STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

Rapport S 340

# Enige problemen in de verdelingsvrije Statistiek

door Prof.dr. C.B. Bell

2^e druk

februari 1965

Printed at the Mathematical Centre at Amsterdam, 49, 2nd Boerhaavestraat.
The Netherlands.

The Mathematical Centre, founded the 11th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for Pure Scientific Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

Enige problemen in de verdelingsvrije statistiek

door

C.B. BELL [1)2)]

(Mathematisch Centrum, Amsterdam and San Diego State College)

(Structuur en bouw van verdelingsvrije toetsen voor

a) de 2-steekproeven hypothese $H_o : F_1 = F_2$; en

b) de onafhankelijkheidshypothese $H_o : F(x,y) = G(x)\ J(y)$ voor alle $x,y$ in $R_1$.)

(A) Noot: Volgens KENDALL & STUART (Vol. II p. 487) is de 2-steekproeven hypothese een hypothese die zegt dat de gecombineerde steekproefwaarde $(X_1,\ldots,X_m; Y_1,\ldots Y_n)$ en de steekproefaantallen (1,2) onafhankelijk zijn. Men kan dit zien als men de relatie bestudeert tussen de toets voor onafhankelijkheid in een 2×2-tabel (dubbele dichotomie) en de mediaan toets voor twee steekproeven. Daarom zullen wij de twee hypothesen tezamen bestuderen.

(B) Notatie: Evenals SCHEFFE (1943), BIRNBAUM en RUBIN (1954), en BELL (1964a) voeren wij de volgende gebruikelijke schrijfwijze in.

$R_n$, de n-dimensionale Euclidische ruimte.

$\mathcal{B}_n$, de klasse van Borelverzamelingen op $R_n$.

$\Omega$, een klasse van verdelingsfuncties op een willekeurige Euclidische ruimte $R_k$.

$\mathfrak{N}(\Omega) = \{A \in \mathcal{B}_k : P_F(A) = 0$ voor alle $F \in \Omega\}$, de nulklasse van $\Omega$.

$\Omega(n) = \{F^{(n)} : F \in \Omega\}$, waar $F^{(n)}$ de verdelingsfunctie (verdfct) is van een aselecte steekproef van n elementen uit een populatie met verdfct F.

1) Current address: School of Statistics, University of Madrid.

2) Work on this paper was primarily supported by the National Science Foundation through grant NSFG-25220.

$\Omega_2$, de klasse van continue verdfcts op $R_1$.

$\Omega^{*}_2$, de klasse van(strikt) monotone, continue verdfcts op $R_1$.

$\Omega_H$, de klasse van normale verdfcts op $R_1$.

$\varepsilon$, de verdfct met $\varepsilon(x) = \begin{cases} 1 \text{ als } x \geq 0 \\ 0 \text{ als } x < 0. \end{cases}$

$(1-p)F$ en $p \sum p_i \ \varepsilon(.-a_i)$, de continue en discrete delen van een verdfct $Q$ op $R_1$. $(0 \leq p, p_i \leq 1, \sum p_i = 1, a_i$ reële getallen).

In het 2-steekproeven probleem heeft men

$(X_1,\ldots,X_m)$, de X-steekproef;
$(Y_1,\ldots,Y_n)$, de Y-steekproef;
$(Z_1,\ldots,Z_N) = (X_1,\ldots,X_m; Y_1,\ldots,Y_n)$, de gecombineerde steekproef, waarbij $N = n+m$;
$R(X_i)$ = het rangnummer van $X_i$ in de gecombineerde steekproef; (na ordening naar grootte).
$Z(1),\ldots,Z(N)$, de "order statistics" van de gecombineerde steekproef;
$W_1,\ldots,W_N$, een aselecte steekproef onafhankelijk van de gecombineerde steekproef en met willekeurige vaste verdfct J.

Voor het onafhankelijkheidsprobleem heeft men
$(X_1,Y_1), (X_2,Y_2),\ldots,(X_n,Y_n)$, de 2-dimensionale steekproef die men in de volgende vorm zal schrijven zodat men de overeenkomst tussen het tweesteekproevenprobleem en het onafhankelijkheidsprobleem kan zien:
$X_1,\ldots,X_n; Y_1,\ldots,Y_n$.
In dit geval heeft men nodig
$R_X(X_i)$, het rangnummer van $X_i$ in de geordende X-steekproef;
$R_Y(Y_j)$, het rangnummer van $Y_j$ in de geordende Y-steekproef; en
$W_1,\ldots,W_n$ en $V_1,\ldots,V_n$, twee onafhankelijke aselecte steekproeven die onafhankelijk zijn van de 2-dimensionale steekproef en die de willekeurige vaste verdfcts J respectievelijk K hebben.
Ook heeft men nodig
$S(k) = \{\Pi\}$, de symmetrische groep van de $k!$ permutaties van getallen $[1,2,\ldots,k]$, met $\Pi[1,2,\ldots,k] = [\Pi(1),\ldots,\Pi(k)]$.

(C) Definitie: In de tijdschriften kan men de woorden "parametervrij" en "verdelingsvrij" door elkaar gebruikt zien. Daarom zullen wij de volgende precieze definities van deze essentiële begrippen van BIRNBAUM en RUBIN (1954) en BELL (1964a) overnemen.

Def. 1(a). Een functie $T(\cdot,\cdot)$ van $\Omega \times R_n$ heet een st.gr.(stochastische grootheid) mbt (met betrekking tot) $\Omega$, een klasse van verdfcts op $R_n$, als $T_F(\cdot) = T(F,\cdot)$ meetbaar is mbt $\mathcal{B}_n$ voor alle F in $\Omega$.

(b) $T_F$ is de marginale st.gr. van T.

[Opm.: De T van Definitie 1 is niet een gewone st.gr. De marginale st. grn. $T_F$ zijn gewone st.grn. en $T = \{T_F : F \in \Omega\}$ is een familie van gewone st.grn.]

Def. 2. Een st.gr. T (mbt $\Omega$) heet DF (verdelingsvrij) mbt $\Omega$, als er één enkele verdfct $Q = Q_T$ bestaat met de eigenschap:

$P\{T(F;Z_1,\ldots,Z_n) \leq t|F^{(n)}\} = Q_T(t)$ voor alle F in $\Omega$ en t in $R_1$.

Voorbeelden

(1) $\Omega = \Omega_2(n)$; $T_1(G; X_1,\ldots,X_n) = \sup_x|F_n(x) - G(x)|$ waar $F_n$ de empirische verdfct $\sum_i \frac{1}{n}\,\varepsilon(\cdot-x_i)$ is.

(2) $\Omega = \Omega_H$; $T_2(G;x) = \dfrac{x-\mu_G}{\sigma_G}$, waar $\mu_G$ en $\sigma_G$ respectievelijk de verwachting en de standaardafwijking zijn van een normale verdfct G.

(3) $\Omega = \Omega_H(n) \times \Omega_H(n)$; $T_3(H^{(n)} \times J^{(n)}; X_1,\ldots,X_n; Y_1,\ldots,Y_n) = r$, waar r de correlatiecoëfficient is.

(4) $\Omega = \Omega_2$; $T_4(G;X) = X$.

De st.grn. $T_1, T_2$ en $T_3$ zijn DF mbt de gegeven families $\Omega$, terwijl $T_3$ en $T_4$ parametervrij zijn, d.w.z. zij bezitten geen parameter.

Daarom is het duidelijk, dat de parametervrije eigenschap verschilt van de DF eigenschap, en dat men een precieze definitie van het begrip parametervrij nodig heeft.

Def. 3. T, een st.gr. mbt $\Omega$, heet NP (parametervrij) mbt $\Omega$ als $T_F = T_G$ voor alle F en G in $\Omega$.

In toepassingen en in de literatuur ziet men dat in het algemeen de st.grn. voor de 2-steekproeven- en de onafhankelijkheidshypothesen NPDF zijn, d.w.z. zij bezitten beide eigenschappen NP en DF.

Voorbeelden

In verband met de onderstaande tabel herinnere men zich dat $T_5$,$T_6$,$T_7$ en $T_8$ de toetsingsgrootheden van SPEARMAN, MANN-WHITNEY-WILCOXON, PITMAN respectievelijk PITMAN zijn. BELL en DOKSUM (1964a,b) bestudeerden $T_{10}$, een bijzonder geval van $T_9$, en bijzondere gevallen van $T_{10}(h)$. Ook is het duidelijk dat $T_5(h)$, $T_6(h)$, $T_7(h)$, $T_8(h)$ en $T_{10}(h)$ de algemeenste st.grn. zijn van de bijbehorende klassen en hypothesen.

## Tabel 1 NPDF Toetsingsgrootheden

| Hypothese | Onafhankelijkheid | 2-steekproeven |
|---|---|---|
| $\Omega$ | $\Omega_2^{(n)} \times \Omega_2^{(n)}$ | $\Omega_2(N)$ |
| Gewone Rang Toetsings-grootheid | $T_5 = \sum_1^n R_X(X_i)R_Y(Y_i)$ of<br>$T_5' = \sum_1^n [R_X(X_i) - R_Y(Y_i)]^2$ | $T_6 = \sum R(X_i)$ of<br>$T_6' = [m^{-1}]\sum_i R(X_i) - [n^{-1}]\sum_j R(Y_j)$ |
| PITMAN Permutatie Toetsings-grootheid | $T_7 = \sum \varepsilon \left[\sum_1^n X_i Y_i - \sum_1^n X_i Y_{\Pi(i)}\right]$, waar de eerste sommatie over $\Pi$ in $S(n)$ is. | $T_8 = \sum \varepsilon \left[\underset{N}{\bar{X} - \bar{Y}}(m^{-1})\sum_1^m Z_{\Pi(i)} + (n^{-1})\sum_{m+1} Z_{\Pi(j)}\right]$ waar de eerste sommatie over $\Pi$ in $S(\mu)$ is. |
| Gerandomizeerde DOKSUMtype Toetsingsgrootheid | $T_9 = \sum_1^n W(R_X(X_i))V(R_Y(Y_i))$ | $T_{10} = (m^{-1}) \sum_i W(R(X_i))$<br>$-(n^{-1}) \sum_j W(R(Y_j))$ |
| GENERALISATIE: Rang | $T_5(h) = h[W(R_X(X_1)), \ldots, W(R_X(X_n));$<br>$V(R_Y(Y_1)), \ldots, V(R_Y(Y_n))]$ | $T_6(h) = h[R(X_1), \ldots, R(X_m);$<br>$R(Y_1), \ldots, R(Y_n)]$ |
| GENERALISATIE: PITMAN-type | $T_7(h) = \sum \varepsilon [h(X_1, \ldots, X_n; Y_1, \ldots, Y_n) - h(X_1, \ldots, X_n; Y_{\Pi(1)}, \ldots, Y_{\Pi(n)})]$<br>waar de eerste sommatie over $\Pi$ in $S(n)$ is. | $T_8(h) = \sum_\Pi \varepsilon [h(Z_1, \ldots, Z_N) -$<br>$-h(Z_{\Pi(1)}, \ldots, Z_{\Pi(N)})]$ |
| GENERALISATIE: DOKSUM-type | $T_9(h) = h[W(R_X(X_1)), \ldots, W(F_X(X_n)); V(R_Y(Y_1)), \ldots, V(R_Y(Y_n))]$ | $T_{10}(h) = h[W(R(X_1)), \ldots, W(R(X_m));$<br>$W(R(Y_1)), \ldots, W(R(Y_n))]$ |

Voordat enige problemen gesteld worden, volgt hieronder nog een laatste definitie.

**Def. 4.** Als $\Omega = \Omega^{**}_2(m) \times \Omega^{**}_2(n)$ en T is een st.gr. NPDF mbt $\Omega$, dan heet T SDF (sterk verdelingsvrij) mbt $\Omega$ wanneer

$$P\{T(X_1,\ldots,X_m; Y_1,\ldots,Y_n) \leq t|F^{(m)} \times G^{(n)}\}$$

$$= P\{T(X_1,\ldots,X_m; Y_1,\ldots,Y_n) \leq t|H^{(m)} \times J^{(n)}\}$$

voor alle t in $R_1$, en voor alle F,G,H,J in $\Omega^{**}_2$ waarvoor geldt: $FG^{-1} = HJ^{-1}$.

De SDF eigenschap is de natuurlijke generalisatie van een belangrijke eigenschap van normale verdfcts. Als F en G elementen zijn van $\Omega_H$ volgt hieruit $FG^{-1}(t) = \Phi|\alpha\Phi^{-1}(t)+\beta|$, waar $\Phi(x)$ de standaardnormale verdfct is en de parameters $\alpha = \sigma_F^{-1}\ \sigma_G$ en $\beta = \sigma_F^{-1}|\mu_G-\mu_F|$ ook de parameters zijn van het onderscheidingsvermogen van de $\bar{X}$-toets en de toets van STUDENT. Gemakkelijk kan men bewijzen (BELL, 1964b) dat de gewone 2-steekproevenrang toetsingsgrootheden SDF zijn mbt $\Omega^{**}_2(m) \times \Omega^{**}_2(n)$. Daarom ziet men dat mbt $\Omega^{**}_2(m) \times \Omega^{**}_2(n)$ en $T_6(h)$ SDF zijn; en dat voor vaste $W_1 = W_1,\ldots,W_N = w_N, T_{10}$ en $T_{10}(h)$ SDF zijn.

Men kan nu enige problemen stellen.

**(D) Problemen** Gewoonlijk bestudeert men slechts niet-gerandomizeerde rang st.grn. voor de onafhankelijkheidhypothese en 2-steekproevenhypothese, d.w.z. st.grn. van de vormen $T_5(h)$ en $T_6(h)$. Men zoekt dan, b.v. de st.gr. van de meest onderscheidende rangtoets of de st.gr. van de lokaal meest onderscheidende rangtoets voor de gegeven hypothese en de gegeven klasse van alternatieven.

Het is bekend, dat de klasse van st.grn. van niet-gerandomizeerde rangtoetsen de st.grn. $T_7(h)$ en $T_8(h)$ van de permutatietoets niet insluit, en ook niet de st.grn. $T_9(h)$ en $T_{10}(h)$ van de gerandomizeerde toetsen. Daarom is het niet zonder meer duidelijk dat de meest onderscheidende (of lokaal meest onderscheidende) niet-gerandomizzerde toets de meest onderscheidende (of lokaal meest onderscheidende) DF toets is.

Wegens de vorige besprekingen kan men nu de volgende vragen stellen.

(I) Zijn de NP st.grn. $T_5(h),\dots,T_{10}(h)$ DF,

(II) In welke zin zijn de gegeven st.grn. op verschillende manieren verdelingsvrij? Heeft men een andere definitie nodig?

(III) Welke andere voorwaarde(n) moet een DF st.gr. vervullen om een rang st.gr. te zijn?

(IV) Is het noodzakelijk dat een gerandomizeerde toets of een permutatietoets minder onderscheidingsvermogen heeft dan de overeenkomstige niet-gerandomizeerde rangtoets?

(V) Wat zijn natuurlijke klassen van alternatieven voor de onafhankelijkheidshypothese?

(VI) Bestaat er een natuurlijke definitie van SDF voor de onafhankelijkheidshypothese?

(VII) Hoe zou men de functie h moeten kiezen, opdat $T_7(h)$, $T_8(h)$, $T_9(h)$ en $T_{10}(h)$ optimaal zijn in enige zin?

## (E) Enige gedeeltelijke antwoorden

(I) Voor de eerste vraag kan men de volgende stelling van BELL (1964a) gebruiken.

Stelling. De volgende twee voorwaarden zijn equivalent:

(a) Er bestaat een st.gr. T NPDF mbt $\Omega$ en met verdfct

$$Q_T = (1-p)F+p \sum p_i \ \varepsilon(\cdot-a_i);$$

(b) er bestaat een deling $\{A_0,A_1,\dots\}$ van $R_k$ en een $\sigma$-algebra $\mathcal{B} \subset \mathcal{B}_k$ zodat

(i) $P\{A_i|F\} = pp_i$ voor $i \geq 1$ en alle F in $\Omega$; en

(ii) $A_o \cap \mathcal{B}$ is een niet-atomaire $\sigma$-ring van verzamelingen $\{D\}$ "similar" mbt $\Omega$, d.w.z. $P(D|F) = P(D|G)$ voor alle F en G in $\Omega$.

Gebruikmakend van de vorige stelling ziet men gemakkelijk dat

(1) $T_6(h)$ constant is op iedere verzameling $A(\Pi) = \{\mathfrak{Z}_{\Pi(1)}<\ldots<\mathfrak{Z}_{\Pi(N)}\}$ waar $\Pi$ een element is van $S(N)$; iedere $A(\Pi)$ is "similar" mbt $\Omega_2(n)$; en dat daarom $T_6$ en elke $T_6(h)$ NPDF zijn mbt $\Omega_2(N)$; en dat

(2) $T_5(h)$ constant is op iedere verzameling $A(\Pi,\gamma) = \{x_{\Pi(1)}<\ldots<x_{\Pi(n)};\ y_{\gamma(1)}\ldots y_{\gamma(n)}\}$, waar $\Pi$ en $\gamma$ elementen zijn van $S(n)$; iedere $A(\Pi,\gamma)$ is "similar" mbt $\Omega_2(n) \times \Omega_2(n)$; en dat daarom $T_5$ en elke $T_5(h)$ NPDF zijn mbt $\Omega_2(n) \times \Omega_2(n)$.

Op analoge wijze vindt men dat

(3) voor iedere vaste $\{W_1 = w_1,\ldots,W_N = w_N\}$ elke $T_{10}(h)$ NPDF is mbt $\Omega_2(N)$; en dat

(4) voor iedere vaste $\{W_1 = w_1,\ldots,W_n = w_n;\ V_1 = v_1,\ldots,V_n = v_n\}$ elke $T_9(h)$ NPDF is mbt $\Omega_2(n) \times \Omega_2(n)$.

Maar, voor de PITMAN-achtige st.grn. heeft men moeilijkheden.

(5) In het bijzonder geval $T_7(h) = T_7$ en n=3, vraagt men zich af: Zijn alle volgende verzamelingen van "constantheid" "similar" als dit nodig is?

$\{T_7=6\} = A(1,2,3)$; $\{T_7=5\}=[A(1,3,2) \cap Q_g] \cup [A(2,1,3) \cap Q_g^o]$

$\{T_7=4\} = [A(1,3,2) \cap Q_g^o] \cup [A(2,1,3) \cap Q_g]$; ......,

waar $A(\Pi(1), \Pi(2), \Pi(3)) = \{ \sum_1^3 X_iY_i = \sum_1^3 X(i)Y(\Pi(i))\}$,

$Q_g^o=R_{2n}-Q_g$ en $Q_g = \{ \dfrac{X(2)-X(1)}{X(3)-X(2)} > \dfrac{Y(3)-Y(2)}{Y(2)-Y(1)} \}$.

[N.B.: Als n=2, is $T_7$ een rang st.gr.]

(6) Ook in het bijzondere geval $T_8(h) = T_8$ en m=2=n, heeft men

$\{T_8=1\}=\{X(2) < Y(1)\}$; $\{T_8=2\}=\{X(1) < Y(1) < X(2) < Y(2)\}$;

$\{T_8=3\}=\{I_X \cup I_Y\} \cap \{X(1)+X(2) < Y(1)+Y(2)\}$;..., waar $I_X=\{Y(1) < X(1) < X(2) < Y(2)\}$ en $I_Y=\{X(1)< Y(1) < Y(2) < X(2)\}$.

Zijn alle vorige verzamelingen "similar" mbt $\Omega_2(N)$?

Vermoedelijk zijn de antwoorden op de laatste twee vragen negatief en zijn daarom de PITMAN-achtige st.grn. niet DF in de zin van Def. 2.

(II) De st.grn. verschillen in de volgende opzichten:

(1) $T_5(h)$ en $T_6(h)$ zijn NPDF mbt hun respectievelijke klassen, en $T_6(h)$ is (BELL 1964a) SDF mbt $\Omega_2^*(m)\times \Omega_2^*(n)$.

(2) $T_7(h)$ en $T_8(h)$ zijn in het algemeen niet DF in de zin van Def. 2, maar zij zijn voorwaardelijk DF in één of andere zin, omdat voor elke vaste $\{X_i = x_i\}$ de voorwaardelijke waarschijnlijkheid van iedere Y-permutatie dezelfde is.

(3) De st.grn. $T_9(h)$ en $T_{10}(h)$ zijn geen st.grn. mbt de gegeven $\Omega$'s in de zin van Def. 1, maar zij zijn families van st.grn. De indexverzamelingen van deze families zijn $R_{2n}$ en $R_N$ met respectievelijk productmaat $P\{\cdot|J^{(n)}\times K^{(n)}\}$, en $P\{\cdot|J^{(N)}\}$. Elk element van iedere familie is NPDF.

Men moet nu misschien een definitie van voorwaardelijke DF-heid zoeken, die $T_7(h)$, $T_8(h)$, $T_9(h)$ en $T_{10}(h)$ omvat.

(III), (V) en (VI). BELL (1964a) bewijst dat, in het 2-steekproeven probleem, elke rang st.gr. SDF is mbt $\Omega^*{}_2(m) \times \Omega^*{}_2(n)$; en dat ls T SDF is en als de grens van elke $T^{-1}(A)$, met A in $\mathcal{B}_1$, een element is van $\mathcal{N}(\Omega_2(N))$ (zie SCHEFFÉ; 1943), dan is het zo dat T equivalent met een rang st.gr. is. Men vraagt zich nu af of het volgende vermoeden juist is:

Vermoeden: T is SDF dan en slechts dan als T equivalent is met een rang st.gr.

Noot; Als men een HAARmaat op de groep $\mathcal{L}^*$ van monotone continue functies van $R_1$ op $R_1$ bouwt, dan (LEHMANN, p.225--27) is het vorige vermoeden juist.

Voor het artikel (BELL, 1964a) was het de idee van de "referee", dat men de methoden van hetzelfde artikel kan gebruiken voor een bewijs van een gelijksoortige karakterisering van de klasse van st.grn. voor de onafhankelijkheidshypothese. Maar het schijnt dat de SDF eigenschap essentieel is voor de bewijzen van bedoeld artikel, en daarom heeft men een natuurlijke definitie van SDF nodig.

Omdat het SDF begrip een generalisatie is van een eigenschap van het onderscheidingsvermogen van de normale verdfcts, schijnt het dat men een definitie van SDF en de natuurlijke klassen van alternatieven onder de eigenschappen van de 2-dimensionale normale verdfcts moet zoeken.

(IV) Slechts weinig waarden van het onderscheidingsvermogen van de gewone DF toets zijn bekend. In navolging van CHAPMAN (1958), hebben BELL, MOSER en THOMPSON (1964) het maximale onderscheidingsvermogen voor enige 2-steekproeven st.grn. en eenzijdige alternatieven berekend - in enkele gevallen met Monte Carlo methoden. Maar het schijnt dat het antwoord op Vraag (IV) niet duidelijk is uit deze berekeningen. Het is de mening van enige statistici dat het gebruik van gerandomizeerde st.grn. meer spreiding introduceert en dat daarom de toets gebaseerd op deze st.grn. minder onderscheidingsvermogen heeft, evenals in het normale geval. Ook deze mening is echter onbewezen.
De conclusie is dat men het antwoord op Vraag (IV) niet weet en dat een verdergaande bestuderingen van de permutatie-en de gerandomiseerde toets nodig is.

(VII) Het is duidelijk, dat Vraag (VII) de belangrijkste is en ook de moeilijkste. Van het boek van LEHMANN (1961), p. 236-240, de artikelen van BELL en DOKSUM (1964a,b) en de geciteerde literatuur daarin, kan men het volgende zeggen:
(1) In het algemeen bouwt men een optimale (locale meest onderscheidende rang, meest onderscheidende permutatie, enz.) toets altijd door gebruik te maken van het NEYMAN-PEARSON-lemma.

(2) In het algemeen zijn de overeenkomstige optimale rangtoets en gerandomizeerde rangtoets asymptotisch equivalent.

Vermoeden: De overeenkomstige permutatietoets voldoet aan (2).

(3) Gerandomizeerde rangtoetsen hebben wenselijke limieteigenschappen, maar het schijnt, dat zij nooit de machtigste of locale machtigste toetsen voor eindige steekproefomvang zijn.

(4) In de literatuur bestudeert men gewoonlijk translatie en schaal alternatieven. In de wiskunde zijn de translatiegroep, de schaalgroep en de "affine"-groep, die de twee vorige insluit, drie van de vele mogelijke groepen. Misschien is hun belangrijkste eigenschap hun eenvoud. Ook in statistische toepassingen is het bijna nooit duidelijk, dat translatie of schaal alternatieven de juiste klassen van alternatieven zijn. Daarom schijnt het wenselijk dat men andere klassen van alternatieven of groepen van alternatieven bestudeert. BELL en DOKSUM (1964b) hebben een deel van deze problemen bestudeerd.

## Geciteerde literatuur

1. BELL, C.B. (1964a). Some basic theorems of distribution-free statistics. Ann. Math. Statist. 35, 150-156.

2. BELL, C.B. (1964b). A characterization of multisample distribution-free statistics. Ann. Math. Statist. 35, 735-738.

3. BELL, C.B. (1964c). An introduction to automatic non-parametric signal detection. To be published by U.S. Navy Electronics Laboratory, San Diego, Calif.

4. BELL, C.B. and DOKSUM, K.A. (1964a). Some new distribution statistics. To be published, Ann. Math. Statist. March, 1965.

5. BELL, C.B. and DOKSUM, K.A. (1964b). Optimal one-sample distribution-free statistics and their two-sample extensions. Ann. Math. Statist. March 1965.

6. BELL, C.B. MOSER, J.N. and THOMPSON, RORY. (1964). Goodness criteria for two-sample distribution-free statistics. Submitted for publication.

7. BIRNBAUM, Z.W. and RUBIN, H. (1964). On distribution-free statistics. Ann. Math. Statist. 25, 593-598.

8. BHUCHONGKUL, S. (1964). A class of nonparametric tests for independence in bivariate populations. Ann. Math. Statist. 35, 138-149.

9. CHAPMAN, D.G. (1958). A comparative study of several one-sided goodness-of-fit tests. Ann. Math. Statist. 29, 655-674.

10. KENDALL, M.G. and STUART, ALAN. (1961). The Advanced Theory of Statistics. Volume II: Inference and Relationship. Griffin, London.

11. KONIJN, H.S. (1956). On the power of certain tests for independence in bivariate populations. Ann. Math. Statist. 27, 300-323.

12. LEHMANN, E.L. (1959). Testing Statistical Hypotheses, Wiley, New York.

13. SCHEFFÉ, H. (1943). On a measure problem arising in the theory of nonparametric tests. Ann. Math. Statist. 14, 227-233.