ON THE DISTRIBUTION OF THE MAXIMUM VALUE OF AN EQUALLY CORRELATED SAMPLE FROM A NORMAL POPULATION

By A. KUDO

Kyushu University, Fukuoka, Japan

and

Indian Statistical Institute, Calcutta

SUMMARY. The density and the distribution function of equivariance multivariate normal sample are derived. The distribution function of the maximum of sample is also derived and the table for it is given.

1. Introduction

The purpose of this paper is to give the distribution of the maximum of an equally correlated sample from a normal universe. As an application we shall give the distribution of the statistic for the testing of outlying observation. The table necessary for the calculation of the distribution of the maximum value has been calculated on a small scale in the Indian Statistical Institute. The scale of the table was decided for the sake of another application, confidence interval for the extreme value in future observations. The confidence interval for it is discussed in Kudô (1954).

2. The derivation of the distribution

Let $(x_1, x_2, \ldots, x_N)$ be distributed normally with a common mean value zero and a common variance σ^2, and we assume that they are equally correlated with correlation coefficient ρ. As the correlation matrix should be positive, we should have $\rho \geqslant 1/N-1$. As the variance-covariance matrix is

$$\Lambda=\sigma^2\begin{pmatrix} 1 & & \rho \\ & 1 & \\ & \ddots & \\ \rho & & 1\end{pmatrix}=(1-\rho)\sigma^2\begin{pmatrix} 1 & & 0 \\ & 1 & \\ & \ddots & \\ 0 & & 1\end{pmatrix}+\rho\sigma^2\begin{pmatrix} 1 & 1 & \ldots & 1 \\ 1 & 1 & \ldots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \ldots & 1\end{pmatrix},$$

we have the characteristic function

$$\varphi(\underset{\sim}{t}) = \exp(-\tfrac{1}{2}\underset{\sim}{t}'\Delta\underset{\sim}{t}) = \exp\left(-\tfrac{1}{2}(1-\rho)\sigma^2\sum_{i=1}^{n}\ {}^2-\tfrac{1}{2}\rho\sigma^2\left(\sum_{i=1}^{n} t_i\right)^2\right)$$

$$=\sum_{K=0}^{\infty}\frac{1}{K!}\left(\frac{-\rho\sigma^2}{2}\right)^K\left(\sum_{i=1}^{N}\quad\right)^{2k}\exp\left[-\tfrac{1}{2}(1-\rho)\sigma^2\sum_{i=1}^{N} t_i^2\right]$$

$$=\sum_{K=0}^{\infty}\frac{(2K)!}{k!}\left(\frac{-\rho\sigma^2}{2}\right)^k\sum_{n_1+n_2+\ldots+n_N=2k}\frac{t_1^{n_1}}{n_1!}\frac{t_2^{n_2}}{n_2!}\cdots\frac{t_N^{n_N}}{n_N}\exp\left[-\tfrac{1}{2}(1-\rho)\sigma^2\sum_{i=1}^{N} t_i^2\right].$$

As this series is dominated by an exponential function, we can take the inverse Fourier transformation term by term.

Let $f(x)$ be an integrable function, $\varphi(t)$ be its Fourier transform, and $f(x)$ be differentiable up to n-th order. Then we can easily verify that the Fourier transform of the n-th derivative of $f(x)$ is $(-it)^n\varphi(t)$. Hence the density function of $(x_1, x_2, \ldots, x_N)$ is given by

$$f(x_1, x_2, \ldots, x_N) = \sum_{K=0}^{\infty} \frac{(2K)!}{K!} \left(\frac{\rho\sigma^2}{2}\right)^K (\sqrt{1-\rho}\,\sigma)^{2k+N} \sum_{n_1+\ldots+n_N=2k} \frac{1}{n_1!} F^{(n_1)}\left(\frac{x_1}{\sqrt{1-\rho}\,\sigma}\right) \ldots \ldots \frac{1}{n_N!} F^{(n_N)}\left(\frac{x_N}{\sqrt{1-\rho}\,\sigma}\right),$$

where $F^{(n)}(x)$ is the n-th derivative of the standardized normal distribution function. Therefore the distribution function of $(x_1, x_2, \ldots, x_N)$ is

$$F(x_1, x_2, \ldots, x_N) = \sum_{K=0}^{\infty} \frac{(2K)!}{K!} \left(\frac{\rho\sigma^2}{2}\right)^K \{(1-\rho)\sigma^2\}^{k+N} \sum_{n_1+\ldots+n_N=2K} \frac{1}{n_1!} F^{(n_1)}\left(\frac{x_1}{\sqrt{1-\rho}\,\sigma}\right) \ldots \ldots \frac{1}{n_N!} F^{(n_N)}\left(\frac{x_N}{\sqrt{1-\rho}\,\sigma}\right).$$

Let $a_n(x) = \frac{1}{n!} F^{(n)}(x)$ and then we have

$$F(x_1, x_2 \ldots, x_N) = \sum_{K=0}^{\infty} \frac{(2K)!}{K!} \left(\frac{\rho\sigma^2}{2}\right)^K ((1-\rho)\sigma^2)^{k+n} \sum_{n_1+n_2+\ldots n_N=2K} a_{n_1}\left(\frac{x_1}{\sqrt{1-\rho}\,\sigma}\right) \ldots \ldots a_{n_N}\left(\frac{x_N}{\sqrt{1-\rho}\,\sigma}\right).$$

In case $x_1 = x_2 = \ldots = x_N = x$, it has a special expression :

$$F(x, \ldots, x) = Pr\,(\max x_i < x) = \sum_{K=0}^{\infty} \frac{(2K)!}{K!} \left(\frac{\rho\sigma^2}{2}\right)^K \{(1-\rho)\sigma^2\}^{k+n} A_{n,2k}\left(\frac{x}{\sqrt{1-\rho}\,\sigma}\right),$$

where $A_{n,K}(x)$ are defined in recurrent ways,

$$A_{1,k}(x) = a_k(x), \quad A_{n,K}(x) = \sum_{i=0}^{K} a_i(x)\, A_{n-1,k-i}(X).$$

3. Tables

The tables were calculated in the following way. At first we tabulated $a_n(x)$ up to the seventh decimal place. Then $A_{n,K}(x)$ were tabulated by I.B.M. machines. The relations

$$\sum_{K=0}^{8} A_{n,K}(x) = \sum_{K=0}^{8} A_{n-1,K}(x) \sum_{i=0}^{8-k} a_i(X)$$

were used for check purpose at each stage of calculation and the numbers were listed upto the fifth decimal place. To examine the accuracy, another tabulation was carried out for $x = 3$ with ten decimal places, and there was found no discrepancy.

TABLES OF $A_{N,n}(X)$

$x/^n$	0	2	4	6	8
			$N = 1$		
1.8	0.96407	−.07106	−.00142	.00136	−.00017
1.9	.97128	−.06234	−.00317	.00140	−.00015
2.0	.97725	−.05399	−.00450	.00135	−.00012
2.1	.98214	−.04618	−.00543	.00124	−.00008
2.2	.98610	−.03902	−.00598	.00108	−.00005
2.3	.98928	−.03258	−.00622	.00090	−.00002
2.4	.99180	−.02687	−.00618	.00070	.00001
2.5	.99379	−.02191	−.00593	.00051	.00003
2.6	.99534	−.01766	−.00553	.00034	.00004
2.7	.99653	−.01407	−.00503	.00019	.00005
2.8	.99744	−.01108	−.00447	.00006	.00005
2.9	.99813	−.00863	−.00389	−.00004	.00005
3.0	.99865	−.00665	−.00332	−.00011	.00004
3.1	.99903	−.00506	−.00279	−.00016	.00004
3.2	.99931	−.00381	−.00230	−.00019	.00003
3.3	.99952	−.00284	−.00187	−.00019	.00002
3.4	.99966	−.00209	−.00149	−.00019	.00002
3.5	.99977	−.00153	−.00118	−.00018	.00001
			$N = 2$		
1.8	0.92043	−.13077	.00696	.00308	−.00074
1.9	.94339	−.11678	.00148	.00362	−.00063
2.0	.95602	−.10261	−.00296	.00361	−.00049
2.1	.96459	−.08878	−.00633	.00343	−.00034
2.2	.97239	−.07570	−.00867	.00306	−.00020
2.3	.97867	−.06365	−.01009	.00258	−.00008
2.4	.98367	−.05280	−.01074	.00206	.00001
2.5	.98762	−.04324	−.01078	.00154	.00007
2.6	.99070	−.03497	−.01035	.00106	.00012
2.7	.99308	−.02793	−.00960	.00065	.00013
2.8	.99490	−.02204	−.00865	.00032	.00014
2.9	.99627	−.01719	−.00761	.00006	.00013
3.0	.99730	−.01326	−.00654	−.00013	.00011
3.1	.99807	−.01011	−.00552	−.00026	.00009
3.2	.99863	−.00762	−.00457	−.00033	.00007
3.3	.99903	−.00568	−.00372	−.00037	.00005
3.4	.99933	−.00419	−.00298	−.00037	.00004
3.5	.99953	−.00305	−.00235	−.00035	.00002

TABLES OF $A_{N,n}(X)$ *(continued)*

$x \backslash n$	0	2	4	6	8
			$N = 3$		
1.8	0.89604	−.18010	.02277	.00373	−.00172
1.9	.91630	−.16387	.01246	.00532	−.00154
2.0	.93329	−.14614	.00373	.00605	−.00124
2.1	.94736	−.12794	−.00321	.00600	−.00088
2.2	.95887	−.11011	−.00833	.00563	−.00054
2.3	.96817	−.09326	−.01178	.00486	−.00026
2.4	.97561	−.07781	−.01376	.00395	−.00003
2.5	.98149	−.06400	−.01457	.00302	.00011
2.6	.98608	−.05193	−.01447	.00214	.00021
2.7	.98964	−.04159	−.01372	.00139	.00025
2.8	.99235	−.03289	−.01255	.00077	.00025
2.9	.99441	−.02569	−.01115	.00028	.00023
3.0	.99596	−.01983	−.00960	−.00007	.00020
3.1	.99710	−.01513	−.00818	−.00030	.00017
3.2	.99794	−.01141	−.00680	−.00045	.00013
3.3	.99855	−.00851	−.00555	−.00052	.00010
3.4	.99890	−.00628	−.00445	−.00054	.00007
3.5	.99930	−.00458	−.00351	−.00052	.00004
			$N = 4$		
1.8	.86384	−.21991	.04393	.00235	−.00294
1.9	.88999	−.20410	.02847	.00599	−.00285
2.0	.91206	−.18485	.01477	.00806	−.00241
2.1	.93043	−.16381	.00346	.00879	−.00180
2.2	.94554	−.14233	−.00525	.00851	−.00117
2.3	.95770	−.12145	−.01142	.00757	−.00061
2.4	.96761	−.10191	−.01532	.00628	−.00018
2.5	.97539	−.08420	−.01735	.00489	.00011
2.6	.98149	−.06855	−.01790	.00355	.00029
2.7	.98620	−.05504	−.01739	.00237	.00037
2.8	.98982	−.04361	−.01616	.00139	.00039
2.9	.99256	−.03412	−.01451	.00064	.00036
3.0	.99461	−.02637	−.01267	.00008	.00031
3.1	.99614	−.02013	−.01079	−.00029	.00026
3.2	.99725	−.01519	−.00899	−.00053	.00020
3.3	.99807	−.01133	−.00736	−.00065	.00014
3.4	.99865	−.00836	−.00591	−.00069	.00010
3.5	.99907	−.00610	−.00467	−.00068	.00006

TABLES OF $A_{.8}(X)$ (*continued*)

x; n	0	2	4	6	8
			$N = 5$		
1.8	.83280	−.25105	.06864	−.00168	−.00409
1.9	.86443	−.23794	.04832	.00496	−.00441
2.0	.89131	−.21901	.02943	.00917	−.00396
2.1	.91381	−.19652	.01323	.01117	−.00310
2.2	.93239	−.17242	.00033	.01145	−.00211
2.3	.94752	−.14824	−.00915	.01054	−.00120
2.4	.95968	−.12512	−.01549	.00896	−.00047
2.5	.96933	−.10384	−.01915	.00709	.00003
2.6	.97691	−.08484	−.02068	.00524	.00035
2.7	.98278	−.06830	−.02063	.00358	.00050
2.8	.98729	−.05422	−.01951	.00220	.00055
2.9	.99071	−.04248	−.01771	.00112	.00051
3.0	.99327	−.03286	−.01557	.00032	.00045
3.1	.99517	−.02511	−.01333	−.00023	.00036
3.2	.99657	−.01896	−.01116	−.00057	.00028
3.3	.99759	−.01415	−.00915	−.00076	.00020
3.4	.99832	−.01044	−.00736	−.00083	.00014
3.5	.99884	−.00762	−.00583	−.00083	.00009
			$N = 6$		
1.8	.80288	−.27428	.09536	−.00864	−.00481
1.9	.83961	−.26583	.07095	.00182	−.00600
2.0	.87103	−.24886	.04701	.00899	−.00580
2.1	.89749	−.22621	.02568	.01294	−.00477
2.2	.91943	−.20045	.00817	.01424	.00339
2.3	.93736	−.17367	−.00512	.01364	−.00205
2.4	.95181	−.14746	−.01434	.01188	−.00093
2.5	.96332	−.12294	−.02002	.00959	−.00014
2.6	.97236	−.10079	−.02282	.00720	.00037
2.7	.97938	−.08135	−.02345	.00501	.00063
2.8	.98477	−.06471	−.02257	.00317	.00071
2.9	.98886	−.05078	−.02074	.00171	.00068
3.0	.99193	−.03933	−.01838	.00064	.00059
3.1	.99421	−.03008	−.01582	−.00011	.00048
3.2	.99588	−.02272	−.01328	−.00058	.00037
3.3	.99710	−.01697	−.01092	−.00085	.00027
3.4	.99798	−.01253	−.00881	−.00096	.00018
3.5	.99861	−.00014	−.00608	−.00097	.00012

TABLES OF $A_{N,n_e}(X)$ (*continued*)

x, n	0	2	4	6	8
			$N = 7$		
1.8	.77403	−.29033	.12275	−.01859	−.00472
1.9	.81550	−.28820	.09542	−.00366	−.00738
2.0	.85122	−.27463	.06690	.00724	−.00780
2.1	.88146	−.25302	.04043	.01384	−.00676
2.2	.90664	−.22651	.01802	.01668	−.00502
2.3	.92730	−.19778	.00055	.01672	−.00317
2.4	.94401	−.16894	−.01195	.01497	−.00150
2.5	.95733	−.14149	−.01908	.01231	−.00043
2.6	.96782	−.11640	−.02433	.00939	−.00033
2.7	.97598	−.09421	−.02585	.00664	−.00073
2.8	.98225	−.07509	−.02537	.00429	.00088
2.9	.98701	−.05901	−.02360	.00242	.00086
3.0	.99050	−.04575	−.02108	.00103	.00076
3.1	.99325	−.03502	−.01824	.00006	.00062
3.2	.99520	−.02647	−.01537	−.00056	.00047
3.3	.99662	−.01978	−.01267	−.00091	.00034
3.4	.99764	−.01460	−.01024	−.00108	.00024
3.5	.99837	−.01066	−.00812	−.00111	.00015
			$N = 8$		
1.8	.74622	−.29087	.14071	−.03137	−.00344
1.9	.79208	−.30546	.12091	−.01159	−.00826
2.0	.83185	−.29657	.08853	.00372	−.00980
2.1	.86571	−.27705	.05712	.01366	−.00900
2.2	.89404	−.25064	.02968	.01861	−.00697
2.3	.91736	−.22060	.00773	.01967	−.00460
2.4	.93627	−.18958	−.00838	.01815	−.00247
2.5	.95139	−.15952	−.01909	.01522	−.00086
2.6	.96331	−.13169	−.02524	.01179	.00022
2.7	.97260	−.10687	−.02785	.00846	.00081
2.8	.97974	−.08535	−.02791	.00557	.00104
2.9	.98517	−.06717	−.02629	.00325	.00105
3.0	.98925	−.05214	−.02368	.00151	.00093
3.1	.99229	−.03994	−.02061	.00029	.00076
3.2	.99452	−.03021	−.01743	−.00050	.00059
3.3	.99614	−.02258	−.01440	−.00096	.00043
3.4	.99731	−.01668	−.01166	−.00118	.00029
3.5	.99814	−.01218	−.00926	−.00124	.00019

TABLES OF $A_{N,n}(X)$ (*continued*)

$x \backslash n$	0	2	4	6	8
			$N = 9$		
1.8	.71941	−.30352	.17530	−.04669	.00068
1.9	.76933	−.31797	.14669	−.02195	−.00838
2.0	.81293	−.31488	.11138	−.00169	−.01161
2.1	.85024	−.29844	.07540	.01224	−.01140
2.2	.88161	−.27291	.04294	.01987	−.00921
2.3	.90752	−.24217	.01631	.02237	−.00631
2.4	.92860	−.20940	−.00370	.02133	−.00357
2.5	.94548	−.17702	−.01736	.01826	−.00144
2.6	.95882	−.14666	−.02556	.01436	.00002
2.7	.96923	−.11933	−.02945	.01045	.00085
2.8	.97724	−.09550	−.03019	.00698	.00120
2.9	.98333	−.07527	−.02882	.00418	.00125
3.0	.98792	−.05849	−.02618	.00206	.00112
3.1	.99133	.04484	−.02291	.00057	.00092
3.2	.99383	−.03394	−.01946	−.00041	.00071
3.3	.99566	−.02537	−.01612	−.00099	.00052
3.4	.99697	−.01875	−.01306	−.00127	.00035
3.5	.99791	−.01369	−.01039	−.00136	.00023
			$N = 10$		
1.8	.69356	−.30187	.19874	−.06415	.00381
1.9	.74724	−.32610	.17211	−.03464	−.00748
2.0	.79443	−.32979	.13490	−.00905	−.01304
2.1	.83505	−.31730	.09497	.00945	−.01385
2.2	.86935	−.29339	.05759	.02034	−.01171
2.3	.89779	−.26251	.02616	.02473	−.00832
2.4	.92098	−.22842	.00203	.02446	−.00492
2.5	.93961	−.19400	−.01483	.02141	−.00218
2.6	.95435	−.16131	−.02531	.01709	−.00027
2.7	.96587	−.13160	−.03065	.01259	.00084
2.8	.97474	−.10558	−.03221	.00853	.00134
2.9	.98150	−.08330	−.03119	.00521	.00144
3.0	.98658	−.06480	−.02859	.00268	.00132
3.1	.99037	−.04972	−.02516	.00089	.00109
3.2	.99315	−.03766	−.02145	−.00029	.00084
3.3	.99518	−.02817	−.01781	−.00099	.00061
3.4	.99664	−.02082	−.01446	−.00135	.00042
3.5	.99768	−.01521	−.01151	−.00148	.00027

4. AN APPLICATION

Let x_i $(i = 1, 2, ..., N_1+N_2)$ be independently and identically distributed as $N(m, \sigma^2)$, where σ is known and m is unknown to us. Let $x_M = \max_{i=1,...,N_1} x_i$ and $\bar{x} = \sum_{i=1}^{N_1+N_2} x_i/N_1+N_2$. The statistic $(x_M - \bar{x})/\sigma$ have been proved to be optimum in a certain sense for the testing of an outlying observation, under such a circumstance as we know *a priori* there may be only one observation among $x_i (i = 1, 2, ..., N_1)$ which is from another normal population with a same variance and a greater mean value (Kudō, 1956).

In case $N_2 = 0$, the distribution was given by Mackey (1935), at first, and then Nair (1948) and Grubbs (1950) gave another form of it and tabulated the tables of this distribution.

As $y_i = x_i - \bar{x}$ $(i = 1, 2, ... N_1)$ are equally correlated we have

$$Pr\left(\frac{X_M - \bar{x}}{\sigma} < \lambda\right) = \sum_{K=0}^{\infty} \frac{(2k)!}{k!} \left(\frac{-1}{2(N_1+N_2)}\right)^k A_{N,2k}(\lambda).$$

The convergence is found to be quicker as λ increases.

The following is the comparison of our approximation with the values given by Grubbs for $N_2 = 0$, $\lambda = 1.8$, 2.5 and 3.5.

	1.8		2.5		3.5	
	Grubbs	approximation	Grubbs	approximation	Grubbs	approximation
$N = 3$	.95877	.95936	.99670	.99642	.99997	1.00000
$N = 4$	.92480	.92530	.99222	.99209	.99989	.99989
$N = 5$	.89037	.89066	.98703	.98696	.99977	.99978
$N = 6$	.85646	.85675	.98151	.98146	.99962	.99963
$N = 7$	.82341	.82364	.97580	.97576	.99945	.99945
$N = 8$	.79139	.79155	.96999	.96997	.99927	.99927
$N = 9$	.76046	.76058	.96412	.96411	.99908	.99908
$N = 10$	.73063	.73071	.95823	.95823	.99888	.99887

5. ACKNOWLEDGEMENTS

The author is much indebted to Professor P. C. Mahalanobis for providing him with a generous scholarship to work at the Indian Statistical Institute. He is grateful to Dr. C. R. Rao and Dr. D. Basu for their encouragement and advice in this work. Finally to all those who did the tabulation part of this paper, especially to Mr. D. Bose and Mr. A Halder, the author's gratitude is equally great.

REFERENCES

GRUBBS, F. E. (1950): Sampling criteria for testing outlying observations. *Ann. Math. Stat.*, **21**, 27.

KUDÔ, A. (1956): On the testing of outlying observations. *Sankhyā*, **17**, 67.

——— (1956a): On the confidence interval for the extreme value of a second sample from a normal universe. Bull. Math. Stat., **6**, 51.

MACKEY, A. T. (1935): The distribution of the difference between the extreme observation and the sample mean in samples of n from a normal universe. *Biometrika*, **27**, 466.

NAIR, K. R. (1948): The distribution of the extreme deviate from the sample mean and its studentized form. *Biometrika*, **35**, 118.

Paper received : December, 1955.