parametriske modellar medfører føresetnader om at fordelinga til observasjonane tilhøyrer ein kjend fordelingsklasse. Det einaste ukjende er eit sett parametarar som ofte uttrykker forventning og varians.

Dersom ein parametrisk modell ikkje kan avleiast på grunnlag av den fysiske karakteren til eksperimentet eller på grunnlag av tidlegare erfaring, vil det vere naturleg å nytta ein ikkje-parametrisk modell for å skildra materialet. I notatet har vi handsama metodar som er utvikla for ikkje-parametriske modellar.

Utgangspunktet for vår interesse i ikkje-parametriske metodar var ønsket om å visa nytten av desse metodane i ein konkret analyse, nemleg Ferieundersøkinga 74. Her er ein interessert i å påvisa eventuell samanheng mellom alder og talet på ferieturdagar i løpet av eit år. Ved å nytta moderne plottingsteknikkat har vi vist at våre data ikkje er normalfordelte, slik at føresetnadene for å nytta Student t-test ikkje er til stades. I den aktuelle situasjonen har vi vist at bruk av Student t-test i staden for Wilcoxon testen vil medføra eit tap av informasjon som tilsvarar ca. 40% færre observasjonar.

1. Innleiing

Problemet med å avgjera om to tilfeldige utval (observasjonssett) er genererte frå same fordelingsfunksjon, kallar vi to-utval problemet. Eit slikt problem har vi for eksempel når vi skal avgjera om eit medikament har effekt på ein viss sjukdom. Frå den aktuelle populasjonen, som altså er samansett av dei personane som lir av denne sjukdomen, ønskjer vi å nytta N personar i eksperimentet vårt. Vi let dei n første personane som vert trekte ut bli behandla med medikamentet. Dei resterande $m=N-n$ som er kontrollar, vert behandla med eit standardpreparat som skal vere utan verknad av noko slag. Effekten blir målt på ein eller annan måte. Vi får da to sett observerte effektar, eitt for behandlingsgruppa Y_1, \dots, Y_n og eitt for kontrollgruppa X_1, \dots, X_m . Vi føreset at X -ene og Y -ene er uavhengige og identisk fordelte med kumulative fordelingsfunksjonar F og G . Vi skal testa om $F(x)=G(x)$ for alle x mot alternativet at $F(x) > G(x)$ for alle x , dvs. vi ser på alternativet at Y -ene har ein tendens til å ha større observasjonsverdiar enn X -ene. Konkret tyder dette at vi tester om "medikamentet har ingen effekt" mot alternativet at "medikamentet har effekt".

For å testa problem av den typen som er skissert ovanfor, er Student t-test den testen som er mest nytta. Dersom vi forutset at begge utvala er normalfordelte med same varians, men ulik forventning, veit vi at Student t-test er den over alt sterkaste testen for to-utval problemet.

Ved sida av å nytta strenge føresetnader, er Student t-test følsom over avvik i føresetnadene. I mange situasjoner vil fordelingane til dei to utvala vere ikkje-normale, og då vil altså Student t-test ofte ha dårlege eigenskaper. Spesielt har vi at nivået på testen ikkje kan kontrollerast. Wilcoxon (1945) føreslo ein alternativ metode (Wilcoxon test) som berre krev at observasjonane er uavhengige og gruppevis identisk fordelte, medan formene på fordelingane for observasjonane i dei to utvala er ukjende. Wilcoxon test er basert på ein observator som er definert ved summen av rangane i behandlingsgruppa når alle observasjonane er rangert saman. Rangane, som framkjem ved å ordna dei to utvala i stigande rekkefølge, er nummera på desse ordna tala frå venstre mot høgre frå 1 til N . Dersom for eksempel observasjonen X_k har den minste verdien blant observasjonane i dei to utvala, så vert altså rangen til X_k lik 1. Ein gunstig eigenskap ved Wilcoxon test, er at nivået er det same uansett hva slags fordeling observasjonane kjem i frå. Vi seier at testen er fordelingsfri. Dersom vi i ein konkret analyse eksempelvis har gått ut i frå at observasjonane stammer frå to normalfordelingar, medan dei i røynda stammer frå to eksponentielle fordelingar, så vil ikkje dette skapa noko forstyrning av nivået ved bruk av Wilcoxon test. Denne gunstige eigenskapen har altså ikkje Student t-test. Sjølv om vi veit at for alle fordelingar vil Wilcoxon test gi kontroll over nivået til testen, er det viktig å avgjera om vi taper mykje i styrke på å nytta denne testen i stadenfor Student t-test, dersom observasjonane verkeleg er normalfordelte. Hodges & Lehmann (1956) har vist at tapet i styrke ikkje er stort.

Dei har vidare samanlikna dei to testane under forskyvningsmodellen ($F(x)=G(x+\theta)$) for alle x) for ulike F , og funne at mogleg vinst ved å nytta Wilcoxon test er uavgrensa. Vi meiner derfor at Wilcoxon test har så gode eigenskaper at metoden fortener ei større merksemd enn den har i dag. Dette heng også saman med at mange av dei fordelingane som ein ønskjer å undersøka nettopp ikkje er normale, og at t -testen, som likevel vert nytta som metode for samanlikning, da ikkje har gode eigenskaper.

Dersom det ikkje fins ein konstant θ slik at $F(x)=G(x+\theta)$ for alle x , dvs. at den lineære modellen (forskyvningsmodellen) ikkje er tilfredsstilt, så vil det eksistera ein funksjon $\Delta(x)$ slik at $F(x)=G(x+\Delta(x))$ for alle x . $\Delta(x)$ kan nyttast som eit mål for differensen mellom dei to populasjonane. Doksum (1974) har vist at det eksisterer berre ein funksjon $\Delta(x)$ som tilfredsstiller $F(x)=G(x+\Delta(x))$ for alle x og på same tid er eit nyttig mål for differensen mellom populasjonane. Denne funksjonen er definert ved $\Delta(x)=G^{-1}(F(x))-x$. Under den lineære modellen ($F(x)=G(x+\theta)$) for alle x) som ligg til grunn for bruk av Student t -test, blir forskyvningen mellom populasjonane estimert ved forskjellen mellom gjennomsnitta i dei to utvala. I situasjonar der den lineære modellen ikkje gjeld, vil dette ofte vere eit dårleg mål. Da vil $\Delta(x)$ vere bedre sidan $\Delta(x)$ fangar opp alle moglege forskyvningar mellom dei to fordelingsfunksjonane og altså ikkje berre lokasjonforskyvning.

I kapittel 2 er det gitt ei matematisk-statistisk presisering av to-utval problemet. Vi har vidare drøfta eigenskapane til Wilcoxon test og Student t -test og gjort samanlikning av dei når observasjonane i dei to utvala stammer frå ulike gammafordelingar.

I kapittel 3 har vi gjort greie for Kruskal-Wallis testen. Denne testen er ein generalisering av Wilcoxon-testen og kan nyttast til å samanlikna fleire utval for å avgjera om det er nokon forskjell på dei (med omsyn på ein særskild faktor).

I kapittel 5 har vi handsama den ikkje-lineære modellen $F(x)=G(x+\Delta(x))$ for alle x . Vi har gitt ein estimator $\hat{\Delta}(x)$ for $\Delta(x)$. Vidare viser vi korleis ein på grunnlag av eit observasjonsmateriale kan fastsetja eit band $[\underline{\Delta}(x), \bar{\Delta}(x)]$, som vi kan påstå inneheld $\Delta(x)$ med ein viss sikkerheit. Eit slikt band vert kalla konfidensband. Saman med estimatoren $\hat{\Delta}(x)$ gir konfidensbandet $[\underline{\Delta}(x), \bar{\Delta}(x)]$ informasjon om skiftfunksjonen $\Delta(x)$, som altså er eit mål for differensen mellom to populasjoner. Legg merke til at under den lineære modellen er $\Delta(x)=\theta=\xi_2-\xi_1$ for alle x , som er forskjellen mellom forventningane i dei to utvala. Konfidensbandet $[\underline{\Delta}(x), \bar{\Delta}(x)]$ kan nyttast til å testa dette. Vi har nemleg at dersom bandet $[\underline{\Delta}(x), \bar{\Delta}(x)]$ ikkje kan innehalda ei horisontal linje, så forkaster vi hypotesen H om ein lineær modell $F(x)=G(x+\theta)$ for alle x . Konfidensbandet kan også nyttast til å dra slutningar om skiftfunksjonen.

Vi kan finna ut for kva slags x $\Delta(x)$ er større enn 0 og for kva slags x $\Delta(x)$ er mindre enn 0 ved å sjå på respektive $\{x: \underline{\Delta}(x)>0\}$ og $\{x: \bar{\Delta}(x)<0\}$. Vi er altså interessert i dei x som er slik at den nedre delen av konfidensbandet er større enn 0 og den øvre delen mindre enn 0.

Ved å nytta materialet frå Ferieundersøkinga 1974, har vi brukt metodane ovanfor til å samanlikna lengden på ferieturar (talet på ferieturdagar) mellom dei to aldersgruppene 15-34 år og 35-74 år. I kapittel 3 har vi nytta Wilcoxon-testen for å testa om det generelt er signifikant forskjell i lengden på ferieturar mellom dei to aldersgruppene. Resultatet viser at materialet ikkje gir grunnlag for å påstå at det er signifikant forskjell. I kapittel 4 har vi vidare vist korleis ein kan testa hypotesen om at observasjonane er genererte frå normalfordelingen. Testar for fordelingsfunksjonar er utførleg omtala i appendiks 1 og appendiks 2.

I kapittel 5c, har vi samanlikna lengden på ferieturar mellom dei to aldersgruppene når modellen er ikkje-lineær, altså når $F(x)=G(x+\Delta(x))$ for alle x , der $\Delta(x)=G^{-1}(F(x))-x$ og F og G er dei to ukjende fordelingsfunksjonane til respektive aldersgruppene 35-74 år og 15-34 år. $\Delta(x)$ vert då eit mål for forskjellen i lengden på ferieturar mellom dei to aldersgruppene. Ved analyse av materialet på grunnlag av denne modellen, kan vi for eksempel vere i stand til å finna ut om det er variasjon i forskjellen i talet på ferieturdagar mellom dei to aldersgruppene, medan modellen som ligg til grunn for bruk av Student t -test berre gjer oss i stand til å avgjera om det generelt er ein forskjell. Det framkjem av plottet $(x, \hat{\Delta}(x))$ at dei minste og mellomstore tala på ferieturdagar i den yngste aldersgruppa er større enn tilsvarande for den eldste gruppa, medan dei lengste ferieturane i den yngste gruppa er mindre enn tilsvarande for den eldste gruppa.

2. Samanlikning av to populasjonar når modellen er lineær

I kapittel VII i Sverdrup (1964) er problemstillingen, som er skissert i innleiinga, presisert på følgjande matematisk-statistisk form:

La X_1, \dots, X_m og Y_1, \dots, Y_n vere to uavhengige utval av statistiske variable der

$$EX_i = \xi_1, \text{ var } X_i = \sigma^2, i=1, \dots, m$$

og

$$EY_j = \xi_2, \text{ var } Y_j = \sigma^2, j=1, \dots, n.$$

Ein ønskjer å testa hypotesen $H_0: \xi_1 = \xi_2$ mot alternativet $A: \xi_2 > \xi_1$. Under H_0 er altså begge utvala generert frå same fordelingsfunksjon.

Modellen ovanfor vert ofte omtala som forskyvningsmodellen, og den vert vanlegvis presentert på følgjande form:

La X_1, \dots, X_m vere uavhengige og $X_i \sim F$ $i=1, \dots, m$. La dessuten Y_1, \dots, Y_n vere uavhengige og $Y_j \sim G$, $j=1, \dots, n$ der F og G er kontinuerlige kumulative fordelingsfunksjoner med $EX_i = \xi_1$, $EY_j = \xi_2$. Vidare føreset vi

$$(2.1) \quad Y = X + \theta \quad (\text{= tyder likskap i fordeling) dvs. } G(y) = F(y - \theta), \text{ der } \theta = \xi_2 - \xi_1.$$

Det følgjer nå lett at vi kan skriva H_0 og A (definert ovanfor) på følgjande form, $H_0: \theta = 0$ mot $A: \theta > 0$. Vi ser at kravet om $\text{var } X_i = \text{var } Y_j$, $i=1, \dots, m$, $j=1, \dots, n$ følgjer frå (2.1).

La oss no tenkja oss at vi har eit observasjonsmateriale som tilfredsstillar forskyvningsmodellen med $F(x) = \Phi\left(\frac{x - \xi_1}{\sigma}\right)$. Vi har altså to utval som er genererte frå normalfordelingar med same varians, men ulike forventningar. For å testa om forskjellen i forventning mellom dei to utvala er signifikant, er det vist at Student test for to utval er den overalt sterkaste testen i denne situasjon. Student test forhaster $H_0: \theta = 0$ mot $A: \theta > 0$ ($\theta = \xi_2 - \xi_1$) dersom

$$T = \frac{\bar{Y} - \bar{X}}{S} \sqrt{\frac{mn}{m+n}} > k$$

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i, \bar{Y} = \frac{1}{n} \sum_{j=1}^n Y_j \text{ og } S^2 = \frac{1}{m+n-2} \left[\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2 \right].$$

Legg merke til at under H_0 er T Studentfordelt med $m+n-2$ fridomsgrader når $F(x) = \Phi\left(\frac{x - \xi_1}{\sigma}\right)$.

Vi skal nå handsama to-utvalsproblemet ved hjelp av ein annan metode, foreslått av Wilcoxon (1945). Denne metoden rangerer dei to utvala saman og nyttar vidare summen av rangane til Y -ene som testobservatorer. Dersom vi let R'_1, \dots, R'_n vere dei ordna rangene til Y -ene blant alle observasjonene, så blir altså Wilcoxon-observatoren definert ved $W' = \sum_{i=1}^n R'_i$. Under nullhypotesen at dei to utvala er generert frå samme fordelingsfunksjon F , er fordelingen til rangene og dermed til W' uavhengige av F . Vi seier at Wilcoxon-testen er ein fordelingsfri test. Dersom vi ser på alternativet at utvalet av Y -ene har ein tendens til å ha større observasjonar og dermed større ranger enn X -ene, så forkaster vi nullhypotesen i favør av dette alternativet, når W' er større eller lik ein konstant, k . Ekvivalent kan vi formulere dette på følgjande måte: Vi forkaster nullhypotesen $H: F=G$ til fordel for alternativet $A: F > G$ dersom $W' > k$. F og G er definert som tidlegare.

Legg merke til at forskyvningsmodellen med $H_0: \theta = 0$ mot $A: \theta > 0$ er eit spesialtilfelle av det problemet som er omtala ovanfor.

Dersom vi ønskjer å testa $H: F=G$ mot $A': F \neq G$ seier vi at vi har ein to-sidig testsituasjon. Då forkaster vi H_0 dersom $W' < c_1$ eller $W' > c_2$ der c_1 og c_2 er konstanter. Nedanfor skal vi ved hjelp av eit eksempel visa korleis ein kan gå fram for å nytta Wilcoxon test når m og n er småe, dvs. når vi har få observasjonar i dei to utvala.

Eksempel 1: Vi skal testa effekten av eit middel. Forsøket omfattar $N=10$ objekt, der $n=5$ objekt er tilfeldig trekte og behandla med middelet. Dei andre $m=5$ er kontrollobjekt. Vi måler effekten og har at "store tal" viser at middelet virkar. La oss tenkja at undersøkinga gir dette resultatet:

Behandlingsgruppa $Y=(10, 13, 8, 14, 16)$

Kontrollgruppa $X=(4, 11, 5, 6, 3)$

Vi får då $R'=(R'_1, \dots, R'_5)=(5, 6, 8, 9, 10)$, som altså er dei ordna rangane til Y -ene. Vidare vert $W'=\sum_{i=1}^5 R'_i=38$. Hypotesen som vi skal testa er H_0 : "Middel virkar ikkje" mot A : "Middel virkar": Under H_0 har alle moglege rangvektorar like stort sannsyn for å inntreffa. Sidan vi har $\binom{N}{n}$ moglege rangvektorar, der rangane er ordna, vert $P_{H_0}(R'=r')=\frac{1}{\binom{N}{n}}=\frac{1}{\binom{10}{5}}=\frac{1}{252}$. Vi kan da finna signifikanssannsynet,

$P_{H_0}(W' \geq 38)$, på følgjande måte: Vi tel opp kor mange av dei moglege rangvektorane som gir $W' \geq 38$, og finn 4 stykker. Sidan alle dei 252 rangvektorane har like stort sannsyn, så vert $P_{H_0}(W' \geq 38)=4 \cdot \frac{1}{252} \approx 0.016$. Vi får altså forkastning for alle nivå større enn 0.016.

Hvis vi i observasjonsmaterialet har observasjonar med like verdjar, løyser vi dette på følgjande måte:

La det vere e ulike verdjar blant observasjonane i dei to utvala ($e \leq N=m+n$). La t_1 =talet på observasjonar lik den minste verdien, og la vidare t_i =talet på observasjonar lik den i -te minste verdien, $i=2, \dots, e$.

La ρ_i vere midrangen til observasjonane lik den i -te minste verdien definert ved

$$\rho_i = \sum_{j=1}^{i-1} t_j + \frac{t_i+1}{2}, \quad i=2, \dots, e \text{ og } \rho_1 \text{ midrangen til observasjonane lik den minste verdien definert ved}$$

$$\rho_1 = \frac{t_1+1}{2}.$$

La R_1, \dots, R_n vere dei ordna midrangane til Y -ene. Da blir Wilcoxon-observatoren definert ved $W = \sum_{i=1}^n R_i$. Vi ser lett at W' er W avgrensa til eit observasjonsmateriale der det ikkje fins like verdjar. Vi skal nå ved hjelp av eit eksempel visa korleis ein kan gå fram for å nytta Wilcoxon-test når vi i materialet vårt har observasjonar med like verdjar.

Eksempel 2: Vi går ut i frå at vi har den situasjonen som er handsama i eksempel 1, men at undersøkinga no har gitt dette resultatet:

Behandlingsgruppa $Y=(10, 13, 10, 14, 13)$

Kontrollgruppa $X=(6, 8, 6, 13, 10)$

Altså blir $t_1=2, t_2=1, t_3=3, t_4=2, t_5=1$ ($e=5$) og vi får midrangane $(\rho_1, \dots, \rho_{10}) = (1.5, 1.5, 3, 5, 5, 5, 8, 8, 8, 10)$. Dei ordna midrangane til Y -ene blir da $R=(R_1, \dots, R_5) = (5, 5, 8, 8, 10)$. Dermed har vi $W = \sum_{i=1}^5 R_i = 36$. For å nytta W til testing av hypotesen (H_0 i eksempel 1), må vi finna fordelingen til W under H_0 . Det vi gjer er å rekna ut den betinga fordelingen gitt midrangane. Vi går da fram på følgjande måte: La oss tenkje oss at vi har ein urne med 10 kuler. På disse kulene står midrangane 1.5, 1.5, 3, 5, 5, 5, 8, 8, 8 og 10. Vi trekker så 5 kuler tilfeldig og let desse representera behandlingsgruppa. På denne måten får vi den betinga fordelingen til (R_1, \dots, R_5) og dermed den betinga fordeling til W . Ved hjelp av følgjande tabell skal vi visa korleis ein kan finna signifikanssannsynet, $P_{H_0}(W \geq 36 \mid (\rho_1, \dots, \rho_{10}))$.

(R_1, \dots, R_5)	(5, 8, 8, 8, 10)	(3, 8, 8, 8, 10)	(5, 5, 8, 8, 10)
W	39	37	36
$P_{H_0}((R_1, \dots, R_5) (\rho_1, \dots, \rho_{10}))$	$\frac{\binom{3}{1}\binom{3}{3}\binom{1}{1}}{\binom{10}{5}} = \frac{3}{252}$	$\frac{\binom{1}{1}\binom{3}{3}\binom{1}{1}}{\binom{10}{5}} = \frac{1}{252}$	$\frac{\binom{3}{2}\binom{3}{2}\binom{1}{1}}{\binom{10}{5}} = \frac{9}{252}$

Dette gir $P_{H_0}(W \geq 36 | (\rho_1, \dots, \rho_{10})) = \frac{13}{252} \approx 0.052$. Vi får altså forkastning av nullhypotesen for alle nivå større enn 0.052.

Tabeller over den betingte fordelingen til W fins av naturlege grunnar ikkje. For små utvalstorleiker (få observasjonar i utvala) må denne finnast ved hjelp av den metoden som er skissert ovenfor. For W' , dvs. for W med eit observasjonsmateriale der det ikkje fins like verdier er fordelingen tabellert i Owen (1962). Når utvala er store kan vi nytta følgjande resultat, sjå Lehmann (1974):

Hvis $\lim_{N \rightarrow \infty} \max \frac{t_i}{N} < 1$, som tyder at ein ikkje kan ha for mange like observasjonar, så er

$$\lim_{\min(m, n) \rightarrow \infty} P_{H_0} \left(\frac{W - EW}{\sqrt{\text{var } W}} \leq z \right) = \Phi(z)$$

der

$$\text{var } W = \frac{n}{12} (N+1) - \frac{m n \sum_{i=1}^e (t_i^3 - t_i)}{12 N (N-1)} \quad \text{og } EW = \frac{n}{2} (N+1)$$

Lehmann (1974) viser at tilnærminga er noko betre dersom vi nyttar kontinuitetskorreksjon, dvs. at vi set $P_{H_0}(W \leq k) \approx \Phi\left(\frac{k + \frac{1}{2} - EW}{\sqrt{\text{var } W}}\right)$ i staden for $P_{H_0}(W \leq k) \approx \Phi\left(\frac{k - EW}{\sqrt{\text{var } W}}\right)$. Numeriske utrekningar viser at tilnærminga er god når $m, n \geq 8$.

Når $m, n \geq 8$ forkaster vi derfor $H_0: F=G$ til fordel for $A: F > G$ dersom $W > k$ der $k = g_{1-\epsilon} \sqrt{\text{var } W} - \frac{1}{2} + EW$ og $g_{1-\epsilon}$ er $1-\epsilon$ fraktilen frå Φ . Når $n, m < 8$ kan vi nytta den framgangsmåten som blei handsama i eksempel 2.

2a. Samanlikning av Student-test og Wilcoxon testen når observasjonane er generert frå gammafordelingen

Vi skal nå føreta ein vurdering av dei to testane, Student (t) og Wilcoxon (W) når vi tester $H_0: \theta=0$ mot $A: \theta>0$ under modellen $G(Y)=F(Y-\theta)$ der F er kontinuerleg kumulativ fordelingsfunksjon. Spesielt skal vi ta for oss tilfellet der $F=F_1$ er gammafordelingen med parametrane $\alpha_1=2.38$ og $\beta_1=9.3$.

Viktige spørsmål å få avklara er om Wilcoxon-testen er:

- (i) effesient når observasjonsmateriale er generert frå normalfordelingen, dvs. at vi ikkje taper for mykje om vi nyttar ein kjend normalitet,
- (ii) meir effesient enn Student t -test for andre rimelege fordelingsklassar.

For å svara på desse spørsmåla innfører vi følgjande definisjon:

Effesiansen av W relativt til t er forholdet $\frac{N_t}{N_W}$ av utvalstorleikane som Student og Wilcoxon testane treng for å oppnå same styrke γ for eit gitt alternativ når begge testane har same nivå ϵ .

$\frac{N_t}{N_W}$ er avhengig av ϵ, γ og av alternativet. Hodges & Lehmann (1956) viste at $0.916 \leq \frac{N_t}{N_W} \leq 0.978$ når observasjonsmateriale er normalfordelt. Altså er (i) stetta.

Sidan t -testen ikkje er fordelingsfri, dvs. nivået er avhengig av kva slags fordeling observasjonane stammer frå, er det vanskeleg å gi svar på (ii) med utgangspunkt i definisjonen av effesians. Da t -testen er asymptotisk fordelingsfri, innfører vi derfor begrepet asymptotisk effesians som er definert på følgjande måte:

La $0 < \epsilon < \gamma < 1$ vere gitte og la ϕ og ϕ^* vere ϵ -nivå testar og $\gamma(\rho, \theta)$ styrkefunksjonen for testen ρ . La dessutan $\{\theta_N\}$ ($N=m+n$) vere ein sekvens slik at $\gamma(\phi_N, \theta_N) \rightarrow \gamma$ når $N \rightarrow \infty$ og $\{r_N\}$ ein sekvens slik at $\gamma(\phi_{r_N}^*, \theta_N) \rightarrow \gamma$ når $N \rightarrow \infty$. La $N \rightarrow \infty$ på ein slik måte at $\frac{m}{N} \rightarrow \lambda$ der $0 < \lambda < 1$, dvs. vi får ein viss brøkdel av observasjonar frå kvart utval og ikkje nokon av utvala får dominera. Anta at grensa for $\frac{r_N}{N}$ eksisterar og er uavhengig av $\{\theta_N\}$. Den asymptotiske effesiensen (Pitman-effesiensen) til ϕ^* med omsyn på ϕ er da gitt ved

$$e(\phi^*, \phi) = \lim_{N \rightarrow \infty} \frac{N}{r_N}$$

Pitman (1948) viste at når ϕ^* er Wilcoxon testen og ϕ Student t-test så vert

$$(2.2) \quad e_{W,t} = e(\phi^*, \phi) = 12 \sigma^2 \left[\int f^2(x) dx \right]^2 \text{ der } f \text{ er tettheten til den kumulative fordelingen } F \text{ og } \sigma^2 = \text{Var } X = \text{Var } Y.$$

Når F er normalfordelingen fann Pitman (1948) at $e_{W,t} = \frac{3}{\pi} = 0.955$. Hodges & Lehmann (1956) viste at den asymptotiske effesiensen har ei nedre grense på 0.864, medan det ikkje eksisterar noko øvre grense. Dette tyder at både (i) og (ii) er tilfredsstillt. La

$$(2.3) \quad f(x) = \frac{1}{\beta_1^{\alpha_1} \Gamma(\alpha_1)} x^{\alpha_1 - 1} \frac{1}{\beta_1} e^{-\frac{x}{\beta_1}}, \quad x > 0 \text{ (gammafordelingen).}$$

Vi får da at

$$\sigma^2 = \text{var } X = \alpha_1 \beta_1^2, \quad EX = \alpha_1 \beta_1$$

og

$$\int_0^{\infty} f^2(x) dx = \frac{\Gamma(2\alpha_1 - 1)}{\beta_1^2 \Gamma^2(\alpha_1) 2^{2\alpha_1 - 1}}$$

og dermed får vi frå (2.2)

$$e_{W,t} = 12\alpha_1 \left[\frac{\Gamma(2\alpha_1 - 1)}{\Gamma^2(\alpha_1) 2^{2\alpha_1 - 1}} \right]^2.$$

Legg merke til at den asymptotiske effesiensen er uavhengig av β_1 . I tabell 1 er $e_{W,t}$ gitt for nokre verdiar av α_1 .

Tabell 1. Asymptotisk effisiens for Wilcoxon testen med omsyn på Student t-test når observasjonane er gammafordelte

α_1	1.00	1.50	2.00	2.38	2.50	3.00	4.00
$e_{W,t}$	3.00	1.83	1.50	1.42	1.35	1.26	1.17

Vi ser frå tabell 1 at tapet ved å nytta Student t-test i staden for Wilcoxon testen kan bli ganske stort når observasjonane er gammafordelte.

På grunnlag av dei eigenskapene ved Wilcoxon-testen som er nemnde ovanfor kan vi trekka følgjande konklusjon: I ein situasjon der den underliggjande fordelingen er ukjend og utvalsstorleiken er stor, vil bruk av Wilcoxon-observatoren aldri medføra eit stort tap i effesiens (i forhold til Student t-test). På den andre sida er mogleg vinst i effesiens uavgrensa.

2b. Wilcoxon testen nytta på data frå Ferieundersøkinga 1974

Frå Ferieundersøkinga 1974 har vi to aldersutval 15-34 år og 35-74 år på respektive $n=659$ og $m=1008$ personer. Kvar person er vorten spurd om talet på ferieturdagar førre året, der det går inn i definisjonen av ferietur at ein må ha overnatta borte minst 4 netter i samanheng. Vi er interessert i å finna ut om talet på ferieturdagar blant personar som reiser på ferietur i aldersgruppa 15-34 år er signifikant større enn i gruppa 35-74 år. For å løysa dette problemet kan vi nytta Wilcoxon test. Denne testen er handsama i kapittel 2.

Eit program som utfører reknearbeidet i samband med Wilcoxon test er innlagt på terminalen. For å nytta programmet må vi ordna observasjonane i kvart utval etter storleik, altså etter talet på ferieturdagar. I vår situasjon har vi 95 kategorier (ulike observasjonsverdiar), nemleg frå 5 til 99 ferieturdagar. Vi let aldersgruppene 35-74 år og 15-34 år vere respektive referanse- og testgruppe.

Tabell 2. Datautlistinga frå programmet for Wilcoxon testen

Gruppestørrelse (tallet på kategorier maks. 100) = 95

Inputdata: Referansegruppa først, testgruppa sist

71	21	52	41	7	41	20	27	17	70	25	26	19	17	35	32	62	27	10	16	10
24	26	40	25	25	18	13	8	14	24	16	7	18	6	12	5	8				
1	6	8	3	4	2	5	0	1	4	3	1	2	3	1	0	1	3	1	2	2
1	3	0	3	2	1	0	0	0	1	0	0	0	0	0	3					
0	1	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	1			
33	17	41	22	6	24	10	26	14	46	13	13	6	14	15						
22	30	12	18	19	11	16	18	23	5	10	11	10	19	8	12					
8	8	4	4	7	4	5	3	4	6	3	3	1	4	5	2	4	1	2	2	3
2	2	3																		
1	2	0	0	1	4	2	0	0	1	2	2	1	1	2	0	1	0	0	0	0
0	0	3	0	0																
0	0	0	0	1	0	1	0	0	0	0	0	0	0							

$W = 563339.00$ $EW = 549606.00$ $VARW = 92209218.00$ $(W+0.5-EW)/STDW = 1.43$

Frå datautlistinga (Tabell 2) ser vi at for referansegruppa er det 71 personar som har 5 ferieturdagar, 21 som har 6 ferieturdagar osv. Programmet reknar så ut $\frac{W+0.5-EW}{\sqrt{\text{var } W}}$, som får verdien 1.43

i vårt tilfelle. Sidan talet på observasjonar i referansegruppa og testgruppa er respektive 1008 (m) og 659 (n) så vil tilnærminga $P\left(\frac{W+0.5-EW}{\sqrt{\text{var } W}} \leq k\right) \approx \Phi(k)$ vere god. Φ er normalfordelingen med forventning

0 og varians lik 1. Vi kan nå finna kritisk verdi k for testen i tabellar over normalfordelingen. Med nivå 0.05 er den kritiske verdi $k=1.64$. Men sidan testobservatoren blir lik 1.43 får vi ikkje forkastning av hypotesen H_0 : "ingen forskjell i talet på ferieturdagar mellom dei to aldersgruppene" mot alternativet A: "Talet på ferieturdagar er større for personar i aldersgruppa 15-34 år".

I kapittel 4 har vi nytta metodar som har gitt oss informasjon om kva slags fordeling observasjonane kan stamma i frå. Resultata frå denne analysen gir forkastning av hypotesen at observasjonane er normalfordelte. I tillegg får vi indikasjonar på at gammafordelingen med parametrar $\alpha_1=2.38$ og $\beta_1=9.30$ gir ei god skildring av materialet frå den eldste aldersgruppa. Tabell 1 frå kapittel 2a gir at asymptotisk effisiens for Wilcoxon test med omsyn på Student t-test blir 1.42 når observasjonane stammer frå gammafordelingen med parametrar $\alpha_1=2.38$ og $\beta_1=9.30$. Dette tyder at vi ville "trenge" 42% fleire observasjonar ved å nytta Student t-test i staden for Wilcoxon test for å oppnå same styrke gitt same alternativ og nivå. I praksis kunne dette medføre forkastning av H_0 : "ingen forskjell i talet på ferieturdagar" hvis ein nytta Wilcoxon test, medan Student t-test ikkje ville gi forkastning av H_0 i same situasjon.

3. Kruskal-Wallis testen, ein ikkje-parametrisk metode for samanlikning av fleire populasjonar

I kapittel 2 drøfta vi samanlikning av to populasjonar. Vi skal nå visa korleis ein kan samanlikna fleire populasjonar for å avgjera om dei er ulike.

La X_{i1}, \dots, X_{in_i} vere utval nr. i med kumulativ fordelingsfunksjon F_i , $i=1, 2, \dots, P$. Vi ønskjer å testa

$$H_0: F_1=F_2=\dots=F_P \text{ mot } A_1: F_i < F_j \text{ for minst eit par } (i, j) \text{ og } F_i \leq F_j \text{ for alle } (i, j).$$

Dette vil seie at F-ene ikkje får "kryssa" kvarandre under A_1 .

La R_{i1}, \dots, R_{in_i} vere rangane i utval nr. i når alle observasjonane (X_{ij} , $j=1, \dots, n_i$, $i=1, \dots, P$) er rangert saman. La vidare $R'_{i1}, \dots, R'_{in_i}$ vere " R_{i1}, \dots, R_{in_i} " ordna. For å testa H_0 mot A_1 vil vi samanlikna

$$R'_i = \frac{1}{n_i} \sum_{j=1}^{n_i} R'_{ij} \text{ med } R'_{00} = \frac{1}{N} \sum_{i=1}^P \sum_{j=1}^{n_i} R'_{ij} = \frac{1}{N} \sum_{\ell=1}^N \ell = \frac{1}{2} (N+1),$$

dvs. vi vil samanlikna gjennomsnittlege ranger for kvar av gruppene med gjennomsnittet av alle rangar. Kruskal-Wallis (1952) føreslo observatoren

$$K' = \frac{12}{N(N+1)} \sum_{i=1}^P n_i (R'_i - \frac{1}{2} (N+1))^2 = \frac{12}{N(N+1)} \sum_{i=1}^P \frac{R_i^2}{n_i} - 3(N+1).$$

Vi forkastar H_0 dersom $K' \geq$ konstant. For $P=2$ er denne testen ekvivalent med den to-sidige Wilcoxon-testen.

Under føresetnad av at alle observasjonane er genererte frå same fordeling (altså under H_0) er fordelingen til K' tabellert i Owen (1962) for $P=3$. Årsaken til at ein ikkje har tabellar for fleire P , er at det fins følgjande gode tilnærming:

Under H_0 har vi at $K' \xrightarrow{d} \chi^2_{P-1}$ (χ^2 -kvadrat fordeling med $P-1$ fridomsgrader), der \xrightarrow{d} tyder konvergens i fordeling. For $P=3$ er tilnærmingen brukbar når $n_i \geq 6$, og for $P > 3$ når $n_i \geq 5$.

Dersom det fins observasjonar med like verdier i materialet vårt, kan vi nytta midranger analogt med Wilcoxon-observatoren i kapittel 2. La R_i vere summen av dei ordna midrangane i utval i , og la e og t_i vere definert som tidlegare. Da blir Kruskal-Wallis observatoren definert ved

$$K = \frac{\frac{12}{N(N+1)} \sum_{i=1}^P \frac{(R_i)^2}{n_i} - 3(N+1)}{1 - \frac{\sum_{i=1}^e (t_i^3 - t_i)}{N^3 - N}}$$

Hvis det ikkje fins like verdjar i observasjonsmaterialet, ser vi lett at $K=K'$. Vi skal nå handsama forskyvningsmodellen $F_i(x) = F_0(x - \theta_i)$, $i=1, \dots, P$ der F_0 er ein kontinuerlig kumulativ fordelingsfunksjon. Vi skal teste $H_0: \theta_1 = \theta_2 = \dots = \theta_P$ mot $A_1: \theta_i \neq \theta_j$ for minst eit par (i, j) . Når $F_0 = \Phi$ blir Fisher's test (3) den klassiske testen i denne situasjonen. Denne er handsama under kapittel VII, 5. Sverdrup (1964 I). Andrews (1954) har vist at Pitman-effisiensen (definert i kap. 2) til K med omsyn på 3 vert

$$e_{K, 3} = 12 \int_{-\infty}^{\infty} \sigma^2 f^2(x) dx^2,$$

altså nøyaktig det same som vi oppnådde for Wilcoxon testen med omsyn på Student test.

Kruskal-Wallis testen har altså same gode eigenskapene som Wilcoxon-testen.

4. Eksempel på bruk av grafiske metodar for å vurdere modellar

Vi skal nå nytta metodar som er omtala i appendiks 2 for å skaffa oss informasjon om fordelingen til ferieturmaterialet. For å skaffa oss informasjon om fordelingen (F) til observasjonane X_1, \dots, X_m i aldersgruppa 35-74 år, konstruerer vi først eit plot av punktane

$$(X_{(i)}, \Phi^{-1}(F_m(X_{(i)}) - \frac{1}{2m})) = (X_{(i)}, \Phi^{-1}(\frac{i - \frac{1}{2}}{m})),$$

der $X_{(i)}$ er ordningsobservator nummer i og $F_m(x) = \frac{1}{m}$ [talet på X_i - er $\leq x$]. Dette plottet blir kalla det normale sannsynplottet. $\Phi^{-1}(F_m(x) - \frac{1}{2m})$ er ein konsistent estimator for $\Phi^{-1}(F(x))$. Dersom $F(x) = \Phi(\frac{x - \mu}{\sigma})$ (normalfordelingen med forventning μ og varians σ^2) så vert $\Phi^{-1}(F(x))$ den rette linja $\frac{x - \mu}{\sigma}$. Ved hjelp av dette og andre merknader omkring $\Phi^{-1}(F(x))$ for ulike fordelingsklassar F (sjå appendiks 2), kan vi gi tolkningar av det normale sannsynplottet.

På grunnlag av observasjonsmaterialet i aldersgruppa 35-74 år får vi eit plot som gir ein figur med konkav form. Dette indikerer at fordelingen (F) er skeiv til høgre. Vi får vidare at høgre hale er tjukkare enn høgre hale til normalfordelingen og venstre hale tunnare. Sidan plottet indikerer at F ikkje er normalfordelt, skal vi testa hypotesen $H_0: F(x) = \Phi(\frac{x - \mu}{\sigma})$. I appendiks 2 har vi konstruert ein klasse av konfidensband for $\Phi^{-1}(F(x))$. Desse banda kan nyttast til å testa H_0 mot $A: F(x) \neq \Phi(\frac{x - \mu}{\sigma})$. Vi får forkastning av H_0 dersom det bandet vi baserer oss på ikkje kan innehalda ei rett linje. Ved å nytta vektfunksjonen $\psi(F_m(x)) = 1$ for alle x ($a=0, b=1$), får vi konfidensbandet $[\Phi^{-1}(F_m(x) - \frac{k}{\sqrt{m}}),$

$\Phi^{-1}(F_m(x) + \frac{k}{\sqrt{m}})]$ for alle x . Det er gitt ei grafisk framstilling av dette bandet og estimatoren $\Phi^{-1}(F_m(x) - \frac{1}{2m})$ for ferieturmaterialet (aldersgruppa 35-74 år) i figur 1. Konfidensbandet kan ikkje

innehalda ei rett linje og vi får derfor forkastning av hypotesen at observasjonane stammer frå normalfordelingen.

Sidan klassen av gammafordelingar er skeiv til høgre og dessutan relativt fleksibel, legg vi denne klassen av fordelingar til grunn for å skildra ferieturmaterialet. Gammafordelingen er definert ved (2.3). Vi vil nytta momentmetoden for å estimera parametrane α_1 og β_1 i gammafordelingen. Denne metoden har rekнемessige fordelar og gir dessutan konsistente estimatorer under relativt generelle føresetnader. Frå Sverdrup (1964), som gir ei handsaming av metoden, får vi

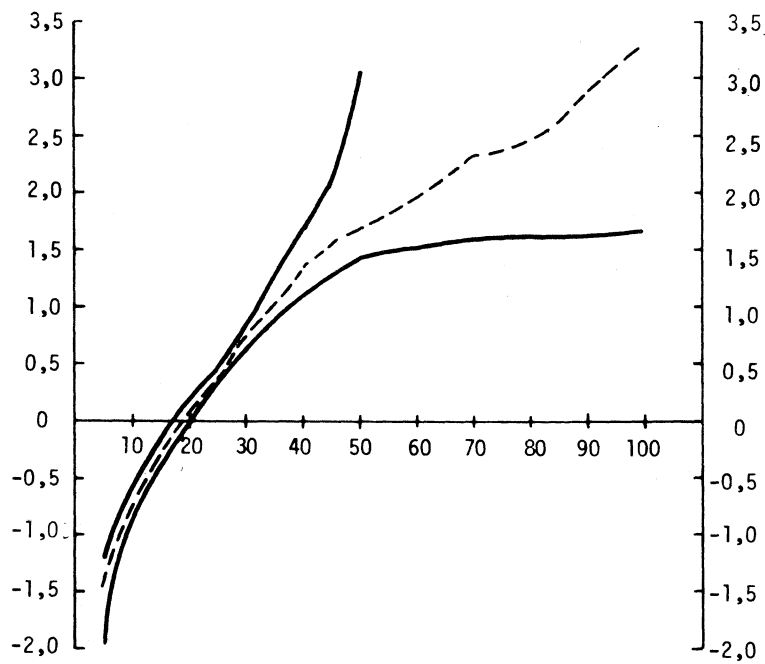
$$\hat{\alpha}_1 = \frac{\bar{X}^2}{S^2} \text{ og } \hat{\beta}_1 = \frac{S^2}{\bar{X}}$$

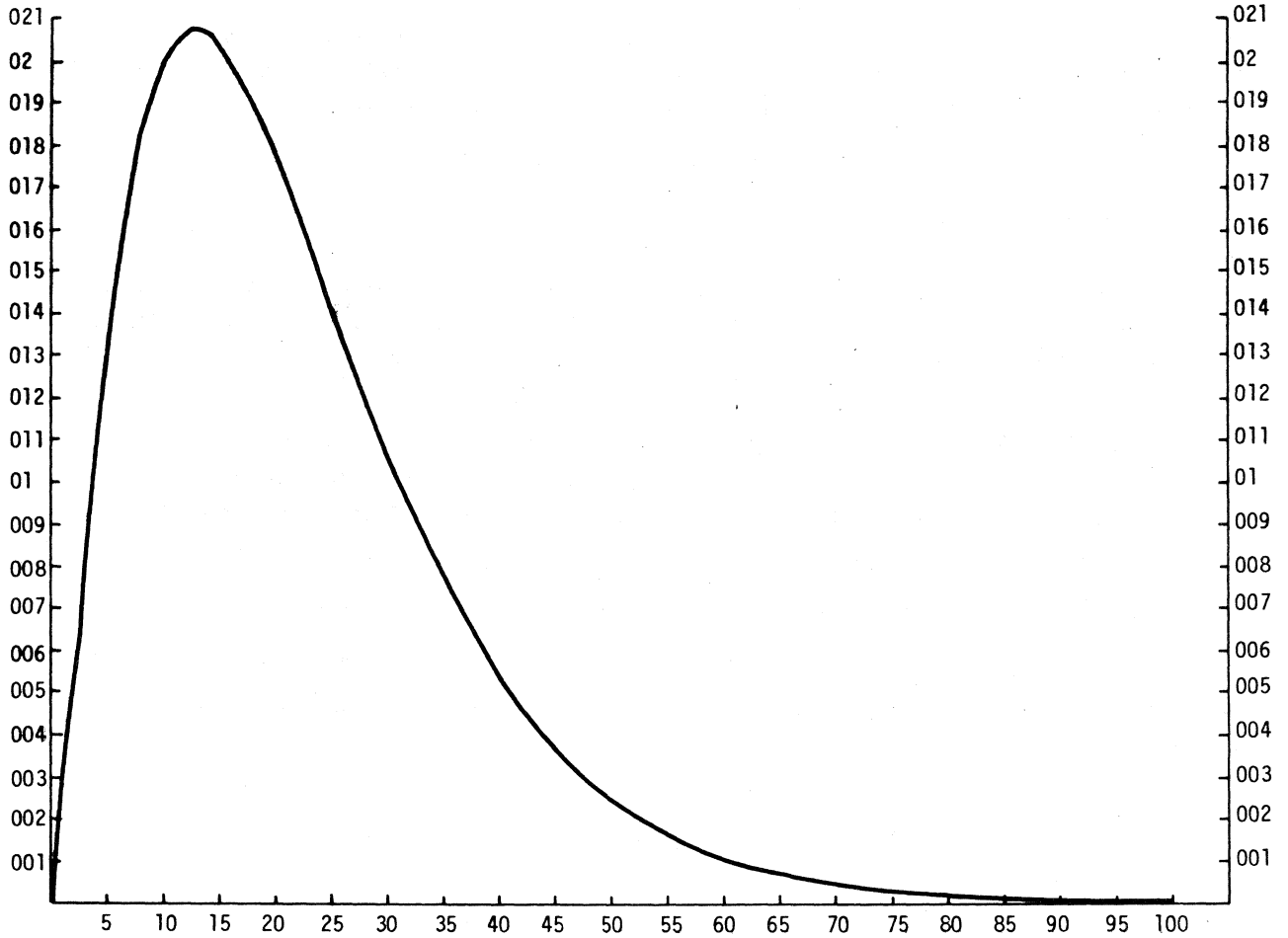
der

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i, \quad S^2 = \frac{1}{m} \sum_{i=1}^m (X_i - \bar{X})^2.$$

Materialet gir $\hat{\alpha}_1 = 2.38$ og $\hat{\beta}_1 = 9.30$. For disse verdiane av α_1 og β_1 har vi gitt ei grafisk framstilling av $f(x)$ i figur 2.

Figur 1. $\phi(F_m(x) - \frac{1}{2m})$ (prikka linje) og konfidensbandet
 $[\phi(F_m(x) - \frac{k}{\sqrt{m}}), \phi(F_m(x) + \frac{k}{\sqrt{m}})]$ (heiltrukne linjer)



Figur 2. Grafisk framstilling av gammafordelingen med parametere $\alpha=2.38$ og $\beta=9.30$ 

5. Samanlikning av to populasjonar når modellen er ikkje-lineær

Vi skal som tidlegare ta føre oss to-utvalsproblemet, altså samanlikning av to populasjonar med fordelingsfunksjonar F og G på basis av dei uavhengige utvala X_1, \dots, X_m og Y_1, \dots, Y_n . I staden for den lineære modellen $F(x)=G(x+\theta)$ for alle x skal vi ta føre oss den meir generelle situasjonen

$$(5.1) \quad F(x)=G(x+\Delta(x)) \text{ for alle } x \text{ for ein eller annan funksjon } \Delta(x).$$

Hvis den lineære modellen ikkje er tilfredsstillt, kan $\Delta(x)$ bli nytta som eit mål for differensen mellom dei to populasjonane. Det er då rimeleg å krevja at $x+\Delta(x)$ skal vere ikkje-avtakande. Doksum (1974) har vist at det eksisterer berre ein funksjon $\Delta(x)$ i (5.1) som tilfredsstiller dette kravet, og vidare at denne funksjonen er definert ved

$$(5.2) \quad \Delta(x)=G^{-1}(F(x))-x.$$

Vi merker oss at $\Delta(x)=\theta$ for alle x når den lineære modellen er oppfylt. Ved å nytta funksjonen $\Delta(x)$ definert ved (5.2), skal vi gi metodar for å avgjera om den lineære modellen er tilfredsstillt og dessutan gi statistiske metodar for samanlikning av dei to populasjonane.

5a. Estimering i generelle modeller

Hvis dei parametriske formene til fordelingsfunksjonane F og G er ukjende, er det naturleg å nytta dei empiriske fordelingsfunksjonane F_m og G_n som estimatorer for F og G . Vi nyttar derfor $\hat{\Delta}(x)=G_n^{-1}(F_m(x))-x$ som estimator for $\Delta(x)$, der venstre inversen G_n^{-1} er definert ved $G_n^{-1}(u)=\inf \{y: G_n(y) \geq u\}$. Ved å la $X(1) \leq \dots \leq X(m)$, og $Y(1) \leq \dots \leq Y(n)$ vere ordningsobservatorene for utvala X og Y og $t < \dots < m$, får vi $\hat{\Delta}(x)=Y(\lfloor n \frac{x}{m} \rfloor) - x$ når $x \in [X(i), X(i+1))$. For å unngå underestimering av $\Delta(x)$ kan vi erstatte venstreinversen G_n^{-1} med gjennomsnittet av venstre og høgre inversen, dvs. vi nyttar estimatoren

$$\hat{\Delta}(x) = \frac{1}{2}(G_n^{-1}(F_m(x)) + G_n^{-1}(F_m(x))) - x = \frac{1}{2} \{Y(\lfloor n \frac{x}{m} \rfloor) + Y(\lceil n \frac{x}{m} \rceil)\} \text{ når } x \in [X(i), X(i+1))$$

der

$$G_n^{-1}(u) = \sup \{y: G_n(y) \leq u\} \text{ og } \lceil t \rceil = \text{største heile tal } \leq t.$$

La $M = \frac{m \cdot n}{m+n}$. Frå Doksum (1974) har vi følgjande uttrykk for den asymptotiske fordelingen til $\hat{\Delta}(x)$:

$$\sqrt{M}(\hat{\Delta}(x) - \Delta(x)) \xrightarrow{d} \frac{W_0(F(x))}{g(G^{-1}(F(x)))}$$

under føresetnad av at $G(y)$ har kontinuerleg derivert $g(y)$ og $0 < g(y) < \infty$ på $\{y: 0 < F(y) < 1\}$. W_0 er Brownske brua, dvs. W_0 er ein Gaussisk prosess der $E\{W_0(t)\} = 0$ og $E\{W_0(s)W_0(t)\} = s(1-t)$, $0 \leq s < t \leq 1$. Vi seier at prosessen W_0 er Gaussisk hvis alle dei endelegdimensjonale fordelingane til prosessen er normale.

Frå uttrykket til den asymptotiske fordelingen til $\hat{\Delta}(x)$ følgjer at

$$\text{var } \hat{\Delta}(x) \approx \frac{F(x)(1-F(x))}{Mg^2(G^{-1}(F(x)))} \text{ for store } m \text{ og } n.$$

På grunnlag av dette resultatet kan vi vurdere kor nøyaktig $\Delta(x)$ vert for ulike F og G.

5b. Konfidensband for $\Delta(x)$

Sidan $\Delta(x) = G^{-1}(F(x)) - x$ er eit mål for differensen mellom to populasjonar kan vi for eksempel vere interessert i å avklara følgjande spørsmål:

- (i) Kva er $\{x: \Delta(x) > 0\}$?
- (ii) Er $\Delta(x)$ lik ein eller annan konstant θ for alle x ?
- (iii) Er $\Delta(x)$ ein lineær funksjon i x ?

Ved å konstruera eit simultant konfidensband $[\underline{\Delta}(x), \bar{\Delta}(x)]$ for $\Delta(x)$ kan vi knytta sannsynsutsagn til desse spørsmåla. Løysinga på spørsmål (i) er: $\{x: \Delta(x) > 0\}$. Vi kan svara nektande på (ii) (hypotesen om forskyvningsmodellen) dersom konfidensbandet ikkje kan innehalda ei horisontal linje. Tilsvarande kan vi nekta for (iii) (hypotesen om lokasjon-skala modellen) dersom bandet ikkje kan innehalda ei rett linje.

La F og G vere kontinuerlege fordelingsfunksjoner. Vi skal gi ein omtale av korleis vi finn konfidensband for $\Delta(x)$. La $G_{\Delta, n}(y) = G_n(y + \Delta(y))$ og la vidare $\Psi(F_m, G_n)$ vere ein testfunksjon med nivå α for å testa $H_0: F = G$. Da er $\{\Delta(\cdot): \Psi(F_m, G_{\Delta, n}) = 0\}$ eit konfidensområde for $\Delta(\cdot)$. Det er altså nytta same ide som blir brukt for å finna konfidensintervall for ein parameter. Dersom vi baserer oss på testobservatoren $A(F_m, G_n)$ med eigenskapene at ulikskapen $A(F_m, G_n) \leq k$ er ekvivalent med

- (5.3) $\underline{h}(F_m(x)) \leq G_n(x) \leq \bar{h}(F_m(x))$ for alle x og ein eller annan \underline{h} og \bar{h} , da vil desse områda bli reduserte til einkle band. Vi utviklar nå konfidensbandet for $\Delta(x)$ frå (5.3) på følgjande måte:

La G_n^{-1} og G_n^{-I} vere respektive venstre- og høgreinversen til G_n , og la k vere valt slik at

- (5.4) $P_{F=G}(A(F_m, G_n) \leq k) = 1 - \alpha$ (fordelingen til $A(F_m, G_n)$ under H_0).

Sidan $G_{\Delta, n}(y) = G_n(y + \Delta(y)) = \frac{1}{n} [\text{talet på } Y_i \leq y + \Delta(y)] = \frac{1}{n} [\text{talet på } Y_i \leq G^{-1}(F(y))]$

$\frac{1}{n} [\text{talet på } F^{-1}(G(Y_i)) \leq y]$ og $F^{-1}(G(Y_i))$ har fordeling F, så har vi vidare

$$\begin{aligned} 1 - \alpha &= P_{F=G}(A(F_m, G_n) \leq k) = P_{F, G}(A(F_m, G_{\Delta, n}) \leq k) = \\ &P_{F, G}(\underline{h}(F_m(x)) \leq G_{\Delta, n}(x) \leq \bar{h}(F_m(x)), \text{ for alle } x) = \\ &P_{F, G}(\underline{h}(F_m(x)) \leq G_n(x + \Delta(x)) \leq \bar{h}(F_m(x)), \text{ for alle } x) = \\ &P_{F, G}(G_n^{-1}(\underline{h}(F_m(x))) - x \leq \Delta(x) \leq G_n^{-I}(\bar{h}(F_m(x))) - x, \text{ for alle } x) \end{aligned}$$

Vi har da følgjande setning frå Doksum & Sievers (1976):

Setning 1: Dersom (5.3) og (5.4) er oppfylt så er

$$(5.5) \quad [G_n^{-1}(\underline{h}(F_m(x))) - x, G_n^{-I}(\bar{h}(F_m(x))) - x] = [\underline{\Delta}(x), \bar{\Delta}(x)] \quad x \in \langle -\infty, \infty \rangle$$

eit simultant fordelingsfritt konfidensband for $\Delta(x)$ med nivå $1 - \alpha$.

Fordelingsfritt vil sei at nivået til konfidensbandet er det same for alle kontinuerlege F og G, dvs. at nivået ikkje vert forstyrra av formen på fordelingsfunksjonane F og G.

Ved hjelp av ordningsobservatorene kan bandet (5.5) uttrykkjast ved

$$[\underline{\Delta}(x), \bar{\Delta}(x)] = [Y(\langle n \underline{h}(\frac{1}{m}) \rangle) - x, Y(\langle n \bar{h}(\frac{1}{m}) \rangle + 1) - x] \text{ når } x \in [X(i), X(i+1)), i=0, \dots, m \text{ med } X(0) = -\infty \text{ og}$$

$$X(m+1) = \infty.$$

La for eksempel $N=m+n$, $M=\frac{mn}{N}$ og $A(F_m, G_n)=\sqrt{M} \sup |F_m(x)-G_n(x)|$ som er Kolmogorov-Smirnov observatoren for to utval. Da blir $h(t)=t-k/\sqrt{M}$, og $\bar{h}(t)=t+k/\sqrt{M}$, der k vert fastsett frå $P(A(F_m, G_n) \geq k) = \alpha$ når α er gitt.

Doksum & Sievers (1976) har handsama konfidensband for $\Delta(x)$ basert på observatorer av typen

$$(5.6) \quad \sqrt{M} \sup_{\{x: a \leq F_m(x) \leq b\}} \frac{|F_m(x) - G_n(x)|}{\phi(H_N(x))}$$

der

$H_N(x) = \lambda F_m(x) + (1-\lambda)G_n(x)$, $\lambda = \frac{m}{N}$, $M = \frac{mn}{N}$, a, b konstnatar og ϕ er ein vektfunksjon.

I situasjoner der F og G er diskrete fordelingar, dvs. vi får observasjonar med like verdjar, vert konfidensbanda for $\Delta(x)$ konservative. Det vil sei at

$$P(\underline{\Delta}(x) \leq \Delta(x) \leq \bar{\Delta}(x)) \geq 1 - \alpha.$$

Noether (1963) har handsama desse problema meir detaljert.

5c. Eksempel på samanlikning av reiseaktiviteten mellom to aldersgrupper når modellen er ikkje-lineær

Vi let som tidlegare X_1, \dots, X_m vere dei $m=1008$ observasjonane i aldersgruppa 35-74 år og Y_1, \dots, Y_n dei $n=659$ observasjonane i gruppa 15-35 år. Ved hjelp av skiftfunksjonen $\Delta(x)$, som altså er eit mål for differensen mellom dei to populasjonane, skal vi gi ein statistisk analyse av materialet. Det er her viktig å leggja merke til at den ikkje-lineære modellen $Y = X + \Delta(X)$ (= tyder likskap i fordeling og $\Delta(x)$ er definert ved (5.2)) ikkje krev noko føresetnader om observasjonsmaterialet, i motsetnad til den lineære modellen som er handsama i kapittel 2. Dersom vi legg den lineære modellen til grunn for analysen av materialet, kan det dessutan vere interessante forskjeller mellom populasjonane som vi ikkje vil oppdaga, sidan vi da har gått ut ifrå at dei ikkje kan vera der.

Ved å la $\phi(t) = \sqrt{t(1-t)}$ i (5.6) (sjå kap. 5b) får vi ein testobservator som gir tilnærma lik vekt til kvar x i den forstand at $\sqrt{M} \left\{ \frac{F_m(x) - G_n(x)}{\phi(H_N(x))} \right\}$ har asymptotisk varians som er uavhengig av x . Vi skal derfor konstruere eit konfidensband for $\Delta(x)$ basert på testobservatoren

$$V_N = V_N(F_m, G_n) = \sqrt{M} \sup_{\{x: a \leq F_m(x) \leq b\}} \left\{ \frac{|F_m(x) - G_n(x)|}{\sqrt{H_N(x)(1-H_N(x))}} \right\}$$

Saman med estimatoren $\hat{\Delta}(x)$ gir bandet informasjon om $\Delta(x)$, som altså er eit ikkje-lineært mål for differensen mellom dei to populasjonane.

For å nytta Setning 1 (kap. 5b) må vi løysa ulikskapen

$$(5.7) \quad V_N(F_m, G_n) \leq k.$$

Ved å la $c = k^2/M$, $u = F_m(x)$, og $\lambda = \frac{m}{N}$ får vi følgjande løysing på (5.7) frå Doksum & Sievers (1976),

$$h^-(F_m(x)) \leq G_n(x) \leq h^+(F_m(x))$$

der

$$h^\pm(u) = \frac{u \pm c(1-\lambda)(1-2\lambda u) \pm \sqrt{c^2(1-\lambda)^2 + 4cu(1-u)}}{1+c(1-\lambda)^2}$$

Da har vi frå Setning 1 at $\Delta(x)$ er i bandet

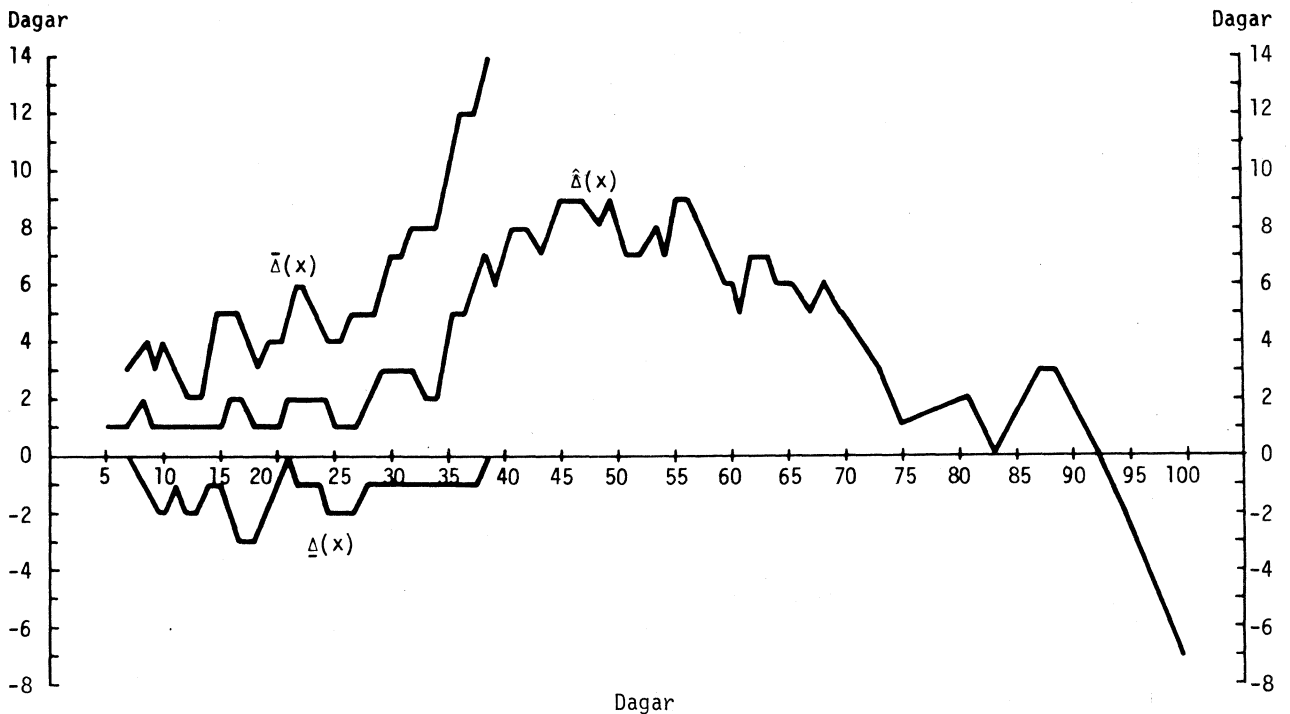
$$(5.8) \quad [\underline{\Delta}(x), \bar{\Delta}(x)] = [G_n^{-1}(h^-(F_m(x))) - x, G_n^{-1}(h^+(F_m(x))) - x]$$

med sannsyn $1-\alpha$ for alle $x \in \{x: a \leq F_m(x) \leq b\}$.

For gitt α , kan vi frå tabell i Borokov & Sycheva (1968) finna asymptotisk kritisk verdi k for ulike verdier av a og b der $0 < a < b < 1$. Ved å velga $a=1-b=0.10$, får vi eit konfidensband for $\Delta(x)$ gjeldande for verdier i intervallet $7 \leq x \leq 38$. Vi vil altså få sannsynutsagn om eventuelle forskjeller i talet på ferieturdagar, $\Delta(x)$, mellom dei to aldersgruppene når talet på ferieturdagar (x) for den eldste gruppa legg i intervallet $[7, 38]$.

I figur 3 har vi gitt ein illustrasjon av bandet (5.8) når $a=1-b=0.10$ og ferieturmaterialet fra ferieundersøkinga 1974 er nytta. Vi har altså $m=1008$, $n=659$ og for nivået $\alpha=0.04$ oppnår vi asymptotisk kritisk verdi $k=3.1384$ frå Borokov & Sycheva (1968).

Figur 3. Estimatoren $\hat{\Delta}(x)$ og konfidensbandet $[\underline{\Delta}(x), \bar{\Delta}(x)]$ for $\Delta(x)$



På grunnlag av formen til $\hat{\Delta}(x)$ (figur 3) får vi indikasjoner på at det er variasjon med omsyn på forskjellen i talet på ferietur dagar mellom dei to aldersgruppene. Det ser ut til at dei minste og mellomstore tala på ferietur dagar i den yngste aldersgruppa er større enn tilsvarande for den eldste gruppa, medan dei lengste ferieturane i den yngste gruppa er mindre enn tilsvarande for den eldste gruppa. Grunnlaget for denne tolkningen ligg i $\Delta(x)$, som måler forskjellen i talet på ferietur dagar mellom dei to aldersgruppene som ein funksjon av responsane (x) i aldersgruppa 35-74 år. Vi ser at den estimerte forskjellen $\hat{\Delta}(x)$ har ein tendens til å auka for $x \in [7, 50]$. For større x avtek denne tendensen. For spesielt store x blir $\hat{\Delta}(x) < 0$. Sidan $\underline{\Delta}(x) \ll 0$ og $\bar{\Delta}(x) \gg 0$ gir ikkje materialet grunnlag for å påstå at det er nokon signifikante forskjeller i talet på ferietur dagar mellom dei to gruppene. Vidare ser vi at konfidensbandet kan innehalda ei horisontal linje. Dermed kan vi ikkje forkasta den lineære modellen, $\Delta(x) = \theta$ for ein eller annan θ og for alle x . Det var denne modellen som låg til grunn for samanlikningen av Student t-test og Wilcoxon test i kapittel 2a.

TESTING AV MODELLAR

Vi skal her gi ei kort utgreiing om observatorer som er nyttige i samband med å testa om eit utval X_1, \dots, X_n av uavhengige stokastiske variable er generert frå ein kjend kontinuerleg fordelingsfunksjon F .

La $X_i \sim F$, $i=1, \dots, n$ der F er ein ukjend kontinuerleg fordelingsfunksjon. Vi ønskjer altså å testa nullhypotesen $H_0: F=F_0$ mot $A: F \neq F_0$. Dette kan gjerast ved å ta utgangspunkt i observatorar av typen

$$W_n = \sqrt{n} \sup_{\{x: a \leq F_0(x) \leq b\}} \left\{ \frac{|F_n(x) - F_0(x)|}{\Psi(F_0(x))} \right\}$$

der Ψ er ein vektfunksjon, a og b konstanter slik at $0 < a < b < 1$ og $F_n(x) = \frac{1}{n} \cdot [\text{talet på } X_i, i=1, \dots, n \leq x]$.

Vi forkaster H_0 dersom $W_n \geq k$.

La $a=0$, $b=1$ og $\Psi=1$. Då vert W_n redusert til Kolmogorov-Smirnov observatoren

$$D_n = \sqrt{n} \sup_x |F_n(x) - F_0(x)|.$$

Denne observatoren gir størst vekt til observasjonane i den sentrale delen av fordelingen F i den forstand at $\text{var} \sqrt{n} \{F_n(x) - F(x)\} = F(x)(1-F(x))$. Frå tabell (f.eks. Owen (1962)), kan vi finna k slik at for gitt α er $P_{H_0}(D_n \geq k) = \alpha$.

Dersom vi ønskjer at W_n skal gi lik vekt til alle observasjonane i den forstand at

$\text{var} \left\{ \frac{F_n(x) - F(x)}{\Psi(F(x))} \right\} = \text{konstant}$, kan vi velja $\Psi(u) = \sqrt{u(1-u)}$. Vi får da observatoren

$$T_n = \sqrt{n} \sup_{\{a \leq F_0(x) \leq b\}} \frac{|F_n(x) - F_0(x)|}{\sqrt{F_0(x)(1-F_0(x))}}.$$

For gitt nivå, α og $0 < a < b < 1$ kan vi finna eksakt eller asymptotisk kritisk verdi k slik at $P(T_n \geq k) = \alpha$ i Brokov & Sycheva (1968). Monte Carlo verdier av T_n når $a=1-b=0$ er gitt i Canner (1975). Frå Borokov & Sycheva (1968) følgjer det dessutan at om vi handsamar klassen av ein-sidede testobservatorer

$$(1) \quad \sqrt{n} \sup_{\{x: a \leq F_0(x) \leq b\}} \left\{ \frac{F_n(x) - F_0(x)}{\Psi(F_0(x))} \right\}$$

med $0 < a < b < 1$, så vil valget $\Psi(u) = \sqrt{u(1-u)}$ i (1) asymptotisk maksimera den minste styrken når vi tester $H: F \neq F_0$ mot $A: F(x) - F_0(x) \geq \delta$ for ein eller annan x og $\delta > 0$.

Ved å la $\Psi(u) = u$ og $\Psi(u) = 1-u$ i W_n får vi testobservatorer som gir størst vekt til respektive venstre og høgre hale til F . Følgjeleg vil observatoren

$$r_n' = \sqrt{n} \sup_{\{x: a \leq F_0(x)\}} \left\{ \frac{|F_n(x) - F_0(x)|}{F_0(x)} \right\}$$

gi størst vekt til dei småe observasjonane og

$$r_n'' = \sqrt{n} \sup_{\{x: F_0(x) \leq b\}} \left\{ \frac{|F_n(x) - F_0(x)|}{1-F_0(x)} \right\}$$

til dei store.

I Rényi (1953) er den asymptotiske fordelingene for r'_n tabellert for nokre verdier av a og dermed for r''_n med $b=1-a$.

Klassen

$$W'_n = \sqrt{n} \sup_{\{a \leq F_n(x) < b\}} \left\{ \frac{|F_n(x) - F_0(x)|}{\psi(F_n(x))} \right\}$$

av testobservatorer er asymptotiskekvivalent med W_n . Altså har vi under H_0 at $\lim_{n \rightarrow \infty} P(W'_n \geq k) = \lim_{n \rightarrow \infty} P(W_n \geq k)$.

Av reknemessige årsaker vil vi ofte basere testing av $H_0: F=F_0$ mot $A: F \neq F_0$ på klassen W'_n istadenfor W_n . Det særmerkte med desse testane går ut på at dei gir eit simultant band for $F(x)$, slik at sannsynet for at $F(x)$ skal ligge i bandet er $1-\alpha$ når nullhypotesen er sann. α er nivået til testen. Vi har altså at metoden vi nyttar er slik at sannsynet er $1-\alpha$ for å komme fram til riktig konklusjon.

La $\phi(F_n)$ vere ein fordelingsfri test med nivå α for å testa $H_0: F=F_0$ mot $A: F \neq F_0$, definert ved

$$\phi(F_n) = \begin{cases} 1 & \text{når } W'_n \geq k \\ 0 & \text{" } W'_n < k. \end{cases}$$

Då vert $\{F: \phi(F_n)=0\}$ eit simultant fordelingsfritt, $1-\alpha$ nivå konfidensband for F .

Enkel rekning gir at

$$\{F(x): \phi(F_n)=0\} = \{F(x): F_n(x) - k \cdot \psi(F_n(x)) / \sqrt{n} \leq F(x) \leq F_n(x) + k \cdot \psi(F_n(x)) / \sqrt{n}\}$$

for alle $x \in [F_n^{-1}(a), F_n^{-1}(b)]$, der inversane er definert ved $F_n^{-1}(a) = \inf\{x: F_n(x) \geq a\}$ og

$F_n^{-1}(b) = \sup\{x: F_n(x) \leq b\}$. Vi forkaster H_0 dersom $F_0(x)$ ikkje er innehelden i konfidensbandet. Dersom vi

for eksempel let

$$W'_n = D_n = \sqrt{n} \sup_x |F_n(x) - F_0(x)| \quad (\psi=1, a=1-b=0),$$

så vert altså

$$\left[F_n(x) - \frac{k}{\sqrt{n}}, F_n(x) + \frac{k}{\sqrt{n}} \right]$$

eit simultant fordelingsfritt konfidensband for F med nivå $1-\alpha$.

DET NORMALE SANNSYNNPLOTET

La X_1, \dots, X_n vere eit sett uavhengige stokastiske variable med kumulativ fordelingsfunksjon F . Dersom den parametriske formen til F er ukjend, er det naturleg å nytta den empiriske fordelingsfunksjon F_n som estimator for F . $F_n(x) = \frac{1}{n}$ [talet på X_i -er, $i=1, \dots, n \leq x$]. Vi har dermed at $nF_n(x) \sim$ binomisk $(n, F(x))$. Altså vert $\text{var}(F_n(x)) = \frac{1}{n} F(x)(1-F(x))$. Vi ser at estimatoren er mest precis i halene, dvs. for dei minste og dei største observasjonane. I Appendiks 1 gav vi eksempel på konstruksjon av konfidensband rundt $F(x)$. Desse kan nyttast til å vurdere kor påliteleg estimatoren $F_n(x)$ er.

Eit viktig problem knytta til $F_n(x)$, er vanskan med å finna ut om den er konsistent med ein gitt parametrisk fordelingsfunksjon. I mange situasjonar vil vi vere særleg interessert i normalfordelingen, da mange metodar byggjer på føresetnaden om at observasjonane stammer frå denne fordelingen. Sidan det er lett å avsløre avvik frå ei rett linje, transformerer vi derfor F slik at vi får ei rett linje når F er normal. Den nemnde transformasjonen er altså $p(x) = \Phi^{-1}(F(x))$. Når $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ (normalfordelingen med forventning μ og varians σ^2), så vert $p(x)$ den rette linja $\frac{x-\mu}{\sigma}$. $p_n(x) = \Phi^{-1}(F_n(x))$ er ein konsistent estimator for $p(x)$ basert på $F_n(x)$. For å unngå at $\Phi^{-1}(1) = \infty$, nyttar vi i staden $\Phi^{-1}(F_n(x) - \frac{1}{2n})$ som estimator for $p(x)$. Ved å la $X_{(1)} \leq \dots \leq X_{(n)}$ vere X_i -ene ordna, får vi $p_n(x) = \Phi^{-1}\left(\frac{i-\frac{1}{2}}{n}\right)$ når $x \in [X_{(i)}, X_{(i+1)}]$. Plottet av punktane $(X_{(i)}, p_n(X_{(i)}))$, $i=1, \dots, n$ vert kalla det normale sannsynplottet. Vi kan tolka slike plot ved å nytta følgjande resultat frå Van Zwet (1963):

(i) Dersom F er symmetrisk omkring μ og $\Phi^{-1}(F(x))$ er konveks for $x < \mu$ og konkav for $x > \mu$, så indikerer dette at F har tjukkare halar enn Φ (normalfordelingen).

Eksempel på dette er når F er logistisk fordeling, studentfordeling og

$$(2) \quad \frac{dF(x)}{dx} = f(x) = \frac{1}{2\pi(1+\frac{1}{\theta})} (e^{-|x|^\theta}) \text{ der } -\infty < x < \infty, \theta > 0.$$

Når $\theta < 2$ så har denne fordelingen tjukkare halar enn normalfordelingen.

(ii) Dersom F er symmetrisk omkring μ og $\Phi^{-1}(F(x))$ er konkav for $x < \mu$ og konveks for $x > \mu$, så indikerer dette at F har tunnare halar enn Φ . Eksempel på dette er den uniforme fordeling og klassen (2) når $\theta > 2$.

(iii) Hvis $\Phi^{-1}(F(x))$ er konkav heile vegen då er F skeiv til høgre. Høgre hale er tjukkare enn den venstre. Dessutan er høgre hale tjukkare enn høgre hale til normalfordelingen og venstre hale tunnare. Eksempel er eksponensial, gamma, lognormal og Weibul fordelingane.

La $Z_i = \Phi^{-1}(F(X_i))$, der F er kontinuert. Da har vi

$$P(Z_i \leq z) = P(\Phi^{-1}(F(X_i)) \leq z) = P(F(X_i) \leq \Phi(z)) = P(X_i \leq F^{-1}(\Phi(z))) = FF^{-1}(\Phi(z)) = \Phi(z).$$

Altså er $Z_i \sim \Phi$. Ved å gi ei grafisk framstilling av $\Phi^{-1}(F(x))$ for dei tre tilfella, får vi derfor lett verifisert resultatene i dei tre punkta ovanfor.

Van Zwet (1963) har også vist at det klassiske målet for tunge (tjukke) halar

$$\gamma_4 = \frac{E(X-\mu)^4}{\sigma^4} - 3$$

er positivt når $\Phi^{-1}(F(x))$ er konveks-konkav og negativt når $\Phi^{-1}(F(x))$ er konkav-konveks. Vidare har

$$\gamma_3 = \frac{E(X-\mu)^3}{\sigma^3},$$

som er eit mål for skeivheit, er positivt når $\phi^{-1}(F(x))$ er konkav heile vegen.

Ved sida av å gi informasjon om formen på fordelingen til observasjonane, gir plottet av $p_n(x)$ også informasjon om lokasjonen til F . Vi har nemleg at $p(x)=0$ hvis og bare hvis $F(x)=\frac{1}{2}$, dvs. x er ein median. Altså får vi at det punktet som $p_n(x)$ krysser null er eit estimat for medianen.

For å undersøka kor nøyaktig estimatoren $p_n(x)=\phi^{-1}(F_n(x))$ er, ser vi på den asymptotiske variansen σ_p^2 for $\sqrt{n} \phi^{-1}(F_n(x))$. Ved å Taylor-utvikla $\phi^{-1}(F_n(x))$ omkring $F(x)$ finn vi

$$\sigma_p^2 = \left[\frac{d}{du} \phi^{-1}(u) \Big|_{u=F(x)} \right]^2 F(x)(1-F(x)) = \frac{F(x)(1-F(x))}{\phi^2(\phi^{-1}(F(x)))}$$

Ved å laga tabell for σ_p^2 som ein funksjon av $q=F(x)$, ser vi at $p_n(x)$ er meir presis i den sentrale delen av F enn i halene. Legg merke til at $F_n(x)$ har dei motsette eigenskapane.

Frå appendiks 1 følgjer at

$$[F_n(x) - k\psi(F_n(x))/\sqrt{n}, F_n(x) + k\psi(F_n(x))/\sqrt{n}] = [h(F_n(x)), \bar{h}(F_n(x))] \text{ for alle } x \in [F_n^{-1}(a), F_n^{-1}(b)],$$

(a, b er konstanter og ψ ein eller annan vektfunksjon) er ein klasse simultane fordelingsfrie konfidensband for $F(x)$ med nivå $1-\alpha$. Dette gir at

$$(3) \quad [\phi^{-1}(h(F_n(x))), \phi^{-1}(\bar{h}(F_n(x)))] \text{ for alle } x \in [F_n^{-1}(a), F_n^{-1}(b)]$$

vert ein klasse simultane fordelingsfrie konfidensband for $p(x)=\phi^{-1}(F(x))$. Desse banda kan nyttast til å testa nullhypotesen

$$H_0: F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) \text{ mot } A: F(x) \neq \Phi\left(\frac{x-\mu}{\sigma}\right).$$

Ein test basert på (3) forkaster H_0 dersom inga rett linje passer i $1-\alpha$ nivå bandet (3) for gitt ψ . Testen har nivå som er mindre eller lik α og seiest derfor å vere konservativ. Testprosedyren ovanfor framkjem på grunnlag av følgjande samanheng:

$$P(\phi^{-1}(h(F_n(x))) \leq cx + d \leq \phi^{-1}(\bar{h}(F_n(x))), \forall x \text{ og ein eller annan } c, d)$$

$$> P(\phi^{-1}(h(F_n(x))) \leq \frac{x-\mu}{\sigma} \leq \phi^{-1}(\bar{h}(F_n(x))), \forall x) = 1-\alpha \text{ sidan}$$

$$p(x) = \phi^{-1}(F(x)) = \frac{x-\mu}{\sigma} \text{ når } F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

Lar vi $\psi=1$ ($a=0, b=1$) får vi frå (3) konfidensbandet

$$(4) \quad \left[\phi^{-1}\left(F_n(x) - \frac{k}{\sqrt{n}}\right), \phi^{-1}\left(F_n(x) + \frac{k}{\sqrt{n}}\right) \right].$$

Bandet er basert på Kolmogorov-Smirnov observatoren. Vi får forkastning av H_0 dersom inga rett linje passer i bandet. Ved å basera oss på konfidensbandet

$$(5) \quad \left[\phi^{-1}\left(F_n(x) - \frac{k'}{\sqrt{n}}\right), \phi^{-1}\left(F_n(x) + \frac{k'}{\sqrt{n}}\right) \right],$$

vil vi oppnå ein bedre testprosedyre. Vi skal nå visa dette:

Vi definerer $\eta(x) = \Phi^{-1}\left(F\left(\frac{x - \bar{x}}{s} + \mu\right)\right)$ der

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \text{ og } x_1, \dots, x_n$$

er det observerte utvalet. Under

$$H_0: F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right) \text{ blir}$$

$$\eta(x) = \frac{x - \bar{x}}{s}.$$

Vi ser derfor på observatoren

$$D'_n = \sqrt{n} \sup_x |F_n(x) - \Phi\left(\frac{x - \bar{x}}{s}\right)|.$$

Den kritiske verdien k' i (5) blir for gitt α fastsett frå likningen $P(D'_n \leq k') = 1 - \alpha$. Testen forkaster H_0 (hypotesen at observasjonane stammer frå ein normalfordeling med ukjend forventning og varians) hvis linja $\frac{x - \bar{x}}{s}$ ikkje passer i bandet (5). Dette er ekvivalent med å forkaste H_0 dersom $D'_n \geq k'$, sidan

$$P\left(\Phi^{-1}\left(F_n(x) - \frac{k'}{\sqrt{n}}\right) \leq \eta(x) \leq \Phi^{-1}\left(F_n(x) + \frac{k'}{\sqrt{n}}\right)\right) = P(D'_n \leq k') = 1 - \alpha.$$

Durbin (1973) har vist at k' er monaleg mindre enn k . Dermed vert konfidensbandet (5) smalare enn (4), og vi oppnår ein meir følsom test ved å basera oss på (5) enn på (4).

Referanser

- Andrews, F.C. (1954): Asymptotic behavior of some rank tests for analysis of variance. Ann. Math.Stat. 25, 724-736.
- Borokov, A.A. & Sycheva, N.M. (1968): On asymptotically optimal non-parametric criteria. Th. Prob. and Appl. 13, 359-393.
- Canner, P.L. (1975): A simulation study of one- and two-sample Kolmogorov-Smirnov statistics with a particular weight function JASA 70, 209-211.
- Doksum, K.A. (1974): Empirical probability plots and statistical inference for nonlinear models in the two sample case. Ann. of Statist. 2, 267-277.
- Doksum, K.A. & Sievers, G.L.: Plotting with confidence. Graphical comparisons of two populations. Biometrika, to appear.
- Durbin, S. (1973): Distribution theory for tests based on the sample distribution function. Regional Conference Series in Applied Mathematics SIAM, Philadelphia.
- Hodges, J.L. & Lehmann, E.L. (1956): The efficiency of some nonparametric competitors of the t-test. Ann. Math. Statist. 27, 324-335.
- Kruskal, W.H. & Wallis, W.A. (1952): Use of ranks in one-criterion variance analysis. JASA 47, 583-612.
- Lehmann, E.L. (1974): Statistical Methods Based on Ranks, Holden-Day, Inc., San Francisco.
- Noether, G.E. (1963): "Note on the Kolmogorof statistics in the discrete case". Metrika 7, 115-116.
- Owen, D.B. (1962): Handbook of Statistical Tables, Addison-Wesley, Inc., Massachusetts.
- Pitman, E.J.G. (1948): Lecture notes on nonparametric statistics. Columbia University, New York.
- Rényi, A. (1953): On the theory of order statistics. Acta. Math. Acad. Sci. Hungar. 4, 191-231.
- Sverdrup, E. (1964): Lov og tilfeldighet I og II, Universitetsforlaget.
- Wilcoxon, F. (1945): Individual comparisons by ranking methods, Biometrics 1, 80-83.