ON A STATISTIC WHICH ARISES IN SELECTION AND RANKING PROBLEMS[1]

By Shanti S. Gupta and Milton Sobel

Bell Telephone Laboratories

1. Summary. The statistic $y = (x_{[p]} - x)/s_\nu$ is studied where $x_{[p]}$ is the maximum of p normal independent chance variables with common mean and common unknown variance σ^2, x is another independent normal chance variable with the same mean and the same variance σ^2, and s_ν^2 (distributed as $\sigma^2\chi_\nu^2/\nu$ with ν degrees of freedom) is an estimate of the common variance which is independent of each one of the above $p + 1$ chance variables. Several different methods are proposed and studied for computing the probability integral of y and percentage points of y; in addition, a method for computing percentage points without first computing the probability integral of y is considered. A table of (upper) percentage points of y is given as Table I at the end of the paper. Applications of the statistic y to several ranking and selection problems are mentioned in Section 2. Moments of y are given in Section 3. In Section 7 it is shown that Table I can be used to obtain an approximation and bounds to the percentage points of a related statistic.

2. Introduction. The statistic y was considered by Gupta [8] for the problem of selecting a subset of normal populations (with a common unknown variance σ^2) which contains the "best" population with probability at least P^*. For an explanation of how the statistic y enters particular problems the reader is referred to [8] and [9]. If we let z_i denote normal chance variables $N(0, \sigma^2)$ with correlations $\rho_{ij} = 1/2$ for $i \neq j$ then it is easy to see that

$$P\left\{\frac{z_i}{s_\nu} < \frac{q}{\sqrt{2}} \, (i = 1, 2, \cdots, p)\right\} = P\{y < q\}. \tag{2.1}$$

The left-hand member of (2.1) is a special case of the probability integral of the multivariate analogue of Student's t-distribution which is treated by Dunnett and Sobel in [6] and [7]. For $p = 2$ tables of the percentage points and the probability integral are given in [6]. Expressions for bounds on the probability integral for all p are given in [7].

An application of the distributions in (2.1) to a problem of comparing several populations with a control is given in [5] where percentage points of $y/\sqrt{2}$ are given to two decimals for $p = 1\,(1)\,9$, selected values of ν and α (or $1 - P^*$) = .05, .01. The distribution of y is also needed for the solution of a problem formulated by Paulson [12]. Three more problems for which the distribution of y is needed are (i) the selection by a sequential procedure of the "best" population

Received December 14, 1956; revised July 1, 1957.

[1] This research was supported in part by the Signal Corps, United States Army under contract DA 36-039 Sc-64618.

with probability at least P^* of being correct, (ii) the selection by a multi-stage procedure of a subset which contains the "best" population with probability at least P^* of being correct, and (iii) the selection by a single stage procedure of a subset which contains all populations "better" than a standard with probability at least P^* of being correct. The last problem is treated in [9].

The reader should distinguish between the statistic y and the Nair's studentized extreme deviate [10] defined by $(x_{[p]} - \bar{x})/s_\nu$ where $\bar{x}$ is the average of the p ordered $x_{[i]}$'s.

The statistic v, which corresponds to y for the σ-known case, is treated in detail in [8]. Percentage points of v are identical with the values of $\lambda\sqrt{N}$ in Table I of Bechhofer's paper [1] for the columns headed ($k = i$, $t = 1$; $i = 2, 3, \cdots, 10$).

3. Moments of y. The moments of y can be written as a product of two expectations. For $r < \nu$

$$\mu_r' = Ey^r = E\left[\left(\frac{s_\nu}{\sigma}\right)^{-r}\right] E\left[\left(\frac{x_{[p]} - x}{\sigma}\right)^r\right] \tag{3.1}$$

$$= A_{-r}\left[\sum_{j=0}^{[r/2]} C_{2j}^r \frac{(2j)!}{2^j j!} a_{p,r-2j}\right], \tag{3.2}$$

where $[r/2]$ is the largest integer less than or equal to $r/2$, $a_{p,i}$ is the ith moment of the largest of p independent standard normal chance variables, and

$$A_\beta = E\left\{\left(\frac{s_\nu}{\sigma}\right)^\beta\right\} = \frac{\Gamma\left(\frac{\nu+\beta}{2}\right)}{\left(\frac{\nu}{2}\right)^{\beta/2}\Gamma\left(\frac{\nu}{2}\right)} \tag{3.3}$$

is the βth moment of $\chi_\nu/\sqrt{\nu}$ for *both positive and negative* values of β, provided that $\beta > -\nu$. For $i = 1$ (1) 10 and $p = 1$ (1) 50 the value of $a_{p,i}$ is obtainable from Ruben's paper [13]. For $i > 10$ the value of $a_{p,i}$ can be obtained with the aid of some results of Bose and Gupta [2]; in particular, for $p < i$ the value of $a_{p,i}$ can be obtained, without integration, as a linear function of the $a_{p,j}$ ($j = i - 2, i - 4 \cdots; j \geqq 0$).

4. The probability integral of y and methods for its evaluation. Consider the expression

$$P = P\left[\frac{x_{[p]} - x}{s_\nu} < q\right]. \tag{4.1}$$

If we hold x and s_ν fixed and integrate first with respect to $x_{[p]}$, then, letting $q' = q/\sqrt{\nu}$, we obtain

$$P = \int_0^\infty g_\nu(\chi)\left[\int_{-\infty}^\infty F^p(u) f(u - \chi q')\, du\right] d\chi, \tag{4.2}$$

where f, F are respectively the standard normal density and the standard normal c.d.f. and g_ν is the chi-density with ν degrees of freedom. Similarly, holding another pair fixed, we obtain two alternative forms which are not given here. Another expression for P in the form of a p-fold integral is given in [7].

METHOD 1. Now we shall derive some series expansions for P using (4.2). If the function f in (4.2) is expanded about $q' = 0$, then it is easy to justify a term by term integration and we obtain

$$P = \frac{1}{p+1} \sum_{j=0}^{\infty} \frac{q^j}{j!} A_j E\{H_j(u)\} = \frac{e^{qAH(a)}}{p+1}, \tag{4.3}$$

where A_j is given by (3.3). Here $H_j(x)$ is the jth Hermite Polynomial defined as in [4], the expectation is taken with respect to the density of $u = x_{[p+1]}/\sigma$ given by $(p+1)F^p(u)f(u)$, and the last expression in (4.3) is a symbolic one in which powers of A, H and a (say A^α, H^α and a^α) are to be replaced by A_α, H_α and $a_{p+1,\alpha}$, respectively. The expansion (4.3) converges for all q and the convergence is rapid for small values of q. For example, for $q = \sqrt{2}/2$, $p = 2$ and $\nu = 2$ using (4.3) with $j = 0$ to 10 and Ruben's table [13] we obtain $P = .520175$ as compared with the value .52017 given in Table 1 of [6] (in the column $h = .50$ and the row $n = 2$.).

METHOD 2. Another expansion is derived by expanding the function f in (4.2) about $s_\nu = \sigma$ (i.e., about $\chi = \sqrt{\nu}$), obtaining

$$\begin{aligned} P = \int_{-\infty}^{\infty} F^p(u) \Big[f(u-q) - q f^{(1)}(u-q) E(\chi' - 1) \\ + \frac{q^2}{2!} f^{(2)}(u-q) E(\chi'-1)^2 - \cdots \Big] du, \end{aligned} \tag{4.4}$$

where $f^{(\alpha)}(x)$ is the αth derivative of the normal density $f(x)$ and

$$\chi' = \chi/\sqrt{\nu} = s_\nu/\sigma .$$

The first term in (4.4) is the integral for the σ-known case tabulated in [8] and [11]. After straightforward integration term-by-term in (4.4) we obtain

$$\begin{aligned} P = \int_{-\infty}^{\infty} F^p(u) f(u-q)\, dq - q(1-A_1) \int_{-\infty}^{\infty} (u-q) F^p(u) f(u-q)\, du \\ + \frac{q^2}{2!} 2(1-A_1) \int_{-\infty}^{\infty} [u^2 - 2qu + (q^2-1)] F^p(u) f(u-q)\, du \\ - \frac{q^3}{3!} \left[\frac{4\nu - (4\nu+1)A_1}{\nu} \right] \int_{-\infty}^{\infty} [u^3 - 3u^2 q + 3u(q^2-1) - (q^3 - 3q)] \\ \cdot F^p(u) f(u-q)\, du + \cdots \end{aligned} \tag{4.5}$$

where A_1 is given by (3.3). Each of the integrals above is evaluated by expanding the factor e^{qu} in $f(u-q)$, thus obtaining a series in terms of $a_{p+1,\alpha}$. In the symbolic notation used above we can write P in the form

$$P = \int_{-\infty}^{\infty} F^p(u) f(u-q)\, du + \frac{e^{-q^2/2}}{p+1} \{e^{q[a+(A-1)(a-H(q))]} - e^{qa}\}. \tag{4.6}$$

For $p = 2$, $\nu = 2$ and $q = \sqrt{2}/2$ using the first six terms of (4.5) or (4.6) we obtain $P = 0.5192$. This is comparable with the result $P = 0.5188$ obtained by using the first six terms of (4.3). Equation (4.5) gives a better answer in this case since the correct value is 0.52017. In general, we expect (4.5) to give better results for large ν than (4.3); however, the computations for (4.5) are more involved.

In writing (4.6) we have used the fact that

$$(4.7)\quad \int_{-\infty}^{\infty} F^{p}(u)f(u-q)\,du = \frac{e^{-q^2/2}}{p+1}\left[1 + qa_{p+1,1} + \frac{q^2}{2!}a_{p+1,2} + \cdots\right].$$

This series converges for all q and converges rapidly for small values of q. For example, for $q = \sqrt{2}$ and $p = 2$ using the first eleven terms of (4.7) we obtain $P = .7442$ as compared with .7452 obtained from [11] by interpolation. It should be noted that the sum of any finite number of terms in (4.7) gives a lower bound to P for all integers $p \geqq 1$, and all non-negative q. This follows from the easily shown result that $a_{j,\alpha} \geqq 0$ for all integers $j \geqq 1$, $\alpha \geqq 0$. If the tables of Ruben [13] were extended to moments higher than the tenth, then better accuracy could be obtained.

METHOD 3. Another method of calculating the probability integral is based on the result of Seal [15] that the distribution of $v = (x_{[p]} - x)/\sigma$ is asymptotically normal as p tends to infinity. It follows from his result that the third and higher central moments of v tend to the corresponding moments of the standardized normal distribution. Since the coefficients involving ν in A_{-r} in (3.2) tend to unity as $\nu \to \infty$ it follows that the third and higher central moments of y tend to the corresponding moments of the standardized normal distribution as both ν and p tend to infinity. Hence y is asymptotically normally distributed as both ν and p tend to infinity. It is therefore reasonable to approximate the distribution of $y_s = (y - E(y))/\sigma(y)$ by a Gram-Charlier expansion in the Edgeworth form where

$$(4.8)\quad E(y) = A_{-1}a_{p,1} \quad\text{and}\quad \sigma(y) = [A_{-2}(a_{p,2}+1) - (A_{-1}a_{p,1})^2]^{1/2}.$$

Using equation (17.7.3) of [4] and letting $q_s = (q - E(y))/\sigma(y)$ we obtain

$$(4.9)\quad \begin{aligned} P = P[y < q] = P[y_s < q_s] = F(q_s) &- \left[\frac{\alpha_3}{3!}f^{(2)}(q_s)\right] \\ &+ \left[\frac{\alpha_4}{4!}f^{(3)}(q_s) + \frac{10\alpha_3^2}{6!}f^{(5)}(q_s)\right] \\ &- \left[\frac{\alpha_5}{5!}f^{(4)}(q_s) + \frac{35\alpha_3\alpha_4}{7!}f^{(6)}(q_s) + \frac{280\alpha_3^3}{9!}f^{(8)}(q_s)\right] \\ &+ \left[\frac{\alpha_6}{6!}f^{(5)}(q_s) + \frac{(35\alpha_4^2 + 56\alpha_3\alpha_5)}{8!}f^{(7)}(q_s) \right.\\ &\qquad \left. + \frac{21\alpha_3^2\alpha_4}{10!}f^{(9)}(q_s) + \frac{154\alpha_3^4}{12!}f^{(11)}(q_s)\right] + \cdots, \end{aligned}$$

where α_r is the rth standardized (i.e. relative) cumulant of y and given by $\alpha_r = \kappa_r/\kappa_2^{r/2}$. The κ's can be obtained from the moments around the origin μ_r'

(given in Section 3) by the usual formulae. For $p = 9$, $\nu = 18$ and $q_s = 1.70$ we obtain $q = 3.703$ and equation (4.9) gives

$$P = .955435 - .008743 + .001485 + .003342 - .001355 = .950162$$

which rounds to .950.

METHOD 4. This probability can also be computed by using the Gauss quadrature method based on the zeros and weight factors of the 15th degree Laguerre polynomial which are given by Salzer and Zucker in [14]. The values of the inside integral in (4.2) at the zeros of the Laguerre polynomial were computed by interpolation in [11]. This gives for $q = 3.703$ the result $P = .9495$ which rounds to .950. This agrees to three decimal places with the result obtained above by Method 3.

5. Expression for the percentage points of y. In practice it is of much greater interest to compute the percentage points corresponding to fixed probability levels rather than the probability integral. Cornish and Fisher [3] have developed a technique for computing the percentage points of a statistic directly without first computing a table of the probability integral values. Since y is asymptotically normal, their technique is applicable here. Applying their result we obtain for the upper percentage point $q_s = q_s(P^*)$

$$\begin{aligned} q_s(P^*) = x_s(P^*) &+ [\alpha_3 I_c] + [\alpha_4 I_d + \alpha_3^2 I_{c^2}] + [\alpha_5 I_e + \alpha_3\alpha_4 I_{cd} + \alpha_3^3 I_{c^3}] \\ &+ [\alpha_6 I_f + \alpha_4^2 I_{d^2} + \alpha_3\alpha_5 I_{ce} + \alpha_3^2\alpha_4 I_{c^2d} + \alpha_3^4 I_{c^4}] + \cdots, \end{aligned} \tag{5.1}$$

where $x_s(P^*)$ is the standard normal deviate corresponding to P^*, α_r is the relative cumulant defined earlier, and I_c, I_d, I_{c^2}, $\cdots$ are tabulated in Table I of [3] for the probability levels $P^* = .75, .90, .95, .975, .99, .995, .9975, .999$ and .9995. Then q can be found from q_s using (4.8).

The same method can be used for lower percentage points except that the 2nd, 4th, $\cdots$ correction terms in (5.1) change sign. Combining the two results we can obtain "2-sided" percentage points with tails of equal probability, $(1 - P^*)/2$. It should be noted that the lower percentage point corresponding to a probability of $1/(p + 1)$ is zero for all values of ν.

6. Construction of tables. The procedure described in Section 5 was used to compute the percentage points of y in Table I for $P^* = .75, .90, .95, .975$ and .99 for $k = p + 1 = 2, 5, 10\ (1)\ 16, 18, 20\ (5)\ 40, 50$ and $\nu = 15\ (1)\ 20, 24, 30, 36, 40, 48, 60, 80, 100, 120, 360$ and ∞.

We may regard the expressions in square brackets in (5.1) as correction terms. The expression (5.1) was calculated to 3 correction terms and also to 4 correction terms and the result is given in Table I to two decimal places. In some cases for small degrees of freedom where this method gives only one decimal agreement one of the methods in Section 4 was used to find the second decimal. For $p = 1$ (or $k = 2$) the percentage point of y has been checked against $\sqrt{2}$ times the corresponding percentage point of Student's t-distribution.

For example, when $P^* = .95$, $\nu = 18$ and $p = 9$, using 3 and 4 correction

TABLE I

Percentage points of y *for selected values of* p *and* ν

$P^* = .75$

ν \ k	2	5	10	11	12	13	14	15	16	18	20	25	30	35	40	50
p	1	4	9	10	11	12	13	14	15	17	19	24	29	34	39	49
15	0.98	1.91	2.35	2.41	2.46	2.49	2.54	2.57	2.61	2.65	2.72	2.82	2.90	2.97	3.03	3.12
16	0.98	1.90	2.35	2.40	2.45	2.49	2.53	2.56	2.60	2.65	2.71	2.81	2.89	2.96	3.02	3.11
17	0.97	1.90	2.34	2.40	2.44	2.48	2.52	2.55	2.60	2.64	2.70	2.80	2.88	2.95	3.02	3.10
18	0.97	1.90	2.34	2.39	2.44	2.48	2.52	2.55	2.59	2.64	2.70	2.80	2.88	2.95	3.01	3.09
19	0.97	1.90	2.34	2.39	2.44	2.48	2.52	2.55	2.59	2.63	2.69	2.79	2.87	2.94	3.00	3.09
20	0.97	1.89	2.33	2.38	2.43	2.47	2.51	2.54	2.58	2.63	2.69	2.79	2.87	2.94	3.00	3.08
24	0.97	1.89	2.32	2.37	2.42	2.46	2.50	2.53	2.57	2.62	2.67	2.77	2.85	2.92	2.98	3.06
30	0.97	1.88	2.31	2.36	2.41	2.45	2.49	2.52	2.55	2.61	2.66	2.76	2.84	2.91	2.96	3.05
36	0.96	1.87	2.30	2.35	2.40	2.44	2.48	2.51	2.54	2.60	2.65	2.75	2.83	2.89	2.95	3.04
40	0.96	1.87	2.30	2.35	2.39	2.43	2.47	2.50	2.54	2.59	2.64	2.74	2.82	2.89	2.94	3.03
48	0.96	1.87	2.29	2.34	2.39	2.43	2.47	2.50	2.53	2.59	2.64	2.74	2.82	2.88	2.93	3.02
60	0.96	1.86	2.29	2.34	2.38	2.42	2.46	2.49	2.52	2.58	2.63	2.73	2.81	2.87	2.93	3.01
80	0.96	1.86	2.28	2.33	2.38	2.42	2.45	2.49	2.52	2.57	2.62	2.72	2.80	2.86	2.92	3.01
100	0.96	1.86	2.28	2.33	2.37	2.41	2.45	2.48	2.51	2.57	2.62	2.72	2.79	2.86	2.91	3.00
120	0.96	1.85	2.27	2.33	2.37	2.41	2.45	2.48	2.51	2.57	2.61	2.71	2.79	2.86	2.91	3.00
360	0.95	1.85	2.27	2.32	2.36	2.40	2.44	2.47	2.50	2.56	2.61	2.70	2.78	2.84	2.90	2.99
∞	0.95	1.85	2.26	2 31	2.36	2.40	2.44	2.47	2.50	2.55	2.60	2.70	2.78	2.84	2.89	2.98

$P^* = .90$

ν \ k	2	5	10	11	12	13	14	15	16	18	20	25	30	35	40	50
p	1	4	9	10	11	12	13	14	15	17	19	24	29	34	39	49
15	1.90	2.77	3.21	3.26	3.31	3.36	3.39	3.43	3.46	3.52	3.58	3.69	3.79	3.84	3.90	4.02
16	1.89	2.76	3.19	3.24	3.29	3.34	3.37	3.41	3.44	3.50	3.55	3.65	3.75	3.81	3.86	3.97
17	1.89	2.75	3.18	3.23	3.27	3.33	3.35	3.39	3.43	3.48	3.53	3.63	3.73	3.80	3.84	3.94
18	1.88	2.74	3.17	3.21	3.26	3.31	3.34	3.38	3.41	3.47	3.52	3.62	3.71	3.78	3.83	3.92
19	1.88	2.73	3.16	3.20	3.25	3.30	3.33	3.37	3.40	3.46	3.51	3.61	3.70	3.77	3.82	3.91
20	1.87	2.72	3.15	3.20	3.24	3.29	3.32	3.36	3.39	3.45	3.50	3.60	3.69	3.75	3.81	3.90
24	1.86	2.70	3.12	3.17	3.21	3.26	3.29	3.33	3.36	3.42	3.47	3.56	3.65	3.72	3.77	3.86
30	1.85	2.68	3.09	3.14	3.19	3.23	3.26	3.30	3.33	3.38	3.43	3.53	3.61	3.68	3.73	3.82
36	1.85	2.67	3.07	3.12	3.17	3.21	3.24	3.27	3.31	3.36	3.41	3.51	3.59	3.65	3.70	3.79
40	1.84	2.66	3.06	3.11	3.15	3.19	3.23	3.25	3.29	3.34	3.39	3.49	3.57	3.63	3.68	3.77
48	1.84	2.65	3.05	3.10	3.14	3.19	3.22	3.25	3.28	3.33	3.38	3.48	3.55	3.62	3.67	3.76
60	1.83	2.64	3.04	3.08	3.13	3.17	3.20	3.23	3.26	3.32	3.36	3.46	3.54	3.60	3.65	3.74
80	1.83	2.63	3.02	3.07	3.11	3.15	3.19	3.22	3.25	3.30	3.35	3.44	3.52	3.58	3.63	3.72
100	1.82	2.62	3.01	3.06	3.10	3.14	3.18	3.21	3.24	3.29	3.34	3.43	3.51	3.57	3.62	3.71
120	1.82	2.62	3.01	3.06	3.10	3.14	3.17	3.20	3.23	3.28	3.33	3.43	3.50	3.56	3.61	3.70
360	1.82	2.61	2.99	3.04	3.08	3.12	3.15	3.18	3.21	3.26	3.31	3.40	3.48	3.53	3.59	3.67
∞	1.81	2.60	2.98	3.03	3.07	3.11	3.14	3.17	3.20	3.25	3.30	3.39	3.46	3.52	3.58	3.66

terms in (5.1), we obtain $q_s = 1.645 + .084 - .013 - .017 - .004 = 1.695$ (1.699 for 3 terms). This gives $q = 3.697$ (3.702 for 3 terms). Since $p = 9$ the values $q = 3.70$ and 3.71 can be checked by a Gauss-Laguerre quadrature (Method 4 in Section 4). The results obtained were $P(3.70) = .9493$ and $P(3.71) = .9501$. This gives $q/\sqrt{2} = 2.62$ which is comparable with the result $t = 2.62$ given in [5].

$P^* = .95$

ν \ k / p	2 / 1	5 / 4	10 / 9	11 / 10	12 / 11	13 / 12	14 / 13	15 / 14	16 / 15	18 / 17	20 / 19	25 / 24	30 / 29	35 / 34	40 / 39	50 / 49
15	2.48	3.34	3.78	3.84	3.89	3.93	3.97	4.01	4.05	4.11	4.17	4.28	4.37	4.45	4.51	4.62
16	2.47	3.32	3.75	3.80	3.85	3.89	3.93	3.97	4.01	4.07	4.13	4.23	4.32	4.39	4.46	4.56
17	2.46	3.29	3.72	3.77	3.82	3.86	3.90	3.94	3.98	4.04	4.09	4.20	4.28	4.35	4.42	4.52
18	2.45	3.28	3.70	3.75	3.80	3.84	3.88	3.92	3.95	4.01	4.06	4.17	4.25	4.32	4.39	4.49
19	2.45	3.26	3.68	3.73	3.78	3.82	3.86	3.90	3.93	3.99	4.04	4.15	4.23	4.30	4.36	4.46
20	2.44	3.25	3.67	3.72	3.77	3.81	3.85	3.89	3.92	3.97	4.02	4.13	4.21	4.28	4.34	4.44
24	2.42	3.22	3.63	3.68	3.73	3.77	3.80	3.84	3.87	3.92	3.97	4.07	4.15	4.23	4.28	4.37
30	2.40	3.19	3.59	3.64	3.68	3.72	3.75	3.79	3.82	3.87	3.92	4.02	4.10	4.17	4.22	4.31
36	2.39	3.16	3.56	3.61	3.65	3.69	3.72	3.76	3.79	3.84	3.89	3.99	4.06	4.13	4.18	4.27
40	2.38	3.15	3.54	3.59	3.63	3.67	3.71	3.74	3.77	3.82	3.87	3.97	4.04	4.11	4.16	4.25
48	2.37	3.14	3.52	3.57	3.61	3.65	3.68	3.72	3.75	3.80	3.85	3.94	4.02	4.08	4.13	4.22
60	2.36	3.12	3.50	3.55	3.59	3.63	3.66	3.69	3.72	3.77	3.82	3.91	3.99	4.05	4.10	4.19
80	2.35	3.10	3.48	3.53	3.57	3.60	3.64	3.67	3.70	3.75	3.80	3.89	3.96	4.02	4.07	4.16
100	2.35	3.09	3.47	3.51	3.55	3.59	3.62	3.66	3.68	3.74	3.78	3.87	3.95	4.01	4.06	4.14
120	2.34	3.09	3.46	3.50	3.55	3.58	3.62	3.65	3.67	3.73	3.77	3.86	3.93	3.99	4.05	4.13
360	2.33	3.07	3.43	3.48	3.52	3.55	3.59	3.62	3.65	3.70	3.74	3.83	3.90	3.96	4.01	4.09
∞	2.33	3.06	3.42	3.46	3.50	3.54	3.57	3.60	3.63	3.68	3.72	3.81	3.88	3.94	3.99	4.07

$P^* = .975$

ν \ k / p	2 / 1	5 / 4	10 / 9	11 / 10	12 / 11	13 / 12	14 / 13	15 / 14	16 / 15	18 / 17	20 / 19	25 / 24	30 / 29	35 / 34	40 / 39	50 / 49
15	3.01	3.87	4.33	4.39	4.44	4.48	4.53	4.56	4.59	4.65	4.72	4.79	4.91	4.96	5.00	5.08
16	3.00	3.82	4.28	4.33	4.39	4.43	4.47	4.50	4.53	4.59	4.66	4.74	4.85	4.90	4.95	5.04
17	2.98	3.79	4.24	4.29	4.34	4.39	4.42	4.46	4.50	4.55	4.61	4.70	4.80	4.85	4.91	5.00
18	2.97	3.77	4.21	4.26	4.31	4.35	4.39	4.43	4.46	4.51	4.57	4.66	4.76	4.82	4.88	4.97
19	2.96	3.75	4.18	4.23	4.28	4.32	4.36	4.40	4.43	4.49	4.54	4.64	4.73	4.79	4.85	4.94
20	2.95	3.74	4.16	4.21	4.26	4.30	4.34	4.38	4.41	4.46	4.52	4.61	4.70	4.77	4.83	4.91
24	2.92	3.69	4.10	4.15	4.20	4.23	4.27	4.31	4.34	4.39	4.44	4.54	4.63	4.69	4.75	4.84
30	2.89	3.64	4.04	4.08	4.13	4.17	4.20	4.24	4.27	4.32	4.37	4.47	4.55	4.61	4.67	4.76
36	2.87	3.61	4.00	4.04	4.09	4.12	4.16	4.19	4.22	4.28	4.32	4.42	4.50	4.56	4.62	4.70
40	2.86	3.60	3.98	4.02	4.07	4.10	4.14	4.17	4.20	4.25	4.30	4.40	4.47	4.53	4.59	4.67
48	2.84	3.57	3.95	3.99	4.03	4.07	4.11	4.14	4.17	4.22	4.26	4.36	4.43	4.49	4.55	4.63
60	2.83	3.55	3.92	3.96	4.00	4.04	4.07	4.10	4.13	4.18	4.23	4.32	4.39	4.46	4.51	4.59
80	2.81	3.52	3.89	3.93	3.97	4.01	4.04	4.07	4.10	4.15	4.19	4.28	4.36	4.42	4.47	4.55
100	2.81	3.51	3.87	3.91	3.95	3.99	4.02	4.05	4.08	4.13	4.17	4.26	4.33	4.39	4.44	4.53
120	2.80	3.50	3.86	3.90	3.94	3.98	4.01	4.04	4.07	4.11	4.16	4.25	4.32	4.38	4.43	4.51
360	2.78	3.47	3.82	3.86	3.90	3.93	3.97	3.99	4.02	4.07	4.11	4.20	4.27	4.33	4.37	4.45
∞	2.77	3.45	3.80	3.84	3.88	3.91	3.95	3.97	4.00	4.05	4.09	4.18	4.24	4.30	4.35	4.43

The percentage points in all cases should be regarded as having been rounded to the *nearest* decimal so that the exact probability may be slightly under or slightly over the desired value P^*.

7. On a related set of percentage points. C. W. Dunnett has proposed in [5] the related problem of finding values of d'' such that

$$P\left\{\left|\frac{z_i}{s_\nu}\right| < d''(i = 1, 2, \cdots, p)\right\} = P^*, \tag{7.1}$$

$P^* = .99$

ν \ k	2	5	10	11	12	13	14	15	16	18	20	25	30	35	40	50
p	1	4	9	10	11	12	13	14	15	17	19	24	29	34	39	49
15	3.68	4.54	4.99	5.05	5.10	5.14	5.19	5.23	5.26	5.32	5.38	5.49	5.58	5.66	5.72	5.82
16	3.65	4.48	4.92	4.98	5.03	5.07	5.11	5.15	5.18	5.24	5.30	5.42	5.51	5.59	5.65	5.75
17	3.63	4.44	4.87	4.92	4.97	5.02	5.06	5.10	5.13	5.19	5.25	5.37	5.46	5.53	5.60	5.69
18	3.61	4.41	4.83	4.88	4.93	4.97	5.01	5.05	5.08	5.14	5.20	5.32	5.41	5.48	5.55	5.64
19	3.59	4.38	4.79	4.84	4.89	4.93	4.97	5.01	5.04	5.10	5.16	5.28	5.37	5.44	5.50	5.59
20	3.58	4.35	4.76	4.81	4.86	4.90	4.94	4.98	5.01	5.07	5.13	5.24	5.33	5.40	5.47	5.55
24	3.52	4.28	4.67	4.72	4.77	4.81	4.85	4.88	4.91	4.97	5.02	5.13	5.22	5.29	5.35	5.43
30	3.47	4.21	4.59	4.63	4.68	4.72	4.75	4.78	4.81	4.87	4.92	5.02	5.10	5.17	5.23	5.30
36	3.44	4.16	4.53	4.58	4.62	4.66	4.69	4 72	4.75	4.80	4.85	4.95	5.03	5.10	5.15	5.22
40	3.43	4.13	4.50	4.55	4.59	4.63	4.66	4.69	4.72	4.77	4.82	4.92	4.99	5.06	5.11	5.19
48	3.41	4.10	4.46	4.50	4.54	4.58	4.61	4.64	4.67	4.72	4.77	4.86	4.93	5.00	5.05	5.13
60	3.38	4.06	4.41	4.46	4.50	4.53	4.56	4.59	4.62	4.67	4.72	4.81	4.88	4.94	4.99	5.07
80	3.36	4.02	4.37	4.41	4.45	4.49	4.52	4.55	4.58	4.63	4.67	4.76	4.83	4.89	4.94	5.02
100	3.34	4.00	4.35	4.39	4.43	4.46	4.49	4.52	4.55	4.60	4.64	4.73	4.80	4.85	4.90	4.98
120	3.33	3.99	4.33	4.37	4.41	4.44	4.47	4.50	4.53	4.58	4.62	4.71	4.77	4.83	4.88	4.96
360	3.30	3.94	4.27	4.31	4.35	4.38	4.41	4.44	4.47	4.51	4.55	4.64	4.71	4.76	4.81	4.88
∞	3.29	3.92	4.25	4.29	4.32	4.35	4.38	4.41	4.44	4.48	4.52	4.61	4.67	4.73	4.77	4.85

where P^* is specified and z_i are normal chance variables $N(0, \sigma^2)$ with correlations $\rho_{ij} = \frac{1}{2}$ for $i \neq j (i, j = 1, 2, \cdots, p)$. It will be convenient to introduce the symbols $m = d''\sqrt{2}$ and $m' = m/\sqrt{\nu}$. The exact integral expression which defines m' and hence also d'' is given by

$$\text{(7.2)} \qquad \int_0^\infty g_\nu(\chi) \left\{ \int_{-\infty}^\infty [F(u + \chi m') - F(u - \chi m')]^p f(u)\, du \right\} d\chi = P^*.$$

It is interesting to note that (7.1) can be written in the form $P\{y_D < d''\} = P^*$ where in terms of independent normal chance variables $x, x_1, \cdots, x_p$ we can write $y_D = \max | x_i - x |/s_\nu$ in contrast with the statistic

$$y = \max (x_i - x)/s_\nu .$$

Let C_p denote the hypercube $| z_i | < d''s_\nu (i = 1, 2, \cdots, p)$, let $P\{C_p \mid m, \nu\}$ denote the left member of (7.1) and let

$$\text{(7.3)} \qquad M_p = M_p(m, \nu) = P\{z_i < d''s_\nu (i = 1, 2, \cdots, p)\},$$

which is given, for example, by (4.2) with $q' = d''\sqrt{2/\nu}$. We shall give upper and lower bounds on $P\{C_p \mid m, \nu\}$ in terms of M_p. Using an identity like (4.3.8) of [4] it can be shown (the proof is omitted) that

$$\begin{aligned} &1 - P\{\text{all } | z_i | < d''s_\nu\} \\ \text{(7.4)} \qquad &= (1 - P\{\text{all } z_i < d''s_\nu\}) + (1 - P\{\text{all } z_i > -d''s_\nu\}) \\ &\qquad -P\{\text{at least one } z_i \geqq d''s_\nu \text{ and at least one } z_i \leqq -d''s_\nu\}. \end{aligned}$$

Since the first two terms on the right-hand side of (7.4) are both equal to $1 - M_p$ we can write (7.4) as

$$\text{(7.5)} \qquad 2M_p - 1 \leqq P\{C_p \mid m, \nu\} = 2M_p - 1 + Q \leqq M_p$$

where $Q \geqq 0$ is the last probability in (7.4). For moderately large p we have the approximation

(7.6) $Q \cong P\{\text{at least one } z_i \geqq d''s_\nu\}\, P\,\{\text{at least one } z_i \leqq -d''s_\nu\} = (1 - M_p)^2$
so that form (7.5) and (7.6) we obtain

$$P\{C_p \mid m, \nu\} \cong M_p^2\,. \tag{7.7}$$

Hence an approximation d_2'' to the percentage points d'' for Dunnett's problem is obtained from a table of M_p-values by setting

$$M_p(m, \nu) = P^{*\frac{1}{2}} \quad \text{and} \quad d_2'' = \frac{m}{\sqrt{2}}. \tag{7.8}$$

An upper bound d_3'' is obtained from a table of M_p-values by setting

$$M_p(m, \nu) = \frac{1 + P^*}{2} \quad \text{and} \quad d_3'' = \frac{m}{\sqrt{2}}. \tag{7.9}$$

Finally a lower bound d_1'' is obtained from a table of M_p-values by setting

$$M_p(m, \nu) = P^* \quad \text{and} \quad d_1'' = \frac{m}{\sqrt{2}}. \tag{7.10}$$

The approximation obtained by using (7.10) is best for large d''; for smaller d'' we might consider another lower bound described below. It is conjectured that the approximations in (7.6) and (7.7) actually give upper bounds for all p so that (7.10) gives a lower bound on the percentage point but this has not been rigorously shown.

Appendix to Section 7

It is also possible to obtain an upper bound for the left member of (7.1) and a lower bound for the percentage point d'' satisfying (7.1). These are obtained by replacing the p-dimensional cube C_p' defined by

$$|\, t_i \,| \leqq d''s_\nu/\sigma; \qquad t_i = z_i/\sigma \qquad (i = 1, 2, \cdots, p) \tag{7.11}$$

by a central p-dimensional ellipsoid E_p

$$t'\Sigma^{-1}t = b^2, \tag{7.12}$$

where b is determined so that they both contain the same hypervolume. By a well-known argument which plays a prominent part in the work of Neyman and Pearson and which is omitted here it is easy to show that

$$\int_{C_p'} \exp\,(-\tfrac{1}{2}t'\,\Sigma^{-1}t)\; dt \leqq \int_{E_p} \exp\,(-\tfrac{1}{2}t'\Sigma^{-1}t)\; dt. \tag{7.13}$$

Using (11.12.3) of [4] and equating hypervolumes, we obtain

$$\frac{\pi^{p/2}b^p\,|\,\Sigma\,|^{1/2}}{\Gamma\left(1 + \frac{p}{2}\right)} = \left(\frac{2\,d''\chi}{\sqrt{\nu}}\right)^p \tag{7.14}$$

which determines $b^2 = b^2(\chi)$. Since $t'\Sigma^{-1}t$ is distributed as χ_p^2 with p degrees of freedom we can write for the exact value P of the left member of (7.1)

$$P \leqq \int_0^\infty P\{\chi_p^2 < b^2(\chi)\} g_\nu(\chi)\, d\chi. \tag{7.15}$$

For even p and any ν the right hand member of (7.15) can be integrated exactly and we obtain

$$P \leqq 1 - \frac{1}{\left(1 + \frac{k^2}{\nu}\right)^{\nu/2}} \sum_{i=0}^{(p/2)-1} C_{(\nu/2)-1}^{(\nu/2)+i-1} \left(\frac{k^2}{\nu + k^2}\right)^i \tag{7.16}$$

where

$$k^2 = \frac{8\, d''^2 \left[\Gamma\left(1 + \frac{p}{2}\right)\right]^{2/p}}{\pi(1 + p)^{1/p}}. \tag{7.17}$$

For $\nu = \infty$ and any p (even or odd) it follows from (7.15) that

$$P \leqq P\{\chi_p^2 < k^2\}. \tag{7.18}$$

If we set the right hand member equal to P^* and solve for d'' using (7.14) then we obtain a lower bound for d''. In particular, we obtain from (7.17), by setting $k^2 = \chi_0^2(p, P^*)$ which is the P^* – percentage point of χ_p^2, the result

$$d'' \geqq \frac{(1 + p)^{1/2p}}{4\left[\Gamma\left(1 + \frac{p}{2}\right)\right]^{1/p}} \sqrt{2\pi\chi_0^2(p, P^*)}. \tag{7.19}$$

Numerical calculations show that the results of this appendix do not give very close bounds for large values of d'' and P^*. They are included here principally for their theoretical interest.

8. Acknowledgements. The authors wish to acknowledge the substantial contribution of Miss Marilyn J. Huyett in the computation of Table I and the examples above. Thanks are also due to Prof. J. W. Tukey for helpful suggestions.

REFERENCES

[1] R. E. Bechhofer, "A single-sample multiple decision procedure for ranking means of normal populations with known variances," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 16–29.

[2] R. C. Bose and S. S. Gupta, "Moments of order statistics from a normal population," *Institute of Statistics Mimeograph Series No. 154* (July, 1956), University of North Carolina, Chapel Hill, N.C.

[3] E. A. Cornish and R. A. Fisher, "Moments and cumulants in the specification of distributions" reprinted in *Contributions to Mathematical Statistics*, John Wiley, N.Y. (1950).

[4] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, 1951.

[5] C. W. Dunnett, "A multiple comparison procedure for comparing several treatments with a control," *J. Amer. Stat. Assn.*, Vol. 50 (1955), pp. 1096–1121.

[6] C. W. DUNNETT AND M. SOBEL, "A bivariate generalization of Student's t-distribution, with tables for certain special cases," *Biometrika*, Vol. 41 (1954), pp. 153–169.

[7] C. W. DUNNETT AND M. SOBEL, "Approximations to the probability integral and certain percentage points of a multivariate analogue of Student's t-distribution," *Biometrika*, Vol. 42 (1955), pp. 258–260.

[8] S. S. GUPTA, "On a decision rule for a problem in ranking means," *Institute of Statistics Mimeograph Series No. 150* (May 1956), University of North Carolina, Chapel Hill, N.C.

[9] S. S. GUPTA AND M. SOBEL, "On selecting a subset which contains all populations better than a standard," submitted for publication in *Ann. Math. Stat.*

[10] K. R. NAIR, "The distribution of the extreme deviate from the sample mean and its studentized form," *Biometrika*, Vol. 351 (1948) pp. 118–144.

[11] National Bureau of Standards, Tables computed for R. E. Bechhofer sponsored by ONR and distributed by the Chemical Corps Engineering Agency, Army Chemical Center, Maryland, in the pamphlet, "Probabilities associated with order statistics in samples from two normal populations with equal variance," by D. Teichroew (1955).

[12] E. PAULSON, "A multiple decision procedure for certain problems in the analysis of variance," *Ann. Math. Stat.*, Vol. 20 (1949), pp. 95–98.

[13] H. RUBEN, "On the moments of order statistics in samples from normal populations," *Biometrika*, Vol. 41 (1954), pp. 200–227.

[14] H. E. SALZER AND R. ZUCKER, "Table of the zeros and weight factors of the first fifteen Laguerre polynomials," *Bull. Amer. Math. Soc.*, Vol. 55 (1949), pp. 1004–1012.

[15] K. C. SEAL, "On a class of decision procedures for ranking means," *Institute of Statistics Mimeograph Series No. 109* (June, 1954), University of North Carolina, Chapel Hill, N.C.