GOODNESS CRITERIA FOR TWO-SAMPLE DISTRIBUTION-FREE TESTS[1]

By C. B. Bell, J. M. Moser, and Rory Thompson

San Diego State College

1. Introduction and summary. Some of the concepts and results of Chapman [3] for one-sample distribution-free tests are extended to the two-sample problem. Chapman's lead will be followed since his work gives a goodness criterion for comparing distribution-free tests, for finite sample sizes, over a large class of one-sided alternatives. Since all two-sample distribution-free statistics known to the authors satisfy Scheffé's [10] boundary condition, and all, except those of Pitman, also are strongly distribution-free (SDF) and therefore [1] are rank statistics, consideration may reasonably be restricted to rank tests. Such tests with the additional property of monotonicity are unbiased, partially ordered, and assume maximum and minimum powers for certain reasonable alternatives.

Some of the maximum powers of the Mann-Whitney-Wilcoxon, Fisher-Yates, van der Waerden, Doksum, Savage, Epstein-Rosenbaum, and Cramér-von Mises statistics are tabulated. (For definitions and references, see Section 6.)

2. Preliminaries. Let $X_1, \cdots, X_m$; $Y_1, \cdots, Y_n$ be independent random samples with population distribution functions (cpf's) F and G, respectively, with F an arbitrary element of Ω_2^*, the class of strictly monotone, continuous cpf's on R_1, and G such that either $H_0 : G = F$ or $H_1 : G < F$. [$G < F$ means $G(x) \leqq F(x)$ for all x, and that $G(x) < F(x)$ on some set of positive measure.] Let $W_1, \cdots, W_N$, where $N = n + m$, be the combined sample, and let $R(X_i)$ and $R(Y_j)$ be, respectively, the rank in the combined sample of X_i and Y_j.

A *test* of size α is a function of W such that $0 \leqq \phi \leqq 1$, and $E\{\phi \mid F, G\} \leqq \alpha$ whenever $F = G$. $\beta(\phi; F, G)$ will denote the power function of ϕ. A test ϕ is said to be SDF if $\beta(\phi; F_1, G_1) = \beta(\phi; F_2, G_2)$ whenever $G_1F_1^{-1} = G_2F_2^{-1}$; i.e., $\beta(\phi; F, G)$ for given ϕ depends only on GF^{-1}. A test ϕ is here said to be a *rank* test if $\phi(X_1, \cdots, Y_n)$ is a function only of the set of ranks $\{R(X_k)\}$. A test ϕ is said to be a *monotone* test if $\phi(X_1, \cdots, Y_n) \geqq \phi(X_1', \cdots, Y_n')$ whenever $X_i' \leqq X_i$ and $Y_j' \geqq Y_j$ for all integer $1 \leqq i \leqq m$, $1 \leqq j \leqq n$, and is said to be *partially ordered* (p.o.) if, whenever $G_2 \geqq G_1 \geqq F_1 \geqq F_2$, then $\beta(\phi; F_1, G_1) \leqq \beta(\phi, F_2, G_2)$.

Let U be the standard uniform distribution (cpf): $U(x) = x$ if $0 \leqq x \leqq 1$. Let $\bar{H}(X, \Delta) = U(X - \Delta)$ for all X.

3. Unbiasedness

Lemma. *Each two-sample rank test is* SDF.

Received 24 May 1964; revised 9 July 1965.

[1] This work was done under National Science Foundation Grant NSF-G-25220.

PROOF. [1]. Essentially, the ranking of W is not changed by a strictly increasing, continuous map F^{-1}, so $\beta(\phi, F, G) = \beta(\phi, FF^{-1}, GF^{-1})$.

THEOREM. *Each monotone two-sample rank test is* p.o.; *and each* p.o. *test is unbiased (for H_0 versus H_1).*

PROOF. SDF implies $\beta(\phi; F, G) = \beta(\phi, U, GF^{-1}) \geqq \beta(\phi, U, U) = \alpha$, since $GF^{-1} \leqq U$ when $G \leqq F$. This theorem is essentially that of Lehmann ([5], p. 187).

4. Power and power bounds. The pseudometric

$$\rho^{-}(F, G) = \sup_x(F(x) - G(x))$$

and the classes $C(F, \Delta) = \{G \mid G < F \text{ and } \rho^{-}(F, G) \leqq \Delta\}$ and $D(F, \Delta) = \{G \mid G < F \text{ and } \rho^{-}(F, G) \geqq \Delta\}$ yield power bounds for monotone, p.o., rank tests. Define $\bar{\beta}(\phi; F; \Delta) = \sup \beta(\phi; F, G)$ with the supremum over $C(F, \Delta)$ and $\underline{\beta}(\phi; F; \Delta) = \inf \beta(\phi; F, G)$ where the infimum is taken over $D(F, \Delta)$. The problem may be reduced to cpf's on the unit interval [0, 1] by the following lemma.

LEMMA. *For a two-sample rank test ϕ and arbitrary F in Ω_{2*}, $\bar{\beta}(\phi; F; \Delta) = \bar{\beta}(\phi; U; \Delta)$, and $\underline{\beta}(\phi; F; \Delta) = \underline{\beta}(\phi; U; \Delta)$, where the class of alternatives to U is restricted to* [0, 1].

PROOF. $\beta(\phi; F, G) = \beta(\phi; U, GF^{-1})$ since ϕ is SDF; and $\sup_x (U - GF^{-1}) = \sup_x (F - G)$ and $\inf_x (U - GF^{-1}) = \inf_x (F - G)$, since F^{-1} is continuous and strictly increasing; GF^{-1} is (any) cpf on [0, 1].

The problem now is to derive a method for computing the maximum and minimum powers. The first step in this direction is to define the Birnbaum alternatives on the unit interval. Let

$$\begin{aligned}
\bar{G}(x; \Delta) &= 0 && \text{if } 0 \leqq x \leqq \Delta \\
&= x - \Delta && \text{if } \Delta < x < 1 \\
&= 1 && \text{if } x \geqq 1, \\
\underline{G}(x; u, \Delta) &= 0 && \text{if } x \leqq 0 \\
&= x && \text{if } 0 < x \leqq u \\
&= u && \text{if } u < x < u + \Delta \\
&= x && \text{if } u + \Delta \leqq x < 1 \\
&= 1 && \text{if } x \geqq 1,
\end{aligned}$$

where $0 \leqq \Delta \leqq 1, 0 \leqq u \leqq 1 - \Delta$.

The next theorem asserts that maximum and minimum powers are attained against intuitively appealing alternatives.

THEOREM. *For each F in Ω_2^*, $\Delta \geqq 0$, and monotone, two-sample test ϕ,*

$$\text{(i)} \qquad \bar{\beta}(\phi; F, \Delta) = \beta(\phi; U, \bar{G}(\cdot, \Delta))$$

$$\text{(ii)} \qquad \underline{\beta}(\phi; F, \Delta) = \inf \beta(\phi; U, \underline{G}(\cdot, u, \Delta)),$$

where the infimum is over u in $[0, 1 - \Delta]$.

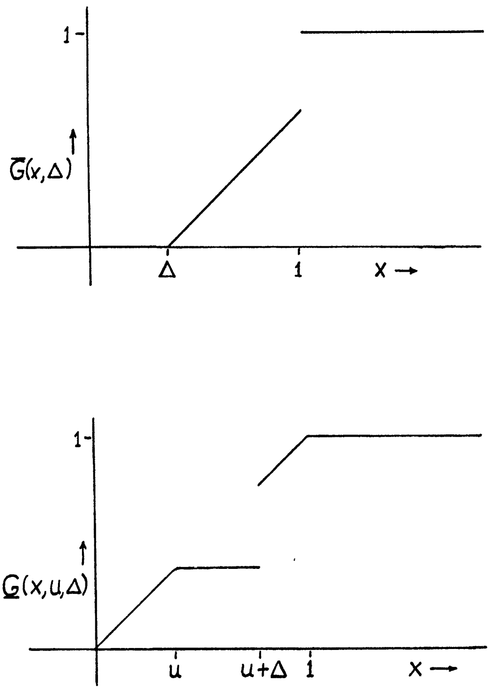

FIG. 1. Sample sketches of maximum and minimum alternatives

PROOF. (i) Let G be a cpf on [0, 1] such that $U(x) - G(x) \leqq \sup_x (U(x) - G(x)) = \Delta' \leqq \Delta$, so $G(x) \geqq x - \Delta' \geqq x - \Delta = \bar{G}(x, \Delta)$ for $\Delta \leqq x < 1$; $G(x) \geqq 0 = \bar{G}(x, \Delta)$ for $x \leqq \Delta$; $G(x) = 1 = \bar{G}(x, \Delta)$ for $x \geqq 1$. Therefore, $G \geqq \bar{G}(\cdot, \Delta)$, so, since ϕ is p.o., $\inf \beta(\phi; U, G) \geqq \beta(\phi, U, \bar{G}) \geqq \inf \beta(\phi, U, G)$ since $\bar{G}$ is in $C(U, \Delta)$.

The proof of (ii) is similar. By the preceding lemma it need only be shown that $\underline{\beta}(\phi; U, \Delta) = \inf_u \beta(\phi; U, \underline{G}(\cdot, u, \Delta))$. Observe that $\underline{G}(\cdot, u, \Delta)$ is in $D(U, \Delta)$, so $\inf_u \beta(\phi; U, \underline{G}(\cdot, u, \Delta)) \geqq \underline{\beta}(\phi; U, \Delta)$.

Now consider any G in $D(U, \Delta)$, so $\sup (U - G) = \Delta' \geqq \Delta$. If G is continuous, $U - G$ assumes its maximum at some point, which may be used as $u + \Delta$ below; in any case, there must exist a sequence $\{x_k\}$ such that $\{U(x_k) - G(x_k)\}$ in-areases to Δ'; so there are an infinite number of distinct x_k's in [0, 1], so there is an cccumulation point. Let $u + \Delta$ denote the supremum ($\leqq 1$) of such accumulation points. Then $u + \Delta$ is also such an accumulation point, so there exists a sequence $\{x_k\}$ increasing or decreasing to $u + \Delta$ such that $\{U(x_k) - G(x_k)\}$ increases to Δ'. If $\{x_k\}$ is increasing, then, for every $\epsilon > 0$, there is an $N > 0$ such that $k \geqq N$ implies $x_k - G(x_k) \geqq \Delta' - \epsilon$, so $G(x_k) \leqq x_k - \Delta' + \epsilon \leqq u + \Delta - \Delta' + \epsilon$, so, since G is increasing, $\sup_{x<u+\Delta} G(x) = \sup_k G(x_k) \leqq u + \Delta - \Delta' + \epsilon$, so $\sup_{x<u+\Delta} G(x) \leqq u + \Delta - \Delta'$. On the other hand, if $\{x_k\}$ decreases to $u + \Delta$ such that $\{U(x_k) - G(x_k)\}$ increases to Δ', then $\{u + \Delta - G(x_k)\}$ increases to Δ', so $\inf_{x>u+\Delta} G(x) = u + \Delta - \Delta' \geqq \sup_{x<u+\Delta} G(x)$.

Fisher-Yates

N	Δ									
	0.00	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.75	0.90
$n = 4$ $m = 5$	.0100	.0155	.0235	.0506	.0989	.1763	.2886	.4362	.7013	.9370
9	.0500	.0772	.1147	.2226	.3700	.5405	.7081	.8468	.9678	.9989
$n = 5$ $m = 5$	.0100	.0163	.0259	.0599	.1234	.2259	.3707	.5489	.8201	.9816
10	.0500	.0801	.1215	.2416	.4057	.5899	.7608	.8892	.9831	.9997
$n = 5$ $m = 6$	.0100	.0171	.0284	.0702	.1496	.2769	.4497	.6460	.8957	.9950
11	.0500	.0838	.1315	.2703	.4543	.6495	.8156	.9262	.9921	.9999
$n = 6$ $m = 6$	.0100	.0179	.0311	.0822	.1815	.3377	.5369	.7396	.9462	.9989
12	.0500	.0856	.137	.288	.485	.686	.846	.944	.9950	1.000
$n = 6$ $m = 7$	.0100	.0188	.0336	.0921	.2046	.3764	.5845	.7812	.9604	.9993
13	.050	.089	.146	.311	.522	.726	.878	.961	.997	1.000
$n = 7$ $m = 7$	.0100	.0196	.0364	.1045	.2374	.4343	.6564	.8431	.9803	.9998
14	.050	.091	.152	.332	.555	.761	.902	.972	.999	1.000
$n = 7$ $m = 8$	.0100	.0204	.039	.116	.265	.478	.705	.879	.9886	.9999
15	.050	.094	.159	.351	.584	.790	.921	.980	.999	1.000
$n = 8$ $m = 8$	.010	.021	.042	.129	.295	.523	.749	.907	.993	1.000
16	.050	.097	.167	.374	.617	.819	.938	.986	.999	1.000
$n = 8$ $m = 9$	.0100	.022	.045	.141	.322	.562	.785	.928	.996	1.000
17	.050	.099	.173	.390	.639	.838	.948	.989	1.000	1.000
$n = 10$ $m = 10$	7	18	44	154	367	631	845	961	1000	1000
20	52	109	211	451	720	913	979	997	1000	1000

van der Waerden

N	0.00	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.75	0.90
$n = 4$ $m = 5$	.0100	.0155	.0235	.0506	.0989	.1763	.2886	.4362	.7013	.9370
9	.0500	.0770	.1139	.2196	.3646	.5333	.7009	.8412	.9659	.9988
$n = 5$ $m = 5$	.0100	.0163	.0259	.0599	.1234	.2259	.3707	.5489	.8201	.9816
10	.0500	.0801	.1215	.2416	.4057	.5899	.7608	.8892	.9831	.9997
$n = 5$ $m = 6$	.0100	.0171	.0284	.0702	.1496	.2769	.4497	.6460	.8957	.9950
11	.0500	.0837	.1313	.2698	.4538	.6489	.8152	.9260	.9921	.9999
$n = 6$ $m = 6$	.0100	.0179	.0311	.0822	.1815	.3377	.5369	.7396	.9462	.9989
12	.0500	.0856	.137	.288	.485	.686	.846	.944	.9950	1.000
$n = 6$ $m = 7$	.0100	.0188	.0336	.0921	.2046	.3764	.5845	.7812	.9604	.9993
13	.050	.088	.142	.303	.511	.716	.871	.958	.997	1.000

van der Waerden—*Continued*

N	Δ 0.00	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.75	0.90
$n = 7$ $m = 7$	.0100	.0195	.0361	.103	.234	.430	.653	.841	.980	.9998
14	.050	.091	.151	.329	.552	.760	.901	.972	.999	1.000
$n = 7$ $m = 8$	.0100	.0204	.039	.115	.262	.473	.700	.876	.988	.9999
15	.050	.093	.158	.347	.579	.785	.918	.979	.999	1.000
$n = 8$ $m = 8$	.010	.021	.042	.127	.292	.519	.746	.906	.993	1.000
16	.050	.096	.165	.369	.610	.813	.935	.985	.999	1.000
$n = 8$ $m = 9$	.010	.022	.044	.139	.320	.560	.784	.928	.996	1.000
17										
$n = 10$ $m = 10$	9	23	52	172	397	661	867	966	1000	1000
20	53	107	201	445	711	899	977	997	1000	1000

Mann-Whitney-Wilcoxon

N	0.00	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.75	0.90
$m = 4$ $m = 5$	.0100	.0155	.0235	.0506	.0989	.1763	.2886	.4362	.7013	.9370
9	.0500	.0767	.1126	.2150	.3561	.5222	.6897	.8326	.9629	.9987
$m = 5$ $m = 5$	.0100	.0163	.0259	.0599	.1234	.2259	.3707	.5489	.8201	.9816
10	.0500	.0796	.1201	.2379	.3992	.5819	.7533	.8842	.9819	.9997
$m = 5$ $m = 6$	.0100	.0171	.0284	.0696	.1481	.2742	.4460	.6424	.8940	.9949
11	.0500	.0808	.1238	.2497	.4209	.6098	.7798	.9025	.9864	.9998
$m = 6$ $m = 6$	.0100	.0179	.0309	.0803	.1753	.3247	.5172	.7175	.9335	.9978
12	.0500	.0832	.1304	.2699	.4568	.6546	.8211	.9301	.9928	.9999
$m = 6$ $m = 7$	.0100	.0187	.0331	.0894	.1979	.3650	.5706	.7692	.9563	.9997
13	.050	.085	.136	.287	.487	.691	.853	.949	.996	1.000
$m = 7$ $m = 7$	.0100	.0193	.0352	.0989	.2226	.4090	.6269	.8201	.9746	.9997
14	.050	.087	.142	.305	.516	.724	.877	.961	.998	1.000
$m = 7$ $m = 8$	.0100	.0198	.037	.108	.245	.446	.670	.854	.983	.9999
15	.050	.090	.148	.324	.547	.756	.901	.972	.999	1.000
$m = 8$ $m = 8$	.010	.021	.039	.118	.270	.487	.715	.887	.990	1.000
16	.050	.091	.153	.338	.570	.780	.920	.979	.999	1.000
$m = 10$ $m = 10$	8	21	44	149	359	605	829	953	999	1000
20	47	93	164	385	639	851	958	992	1000	1000

Doksum-Bell

N	Δ 0.00	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.75	0.90
$n = m = 5$	12	16	22	51	105	170	232	323	401	454
10	54	98	131	238	360	506	693	751	815	885
$n = m = 6$	6	19	28	46	128	213	292	429	522	625
12	59	97	139	248	416	584	741	860	930	950
$n = m = 7$	10	21	31	89	155	293	425	552	726	792
14	64	103	136	256	474	660	798	908	975	980
$n = m = 10$	4	22	40	140	335	529	720	831	951	977
20	57	101	201	406	662	868	956	986	999	1000

Similar Savage

N	0.00	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.75	0.90
$n = 5$ $m = 5$	.0100	.0163	.0259	.0599	.1234	.2259	.3707	.5489	.8201	.9816
10	.0500	.0776	.1157	.2271	.3820	.5613	.7346	.8717	.9789	.9996
$n = 6$ $m = 6$	.0100	.0179	.0306	.0787	.1705	.3148	.5020	.7001	.9233	.9970
12	.0500	.0817	.1273	.264	.448	.645	.813	.925	.992	.9999
$n = 7$ $m = 7$	.0100	.0189	.0340	.0948	.2135	.3946	.6104	.8073	.9715	.9997
14	.050	.087	.141	.305	.517	.725	.878	.962	.998	1.000
$n = 8$ $m = 8$	.010	.020	.038	.115	.264	.480	.709	.883	.989	1.0000
16	.050	.092	.155	.345	.581	.789	.922	.981	.999	1.000

Cramér - von Mises

N	0.00	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.75	0.90
$n = 4$ $m = 5$	.0100	.0109	.0137	.0262	.0510	.0928	.1562	.2437	.4129	.5802
9	.0500	.0544	.0677	.1214	.2134	.3431	.5019	.6712	.8868	.9903
$n = 5$ $m = 5$	.0100	.0111	.0147	.0312	.0660	.1285	.2282	.3707	.6511	.9232
10	.0500	.0553	.0713	.1365	.2486	.4031	.5822	.7559	.9388	.9975
$n = 5$ $m = 6$	.0100	.0114	.0158	.0365	.0813	.1621	.2892	.4631	.7648	.9733
11	.0500	.0547	.0691	.1305	.2405	.3965	.5801	.7584	.9423	.9979
$n = 6$ $m = 6$	.0100	.0117	.0170	.0431	.1011	.2064	.3667	.5696	.8611	.9927
12	.0500	.0558	.0736	.149	.282	.461	.656	.825	.969	.9994
$n = 6$ $m = 7$	.0100	.0120	.0182	.0484	.1156	.2362	.4139	.6257	.8961	.9957
13	.050	.057	.077	.163	.311	.504	.704	.865	.982	1.000
$n = 7$ $m = 7$	.010	.012	.018	.049	.119	.250	.442	.664	.923	.998
14	.050	.057	.080	.176	.341	.550	.753	.900	.990	1.000

Epstein - Rosenbaum - Moses

N	Δ									
	0.00	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.75	0.09
$n = 4$ $m = 5$	.0100	.0152	.0227	.0476	.0914	.1617	.2651	.4046	.6684	.9247
9	.0500	.0717	.1005	.1815	.2938	.4315	.5828	.7312	.9087	.9920
$n = 5$ $m = 5$	.0100	.0154	.0231	.0489	.0949	.1686	.2762	.4190	.6825	.9296
10	.0500	.0719	.1009	.1831	.2971	.4365	.5885	.7365	.9114	.9923
$n = 5$ $m = 6$	.0100	.0159	.0248	.0553	.1102	.1968	.3187	.4725	.7328	.9469
11	.0500	.0749	.1091	.2071	.3401	.4947	.6520	.7925	.9389	.9955
$n = 6$ $m = 6$	.0100	.0162	.0255	.0585	.1201	.2204	.3631	.5401	.8135	.9805
12	.0500	.0752	.1100	.2112	.3512	.5153	.6809	.8240	.9589	.9984
$n = 6$ $m = 7$	.0100	.0167	.0271	.0647	.1341	.2433	.3922	.5692	.8301	.9826
13	.0500	.0784	.1186	.2352	.3901	.5612	.7230	.8537	.9680	.9988
$n = 7$ $m = 7$	.0100	.0170	.0281	.0699	.1512	.2823	.4582	.6537	.8964	.9941
14	.0500	.0788	.1200	.2431	.4106	.5956	.7645	.8904	.9829	.9997
$n = 7$ $m = 8$	.0100	.0176	.0299	.0769	.1668	.3069	.4877	.6815	.9111	.9959
15	.0500	.0822	.1291	.2668	.4452	.6310	.7916	.9057	.9857	.9997
$n = 8$ $m = 8$	.0100	.0178	.0306	.0810	.1797	.3335	.5265	.7220	.9313	.9973
16	.0500	.0830	.1317	.2792	.4739	.6725	.8333	.9355	.9933	.9999
$n = 8$ $m = 9$	.0100	.0185	.0329	.0919	.2074	.3817	.5884	.7814	.9587	.9992
17	.0500	.0859	.1399	.3008	.5022	.6979	.8496	.9427	.9942	.9999
$n = 9$ $m = 9$	.0100	.0188	.0338	.0952	.2158	.3959	.6051	.7951	.9629	.9993
18	.0500	.0878	.1449	.3192	.5386	.7416	.8862	.9635	.9973	.9999
$n = 9$ $m = 10$	.0100	.0205	.0385	.1150	.2629	.4703	.6889	.8615	.9826	.9998
19	.0500	.0947	.1611	.3520	.5759	.7691	.9012	.9683	.9975	.9999
$n = 10$ $m = 10$	.0100	.0200	.0384	.1145	.2624	.4702	.6886	.8614	.9825	.9998
20	.0500	.0933	.1623	.3656	.6043	.8028	.9244	.9790	.9985	.9999

Therefore, $G(x) \leqq \underline{G}(x, u, \Delta)$, since $G(x) \leqq u = \underline{G}(x, u, \Delta)$ for $u \leqq x < u + \Delta$, and $G(x) \leqq U(x) = \underline{G}(x, u, \Delta)$ for $x < u + \Delta$ and for $u + \Delta \leqq x$. But ϕ is p.o., so $\beta(\phi; U, G) \geqq \beta(\phi; U, \underline{G}(\cdot, u, \Delta))$, so $\underline{\beta}(\phi; U; \Delta) \geqq \inf_u \beta(\phi; U, \underline{G}(\cdot; u, \Delta))$, which finishes the proof of (ii)

The maximum power is equalled by uniform shift power.

LEMMA. $\beta(\phi; U, \bar{G}(\cdot, \Delta)) = \beta(\phi; U, \bar{H}(\cdot, \Delta))$.

PROOF. Consider any arrangement of X's and Y's (from $\bar{H}$) with j of the Y's $\geqq 1$: similarly, consider the same arrangement of X's and Y's (from $\bar{G}$) with

j of the Y's $= 1$. Since $\Pr\{Y \geqq 1\} = \Delta$ in both cases and otherwise $\bar{H}$ and $\tilde{G}$ are identical, it is obvious that the probability densities of the two arrangements are equal. But these yield the same ranking, whence the lemma.

5. Maximum power results. The numbers in the tables indicate the supremum of the powers against $C(F, \Delta)$ for various statistics defined below. An expression for the probability of any ordering for m X's from U and n Y's from $\tilde{G}(\cdot, \Delta)$ was summed over the rejection region of each statistic, randomized to sizes .05 and .01, except that the factorial increase of the rejection region forced the use of monte carlo approximation for $N = 20$, and also for the Doksum-Bell statistic, since it does not have a fixed rejection region. The results of the monte carlo are shown as the number of rejections out of 1000 trials. The round-off errors for $\Delta = 0$ on the exact calculations indicate that the powers are correct to the number of decimal places given.

With $X(j)$ the jth order statistic of m X's from U and $Y(k)$ the kth order statistic of n Y's from $\tilde{G}(\cdot, \Delta)$,

$$\Pr_\Delta\{R(X(m)) = m + p, R(Y(1)) = m + 1 - q\} = \sum_{k=0}^{n} \sum_{j=0}^{m} \Pr\{k \text{ of } Y\text{'s} \geqq 1, j \text{ of } X\text{'s} \leqq \Delta\}.$$

$$\Pr\{R(X(m)) = m + p, R(Y(1)) = m + 1 - q \mid k \text{ of } Y\text{'s} \geqq 1, j \text{ of } X\text{'s} \leqq \Delta\} = \sum_{k=0}^{n} \sum_{j=0}^{m} \binom{n}{k}\Delta^k(1-\Delta)^{n-k}\binom{m}{j}\Delta^j(1-\Delta)^{m-j}\cdot B(j, k)$$

where

$$B(j,k) = 0, \quad \text{if } k > n - p \text{ or if } j > m - q,$$
$$= \binom{p-q-2}{p-1}\binom{N-j-k}{n-k}^{-1}, \quad \text{otherwise.}$$

P. van der Laan [13] showed that, for fixed Δ, the probability under uniform shift of a particular ranking is a function only of $R(X(m))$ and $R(Y(1))$. Therefore, the probability of any one of the $\binom{p-q-2}{p-1}$ rankings is given by

$$\sum_{k=0}^{n-p} \binom{N}{k}\Delta^k \sum_{j=0}^{m-q} \binom{N-k}{j}\Delta^j(1-\Delta)^{N-k-j},$$

with $p = R(x(m)) - m$, $q = m + 1 - R(y(1))$. This formula checks correctly against P. van der Laan's expression ([13], Table 8) for the special case $m = n = 3$.

Letting $A(p, q, N, \alpha)$ be the number (which may be non-integer) of distinct orderings which lead to rejection with $R(X(m)) = m + p$, $R(Y(1)) = m + 1 - q$, the following lemma has been shown:

LEMMA. *The maximum power against* $C(F, \Delta)$ *is*

$$\beta(\Delta) = \sum_{p=0}^{n} \sum_{q=0}^{m} A(p,q,N,\alpha)\binom{N}{n}^{-1} \sum_{k=0}^{n-p} \binom{N}{k}\Delta^k \sum_{j=0}^{m-q} \binom{N-k}{j}\Delta^j(1-\Delta)^{N-k-j}.$$

If this is differentiated with respect to Δ and Δ is set to zero, one has

$$\beta'(0) = N\binom{N}{n}^{-1}\left[\sum_{p=0}^{n} \sum_{q=0}^{m} A(p, q, N, \alpha) - \sum_{p=1}^{n} A(p, m, N, \alpha) - \sum_{q=1}^{m} A(n, q, N, \alpha)\right],$$

and since $\beta(0) = \alpha$, $\beta'(0) = N(\alpha - \Pr\{\text{reject and a } Y \text{ smallest} \mid H_0\} - \Pr\{\text{reject and an } X \text{ largest} \mid H_0\})$, so

LEMMA. *$\beta'(0) = N\alpha - m \Pr_0\{reject \mid Y(1) < X(1)\} - n \Pr_0\{reject \mid Y(n) < X(m)\}$ is the initial power slope against uniform shift, with size α.*

6. Statistics used. Let $Z(k)$ denote the kth order statistic from the standard normal cpf, Φ. Let $\zeta(k)$ be the expected value of $Z(k)$. Let R be the ranking function for the combined sample. Then the statistics were used in the forms:

Fisher-Yates [12]: $\sum_{k=1}^{n} \zeta(R(Y_k))$
van der Waerden [14]: $\sum_{k=1}^{n} \Phi^{-1}(R(Y_k)/(N+1))$
Mann-Whitney-Wilcoxon [8]: $\sum_{k=1}^{n} R(Y_k)$
Doksum-Bell [2]: $\sum_{k=1}^{n} Z(R(Y_k)) - \sum_{k=1}^{n} Z(R(X_k))$, $(n = m)$
Similar Savage: $\sum_{k=1}^{n} S(R(Y_k))$, with $S(k) = \sum_{j=N+1-k}^{N} j^{-1}$.
Cramér-von Mises [4]: $\sum_{k=1}^{n} \sum_{j=1}^{m} (R(Y_k) - R(X_j))^2$

The test corresponding to each of the above statistics was: "reject if too large". Therefore, the Similar Savage is not equivalent to the test proposed by Savage [9], but the results are included since the work had been done, and the test is undoubtedly the most powerful rank test against *some* class of alternatives. The Epstein-Rosenbaum-Moses test [8] used was, "reject if too many Y's are greater than all of the X's", i.e., if $X_{(m)} < Y_{(k)}$ for a specified k. The Cramér-von Mises test is not one-sided, but was included for a few cases for general interest.

Perhaps it should be commented that, as well as being asymptotically equivalent, the Fisher-Yates and van der Waerden tests were equivalent up to $n = 6$, $m = 7$ and also coincided with the Similar Savage up to $n = 5$, $m = 5$.

7. Acknowledgments. Thanks are due to the referee and D. G. Chapman for helpful comments. The work was done in part at the computer center of San Diego State College, and in part at the Western Data Processing Center of the Graduate School of Business Education at the University of California at Los Angeles.

REFERENCES

[1] BELL, C. B. (1964). A characterization of multisample distribution-free statistics. *Ann. Math. Stat.* **35** 735–738.
[2] BELL, C. B., and DOKSUM, K. A. (1965). Some new distribution-free statistics. *Ann. Math. Statist.* **36** 203–214.
[3] CHAPMAN, D. G. (1958). A comparative study of several one-sided goodness-of-fit tests. *Ann. Math. Statist.* **29** 655–674.
[4] DARLING, D. A. (1957). The Kolmogorov, Cramér-von Mises tests. *Ann. Math. Statist.* **28** 834.
[5] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.
[6] LOÉVE, M. (1960). *Probability Theory*, 2d ed. van Nostrand, Princeton, New Jersey.
[7] MANN, H. B., and WHITNEY, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.* **18** 50–60.
[8] ROSENBAUM, S. (1954). Tables for a nonparametric test of location. *Ann. Math. Statist.* **25** 146–150.

[9] SAVAGE, I. R. (1956). Contributions to the theory of rank order statistics—the two-sample case. *Ann. Math. Statist.* **27** 590–615.

[10] SCHEFFÉ, H. (1943). On a measure problem arising in the theory of non-parametric tests. *Ann. Math. Statist.* **14** 227–233.

[11] TERRY, M. E. (1952). Some rank order tests which are most powerful against specific parametric alternatives. *Ann. Math. Statist.* **23** 346–366.

[12] THOMPSON, RORY (1964). Slope at $\Delta = 0$ of power curve of MWW versus uniform alternatives. Unpublished Memorandum.

[13] VAN DER LAAN, P. (1964). Exact power of some rank tests. To be published in Publ. de l'Inst. de Statistique de l'Université de Paris.

[14] VAN DER WAERDEN, B. L. (1953). Ein Neuer Test für das Problem der Zwei Stichproben. *Math. Ann.* **126** 93–107.