BAYES RULES FOR A COMMON MULTIPLE COMPARISONS PROBLEM AND RELATED STUDENT-t PROBLEMS[1]

By David B. Duncan

University of North Carolina[2]

0. Summary. The paper is mainly concerned with the following multiple comparisons problem in the analysis of variance setting. In a balanced experiment n treatments are to be compared. Each of the $\frac{1}{2}n(n-1)$ pairwise comparisons is to be made, adjudging each difference as "positive", "negative", or "not significant"; overall decisions involving intransitivities are barred. The loss for each difference is proportional to the error; if a difference is asserted incorrectly the loss has proportionality constant c_1, if "not-significant" is the incorrect conclusion the proportionality constant is c_0; where $c_1 = k_1 + k_0$, $c_0 = k_0$ and $k_1 > k_0 > 0$. Total loss for the experiment is taken as the sum of the $\frac{1}{2}n(n-1)$ component losses. The Bayes rule for any prior distribution is shown as a result to consist in the simultaneous application of Bayes rules to the $\frac{1}{2}n(n-1)$ component problems. Each of these in turn is shown similarly to consist in the simultaneous application of Bayes rules to two subcomponent problems. The subcomponent Bayes rule for a normal prior density of treatment means is explicitly derived. The dependencies of the solution on the variance of the prior density, the degrees of freedom and the loss ratio k_1/k_0 are discussed. A principal finding is that the Bayes solution for the multiple comparisons problem corresponds to a tolerated error probability "of the first kind" for each single difference, that is independent of the number of treatments being compared.

1. Introduction. Many procedures have been proposed for the multiple comparisons problem herein considered. These include, for example, a "least-significant-difference" rule due to Fisher [5], an "honest-significant-difference" rule due to Tukey [17], [18] and multiple range testing procedures due to Newman [10], and the author [3]. Some of these have also been described in recent texts such as Federer [4], Li [9], Snedecor [13], Scheffé [12] and Steel and Torrie [14]. With much help from the recent more general work of Lehmann [7] it has now been possible to solve a Jeffrey's-like Bayes formulation of the problem. This is more complete than any of the previous formulations and leads to a simple solution with properties that are better defined and that appear to be appropriate to an appreciable class of practical situations. In the process, similar Jeffrey's-like Bayes formulations and their solutions are presented and obtained for two com-

Received November 21, 1960; revised June 16, 1961.

[1] Research sponsored jointly by the United States Air Force through the Office of Scientific Research of the Air Research and Development Command under Contracts No. AF 49(638)-929 and No. AF 49(638)-261, by the Office of Naval Research under contract No. Nonr 355(06), and by the U. S. Public Health Service. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[2] Now at the Johns Hopkins University.

mon types of Student-t problems. These are developed first as problems of separate interest in Sections 2 and 3. The main problem is then fully developed in Section 4. A discussion of the more applied aspects, together with further illustrations is planned for a paper to be submitted to *Biometrics*.

2. A two-decision Student-t problem. Given a random observation t from a non-central t distribution with non-centrality parameter τ and v degrees of freedom, a common problem is that of choosing between the two decisions

$$d_0\text{:decide } \tau \leqq \Delta \quad \text{and} \quad d_1\text{:decide } \tau > \Delta, \tag{2.1}$$

where Δ is some unspecified positive boundary value. In the language of the experimenter, d_0 is the decision that τ is not significantly greater than zero and d_1 the decision that it is. In the theory of hypothesis testing the same problem is often more loosely regarded as that of testing $H_0{:}\tau \leqq 0$ with the alternative $H_1{:}\tau > 0$, the decisions thus being

$$d_0\text{:decide } \tau \leqq 0 \quad \text{and} \quad d_1\text{:decide } \tau > 0 \tag{2.2}$$

Strictly speaking however the null decision does not deny the possibility of positive though relatively small values for τ and some such formulation as (2.1) is more precise. The change is relatively trivial in this problem by itself. It is essential however to our subsequent developments as is brought out shortly after (4.12). (See Lehmann [7] also for a similar change).

Our first result is a Bayes rule $\phi(t)$ for this problem with respect to a simple linear loss function

$$\begin{aligned} L_0(\tau) &= L(\tau, d_0) = \begin{cases} 0, & \tau \leqq 0, \\ k_0\tau, & \tau > 0, \end{cases} \\ L_1(\tau) &= L(\tau, d_1) = \begin{cases} k_1\,|\tau|\,, & \tau \leqq 0, \\ 0, & \tau > 0, \end{cases} \end{aligned} \tag{2.3}$$

where k_0 and k_1 are positive constants such that $k_1 > k_0$, and with respect to a normal prior density for τ,

$$\xi(\tau) = (2\pi\gamma^2)^{-\frac{1}{2}}e^{-\frac{1}{2}\tau^2/\gamma^2}, \qquad -\infty < \tau < \infty, \tag{2.4}$$

with mean zero and variance γ^2. The rule is of the common form

$$\phi_*(t) = \begin{cases} 0, & t < t_*, \\ 1, & t > t_*, \end{cases} \tag{2.5}$$

where $\phi(t) = 0$ or 1 is the usual indicator function for making the decision d_0 and d_1 respectively, and $t_* = t_*(k, v, \gamma^2)$ is a significant or critical t ratio for which a set of values are given in Table 1. The arguments determining the significant value t_*, are the ratio $k = k_1/k_0$ from the loss function, the degrees of freedom v for t and the variance γ^2 of the prior density for τ.

The rule $\phi_*(t)$ may be derived as follows. For the average risk of any rule

TABLE 1

Minimum-Average-Risk Significant t Values (t_* Values)

γ^2	v	Log k 0.0	.5	1.0	1.5	2.0	2.5	3.0	3.5
∞	1	0.0	.375	.807	1.353	2.102	3.160	4.685	6.854
	2	0.0	.413	.860	1.379	2.012	2.814	3.851	5.208
	4	0.0	.434	.884	1.367	1.900	2.502	3.197	4.010
	6	0.0	.443	.891	1.356	1.848	2.374	2.948	3.580
	14	0.0	.451	.898	1.340	1.779	2.217	2.654	3.099
	∞	0.0	.457	.902	1.326	1.721	2.091	2.436	2.759
3	1	0.0	.444	1.053	2.503	∞	∞	∞	∞
	2	0.0	.484	1.060	1.926	4.077	∞	∞	∞
	4	0.0	.506	1.056	1.718	2.623	4.178	9.595	∞
	6	0.0	.515	1.053	1.653	2.370	3.308	4.732	7.706
	14	0.0	.522	1.047	1.582	2.136	2.724	3.360	4.074
	∞	0.0	.528	1.041	1.531	1.987	2.414	2.813	3.186
1	1	0.0	.572	1.930	∞	∞	∞	∞	∞
	2	0.0	.610	1.532	8.741	∞	∞	∞	∞
	4	0.0	.629	1.395	2.648	8.592	∞	∞	∞
	6	0.0	.637	1.353	2.303	3.980	13.625	∞	∞
	14	0.0	.642	1.308	2.030	2.859	3.891	5.326	7.818
	∞	0.0	.646	1.275	1.875	2.433	2.957	3.445	3.902
.5	1	0.0	.767	∞	∞	∞	∞	∞	∞
	2	0.0	.785	2.292	∞	∞	∞	∞	∞
	4	0.0	.791	1.963	9.243	∞	∞	∞	∞
	6	0.0	.794	1.800	3.777	∞	∞	∞	∞
	14	0.0	.792	1.653	2.693	4.162	6.670	∞	∞
	∞	0.0	.792	1.562	2.296	2.980	3.622	4.219	4.779

$\phi(t)$, we have

$$A(\xi, \phi) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [L_0(\tau)(1 - \phi(t)) + L_1(\tau)\phi(t)]f(t \mid \tau)\, dt\, \xi(\tau)\, d\tau \tag{2.6}$$
$$= \text{constant} + \int_{-\infty}^{\infty} \phi(t)h_1(t)\, dt,$$

where $f(t \mid \tau)$ is the non-central t density function, (2.12) below, and

$$h_1(t) = \int_{-\infty}^{\infty} [L_1(\tau) - L_0(\tau)]f(t \mid \tau)\xi(\tau)\, d\tau. \tag{2.7}$$

The minimum average risk rule may thus be written

$$\phi_*(t) = \begin{cases} 0, & h_1(t) > 0, \\ 1, & h_1(t) < 0. \end{cases} \tag{2.8}$$

From the inequality $h_1(t) < 0$ we get

$$\int_{-\infty}^{0} k_1 |\tau| f(t \mid \tau)\xi(\tau)\, d\tau < \int_{-0}^{\infty} k_0\, \tau f(t \mid \tau)\xi(\tau)\, d\tau, \tag{2.9}$$

and hence

$$h_2(t) = \frac{\int_{-0}^{\infty} h_3(\tau, t)\, d\tau}{-\int_{-\infty}^{0} h_3(\tau, t)\, d\tau} > \frac{k_1}{k_0} = k, \tag{2.10}$$

where

$$h_3(\tau, t) = \tau f(t \mid \tau)\xi(\tau). \tag{2.11}$$

Now, introducing the non-central t density in the form

$$f(t \mid \tau) = \int_0^{\infty} \frac{e^{-\frac{1}{2}(ut-\tau)^2}}{(2\pi)^{\frac{1}{2}}} u\psi(u \mid v)\, du, \qquad -\infty < t < \infty, \tag{2.12}$$

where $\psi(u \mid v)$ is the χ^2-related density function of $u = \chi(v)/v$, that is

$$\begin{aligned} \psi(u \mid v) &= (v^{\frac{1}{2}v} u^{v-1} e^{-\frac{1}{2}vu^2})/[(\tfrac{1}{2}v - 1)!2^{\frac{1}{2}v-1}], \qquad u > 0, \\ &= 0, \quad \text{otherwise}, \end{aligned} \tag{2.13}$$

and discarding constants in $h_3(\tau, t)$ which will cancel and not affect the value of $h_2(t)$, we have

$$h_3(\tau, t) \propto \tau \int_0^{\infty} \exp\{\tfrac{1}{2}[(ut - \tau)^2 + \tau^2/\gamma^2]\} u\psi(u|v)\, du. \tag{2.14}$$

Putting $\beta^2 = 1 + 1/\gamma^2$ this becomes

$$\begin{aligned} h_3(\tau, t) &\propto \tau e^{-\frac{1}{2}\tau^2\beta^2} \int_0^{\infty} \exp\{ut\tau - \tfrac{1}{2}u^2(v + t^2)\} u^v\, du \\ &= \tau e^{-\frac{1}{2}\tau^2\beta^2} \sum_{i=0}^{\infty} \frac{(t\tau)^i}{i!} \int_0^{\infty} u^{v+i} \exp\{-\tfrac{1}{2}u^2(v + t^2)\}\, du \\ &= \tau e^{-\frac{1}{2}\tau^2\beta^2} \sum_{i=0}^{\infty} \frac{(t\tau)^i}{i!} \left(\frac{2}{v + t^2}\right)^{\frac{1}{2}i} \left(\frac{v + i - 1}{2}\right)!. \end{aligned} \tag{2.15}$$

Next, putting $h_4(t) = \int_0^{\infty} h(\tau, t) d\tau$ and $y = t/\beta(v + t^2)^{\frac{1}{2}}$ and integrating term by term, we get

$$\begin{aligned} h_4(t) &\propto \sum_{i=0}^{\infty} \frac{y^i}{i!} 2^{i/2} \left(\frac{v + i - 1}{2}\right)! \int_0^{\infty} (\tau\beta)^{i+1} e^{-\frac{1}{2}(\tau\beta)^2}\, d(\tau\beta) \\ &\propto \sum_{i=0}^{\infty} \frac{y^i}{i!} 2^i \left(\frac{v + i - 1}{2}\right)! \left(\frac{i}{2}\right)! \\ &= \sum_{i=0}^{\infty} \frac{y^i}{i!} 2^i \left(\frac{i + 1}{2}\right)! \left(\frac{i}{2}\right)! \left(\frac{v + i - 1}{2}\right)! \Big/ \left(\frac{i + 1}{2}\right)! \end{aligned} \tag{2.16}$$

But $2^i[(i+1)/2]!(i/2)! = (i+1)!\frac{1}{2}(\pi)^{\frac{1}{2}}$, from which

$$(2.17)\qquad \begin{aligned} h_4(t) &= \frac{d}{dy}\left[y \sum_{i=0}^{\infty} y^i \left(\frac{v+i-1}{2}\right)! \Big/ \left(\frac{i+1}{2}\right)!\right] \\ &= \frac{d}{dy}\,[yh_5(y) + yh_6(y)], \end{aligned}$$

where $h_5(y)$ is the sum of even terms,

$$h_5(y) = \sum_{i=0}^{\infty} y^{2i}(p+i-\tfrac{1}{2})!/(i+\tfrac{1}{2})!,$$

with $p = v/2$, and $h_6(y)$ is the sum of odd terms,

$$h_6(y) = \sum_{i=0}^{\infty} y^{2i+1}(p+i)!/(i+1)!.$$

Working first with $h_5(y)$, it may be written as

$$(2.18)\qquad h_5(y) = (p-\tfrac{1}{2})!F(p+\tfrac{1}{2}, 1; \tfrac{3}{2}; y^2)/(\tfrac{1}{2})!,$$

where $F(a, b; c; x)$ denotes the hypergeometric function

$$1 + \frac{abx}{c} + \frac{a(a+1)b(b+1)}{c(c+1)}\frac{x^2}{2} + \cdots.$$

Applying Euler's transformation

$$F(a, b; c; x) = (1-x)^{c-a-b}F(c-a, c-b; c; x)$$

and reducing, we get

$$(2.19)\qquad h_5(y) = \frac{1}{y}(1-y^2)^{-p}\int_0^y (1-u^2)^{p-1}\,du(p-\tfrac{1}{2})!/(\tfrac{1}{2})!.$$

Hence

$$(2.20)\qquad \frac{d}{dy}\,[yh_5(y)] = (1-y^2)^{-(p+1)}\left[(1-y^2)^p + 2py\int_0^y (1-u^2)^{p-1}\,du\right] \cdot (p-\tfrac{1}{2})!/(\tfrac{1}{2})!.$$

Next, $h_6(y)$ sums to $(p-1)![(1-y^2)^{-p} - 1]/y$ so that

$$(2.21)\qquad (d/dy)[yh_6(y)] = (1-y^2)^{-(p+1)}2p!y.$$

Treating the denominator of $h_2(t)$ in the same way and combining results we get

$$(2.22)\qquad h_2(t) = h_4(t)/h_4(-t) = g(y)/g(-y),$$

where

$$(2.23)\qquad g(y) = (1-y^2)^{\frac{1}{2}v} + vy\int_0^y (1-u^2)^{\frac{1}{2}(v-2)}\,du + \frac{2(\frac{1}{2})!(\frac{v}{2})!y}{[\frac{1}{2}(v-2)]!}.$$

From (2.8), (2.10) and (2.22) we have

$$\phi_*(t) = \begin{cases} 0, & g(y)/g(-y) < k, \\ 1, & g(y)/g(-y) > k. \end{cases} \tag{2.24}$$

Hence, since $g(y)/g(-y)$ is monotone increasing with respect to y,

$$\phi_*(t) = \begin{cases} 0, & y < y_*, \\ 1, & y > y_*, \end{cases} \tag{2.25}$$

where $y_* = y_*(k, v)$ is the solution for y in $g(y)/g(-y) = k$. Finally, putting $t_* = t_*(k, v, \gamma^2)$ for the solution for t in $y_* = t/\beta(v + t^2)^{\frac{1}{2}}$ we have

$$\phi_*(t) = \begin{cases} 0, & t < t_*, \\ 1, & t > t_*, \end{cases} \tag{2.26}$$

as was to be shown. More specific details of the computation of the significant t ratios in Table 1 are given in Section 6.

EXAMPLE 1. To illustrate suppose the following: A standard treatment is modified in the hopes of producing an increased yield. An experiment is run giving r yield observations $x_{11}, \cdots, x_{1r}$ and $x_{21}, \cdots, x_{2r}$ for the new and control (standard) treatment respectively. It can be assumed that the respective sets of data are random independent samples from normal populations with means μ_1 and μ_2 and with the same, but unknown, variance σ^2. It is required to decide whether the new treatment is significantly superior (in yield) or not significantly superior than the standard; whether to generally recommend it as the superior or to withhold such a general recommendation. Type-1-like errors of recommending a non-superior new treatment (making d_1 when $\delta \leqq 0$, where $\delta = \mu_1 - \mu_2$) are thought to increase in seriousness in direct proportion to the degree, $-\delta$, of inferiority involved. Type-2-like errors of failing to recommend a superior new treatment (making d_0 when $\delta > 0$) are similarly thought to increase in seriousness in direct proportion to the degree, δ, of superiority involved. For any absolute difference $\delta_0 = |\delta|$, recommendation of an inferior new treatment with $\delta = -\delta_0$ is considered k times as serious as the corresponding failure to recommend a superior new treatment with $\delta = \delta_0$. In the averaging of risks it is desired to weight risks symmetrically at $\delta = \pm\delta_0$ for all possible differences $\delta_0 \geqq 0$, with weights decreasing with respect to δ_0 as given by a normal density for $\tau = \delta/(2\sigma^2/r)^{\frac{1}{2}}$ with mean zero and variance γ^2. A minimum-average-risk rule is required which would be invariant with respect to any changes of scale or location in the observations.

Because of sufficiency and invariance considerations the required rule can be restricted to depend on the observations through only the t ratio

$$t = (\bar{x}_1 - \bar{x}_2)/(2s^2/r)^{\frac{1}{2}}, \tag{2.27}$$

where $\bar{x}_1$ and $\bar{x}_2$ are the respective sample means and s^2 is the pooled within-sample variance estimate

$$s^2 = \sum_{i=1}^{2} \sum_{j=1}^{r} (x_{ij} - \bar{x}_i)^2/2(r - 1). \tag{2.28}$$

The required rule is then given by (2.26) where $t_* = t_*(k, v, \gamma^2)$ with $v = 2(r - 1)$.

In practice a similar problem could well arise in which it is desired to use a non-zero mean μ_δ for the prior density. With infinite error degrees of freedom ($v = \infty$) the significant t ratio for such an asymmetric problem can be shown to be given by subtracting a correction of $\mu_\delta/(2\sigma^2/r)^{\frac{1}{2}}$ from the corresponding value in Table 1. For v finite the derivation is more difficult. The use of a similar correction of $\mu_\delta/(2s^2/r)^{\frac{1}{2}}$ would no doubt suffice however for practical purposes except for very small values of v. Since the extensions of this problem in the later sections concern only the symmetric case, a more detailed treatment of the asymmetric case will not be taken up here.

From the roles they play the parameters k, v, and γ^2 determining the minimum-average-risk significant t values may be usefully termed the *loss* or *error seriousness ratio*, the *error degrees of freedom* and the *risk-weighting variance ratio* respectively. Before going on it is of interest to note in Table 1 that a loss ratio of 100 ($\log k = 2$) infinite error degrees of freedom ($v = \infty$), and a risk weighting variance ratio of 3 ($\gamma^2 = 3$) give a t_* of 1.987 close to that 1.960 of a .025 level test of $H_0: \tau \leqq 0$.

3. A related three-decision Student-*t* problem. Given a similar observed t value, a problem related to that of Section 2 is one of choosing between the three common decisions

$$d_0\text{:decide } |\tau| \leqq \Delta, \qquad d_1\text{:decide } \tau > \Delta \quad \text{and} \quad d_2\text{:decide } \tau < -\Delta, \tag{3.1}$$

where, as before, Δ is some unspecified positive boundary value. In the language of the experimenter d_0, d_1 and d_2 are the decisions that τ is not significantly different from zero, that τ is significantly greater than zero and that τ is significantly less than zero, respectively.

Our second result is a Bayes rule for this problem with respect to a similar linear loss function

$$\begin{aligned} L_0^{(2)}(\tau) &= L^{(2)}(\tau, d_0) = \begin{cases} 0, & \tau = 0, \\ c_0\,|\tau|, & \tau \neq 0, \end{cases} \\ L_1^{(2)}(\tau) &= L^{(2)}(\tau, d_1) = \begin{cases} c_1\,|\tau|, & \tau \leqq 0, \\ 0, & \tau > 0, \end{cases} \\ L_2^{(2)}(\tau) &= L^{(2)}(\tau, d_2) = \begin{cases} 0, & \tau < 0, \\ c_1\tau, & \tau \geqq 0, \end{cases} \end{aligned} \tag{3.2}$$

where c_0 and c_1 are positive constants such that $c_1 - c_0 > c_0$ and with respect to the same normal prior density (2.4) for τ. The rule is

$$\boldsymbol{\phi}_*^2(t) = (\phi_{0*}^2(t)\ \phi_{1*}^2(t)\ \phi_{2*}^2(t)) = \begin{cases} (1\,0\,0), & |t| < t_*, \\ (0\,1\,0), & t > t_*, \\ (0\,0\,1), & t < -t_*, \end{cases} \tag{3.3}$$

where the significant t ratio $t_* = t_*(k, v, \gamma^2)$ is the same as that of the previous section with the loss ratio now given by $k = (c_1/c_0) - 1$, and where $\phi_i^{(2)}(t) = 0$ or 1 denotes the not making or making of the decision d_i, $i = 0, 1, 2$.

This result can be obtained as follows: First the three-decision subset system

$$\omega_0: |\tau| \leqq \Delta, \qquad \omega_1: \tau > \Delta, \qquad \omega_2: \tau < -\Delta, \tag{3.4}$$

can be expressed as the restricted product (the full product less empty intersections, see Lehmann [7]) of two component two-decision subset systems like that of the previous problem in Section 2, namely

$$\begin{aligned} &\text{Component system for } +\tau\text{: } \omega_0^+\text{: } \tau \leqq \Delta, \qquad \omega_1^+\text{: } \tau > \Delta, \\ &\text{Component system for } -\tau\text{: } \omega_0^-\text{: } -\tau \leqq \Delta, \qquad \omega_1^-\text{: } -\tau > \Delta. \end{aligned} \tag{3.5}$$

Thus

$$\omega_0 = \omega_0^+ \cap \omega_0^-, \; \omega_1 = \omega_1^+ \cap \omega_0^-, \; \omega_2 = \omega_0^+ \cap \omega_1^-. \tag{3.6}$$

The intersection $\omega_1^+ \cap \omega_1^-$ is excluded since it is empty. Put in other words, each of the main decisions is equivalent to two joint component decisions

$$d_0 \text{ to } d_0^+ \text{ with } d_0^-, \quad d_1 \text{ to } d_1^+ \text{ with } d_0^- \quad \text{and} \quad d_2 \text{ to } d_0^+ \text{ with } d_1^-; \tag{3.7}$$

the joint decision d_1^+ with d_1^- is excluded since it has mutually incompatible components; d_i^α is the decision $\tau \, \varepsilon \, \omega_i^\alpha$; $\alpha = +, -$; $i = 0, 1$.

Second, by putting $k_1 = c_1 - c_0$ and $k_0 = c_0$ the losses for the main decisions can be expressed as the sums of losses for its component decisions as given by the two-decision loss function (2.3) in the previous section. Demonstrating this

$$\begin{aligned} L_0(\tau) + L_0(-\tau) &= \begin{Bmatrix} 0 + k_0(-\tau) = c_0\,|\tau|, & \tau < 0 \\ 0 + 0 = 0, & \tau = 0 \\ k_0\tau + 0 = c_0\,|\tau|, & \tau > 0 \end{Bmatrix} = L_0^{(2)}(\tau), \\ L_1(\tau) + L_0(-\tau) &= \begin{Bmatrix} k_1\,|\tau| + k_0(-\tau) = c_1\,|\tau|, & \tau < 0 \\ k_1\,|\tau| + 0 = c_1\,|\tau|, & \tau = 0 \\ 0 + 0 = 0, & \tau > 0 \end{Bmatrix} = L_1^{(2)}(\tau), \\ L_0(\tau) + L_1(-\tau) &= \begin{Bmatrix} 0 + 0 = 0, & \tau < 0 \\ 0 + k_1\,|-\tau| = c_1\tau, & \tau = 0 \\ k_0\tau + k_1\,|-\tau| = c_1\tau, & \tau > 0 \end{Bmatrix} = L_2^{(2)}(\tau). \end{aligned} \tag{3.8}$$

Next, any rule $\boldsymbol{\phi}^{(2)}(t)$ for the three-decision problem can also be expressed in terms of two component two-decision rules. For this purpose it is convenient to first re-express the two-decision function $\phi(t)$ in the two-element vector form

$$\boldsymbol{\phi}(t) = (\phi_0(t) \; \phi_1(t)), \tag{3.9}$$

where $\phi_0(t) = 1 - \phi(t)$ and $\phi_1(t) = \phi(t)$. In this form, for example, the Bayes rule $\phi_*(t)$ of the previous section appears as

$$\boldsymbol{\phi}_*(t) = \begin{cases} (0\;1), & t < t_*, \\ (1\;0), & t > t_*. \end{cases} \tag{3.10}$$

With this vector notation (3.9) for the two-decision function and the components (3.7) of the three decisions in mind we can write

$$\begin{aligned}\boldsymbol{\phi}^{(2)}(t) &= (\phi_0^{(2)}(t) \quad \phi_1^{(2)}(t) \quad \phi_2^{(2)}(t)) \\ &= (\phi_0^+(t)\phi_0^-(t) \quad \phi_1^+(t)\phi_0^-(t) \quad \phi_0^+(t)\phi_1^-(t)),\end{aligned} \tag{3.11}$$

where $\phi_i^\alpha(t) = 0$ or 1 denotes the not making or making of decision d_i^α; $\alpha = +, -$; $i = 0, 1$.

Now, working with the average risk of $\boldsymbol{\phi}^{(2)}(t)$ we have

$$\begin{aligned}A(\xi, \boldsymbol{\phi}^{(2)}) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sum_{i=0}^{2} L_i^{(2)}(\tau)\phi_i^{(2)}(t)f(t \mid \tau)\, dt\, \xi(\tau)\, d\tau \\ &= \int_{-\infty}^{-\infty}\int_{-\infty}^{\infty} \sum_{i=0}^{1}\sum_{j=0}^{1} (L_i(\tau) + L_j(-\tau))\phi_i^+(t)\phi_j^-(t)f(t \mid \tau)\, dt\, \xi(\tau)\, dt,\end{aligned} \tag{3.12}$$

provided that the condition

$$\phi_1^+(t)\phi_1^-(t) = 0, \qquad -\infty < t < \infty, \tag{3.13}$$

is satisfiied. (Following Lehmann [7] this may be termed a compatibility condition since, if it were not satisfied, the component rules would give incompatible decisions). Assuming pro tempore that this is satisfied, the average risk readily reduces to

$$A(\xi, \boldsymbol{\phi}^{(2)}) = A(\xi, \boldsymbol{\phi}^+(t)) + A(\xi, \boldsymbol{\phi}^-(t)), \tag{3.14}$$

where the component average risks $A(\xi, \boldsymbol{\phi}^\alpha(t))$, $\alpha = +, -$, are the same functions of t and $-t$ as was $A(\xi, \phi) = A(\xi, \phi(t))$ of t in Section 2. To minimize (3.14) it is sufficient to minimize the component average risks separately subject still, of course, to the compatibility condition (3.13). From the result of the previous section, as expressed in (3.10), the component solutions are

$$\boldsymbol{\phi}_*^+(t) = \begin{cases}(1, 0), & t < t_* \\ (0, 1), & t > t_*\end{cases} \quad \text{and} \quad \phi_*^-(t) = \begin{cases}(1\ 0), & -t < t_* \\ (0\ 1), & -t > t_*\end{cases}. \tag{3.15}$$

Since $\phi_{1*}^+(t) = 1$ only in $t > t_*$, $\phi_{0*}^+(t) = 1$ only in $-t > t_* = t < -t_*$, and since t_* is positive (from $k_1 > k_0$) we do have $\phi_{1*}^+(t)\phi_{1*}^-(t) = 0$ for all t, that is, the solutions are compatible. Thus the required Bayes rule is given by

$$\begin{aligned}\boldsymbol{\phi}_*^{(2)}(t) &= (\phi_{0*}^+(t) \quad \phi_{0*}^-(t)\phi_{1*}^+(t)\phi_{0*}^-(t) \quad \phi_{0*}^+(t)\phi_{1*}^-(t)) \\ &= \begin{cases}(1\ 0\ 0), & (t < t_*)(t > -t_*) = |t| < t_*, \\ (0\ 1\ 0), & (t > t_*)(t > -t_*) = t > t_*, \\ (0\ 0\ 1), & (t < t_*)(t < -t_*) = t < -t_*,\end{cases}\end{aligned} \tag{3.16}$$

as was to be shown.

EXAMPLE 2. Suppose that two samples of yields have been observed as in Example 1 except that now they are for two new treatments. It is required to decide whether the first can be recommended as the superior (in yield), whether the second can, or whether to withhold recommendations on both. Errors of making a wrong recommendation or of failing to make an appropriate recommendation are again scorable in direct proportion to the degree of inferiority or superiority involved respectively. For any difference that may exist the error of a wrong recommendation is considered c times as serious as that of just failing to make the right recommendation. The requirements for weighting of risks and invariance are the same. The required rule is then given by (3.16) where t is as obtained (2.27) in Example 1, $t_* = t_*(k, v, \gamma^2)$ with $k = c - 1$ and $v = 2(r - 1)$. The subtraction of one from the loss ratio c will be trivial and unnecessary in most practical situations with c not small. The need for it here and not in Example 1, it may be said, comes from the fact that a wrong recommendation now includes implicitly a failure to make an appropriate recommendation as well.

4. A symmetric multiple comparisons problem. Given $N = \frac{1}{2}n(n - 1)$ t statistics of the form

$$t_{pq} = (\hat{\mu}_p - \hat{\mu}_q)/s_{\hat{\mu}_p - \hat{\mu}_q}, \qquad pq \; \varepsilon \; N, \tag{4.1}$$

with non-centrality parameters of the form

$$\tau_{pq} = (\mu_p - \mu_q)/\sigma_{\hat{\mu}_p - \hat{\mu}_q}, \qquad pq \; \varepsilon \; N, \tag{4.2}$$

where N is used for convenience to denote the set of pairs

$$\{1, 2; 1, 3; \cdots ; (n - 1), n\}$$

as well as its size, a common multiple comparisons problem is that of choosing between the three decisions

$$\overset{0}{d}_{pq}: \tau_{pq} \; \varepsilon \; \overset{0}{\omega}_{pq}, \qquad \overset{1}{d}_{pq}: \tau_{pq} \; \varepsilon \; \overset{1}{\omega}_{pq}, \qquad \overset{2}{d}_{pq}: \tau_{pq} \; \varepsilon \; \overset{2}{\omega}_{pq}, \tag{4.3}$$

simultaneously for all $pq \; \varepsilon \; N$, where the subsets are of the previous (3.1) form

$$\overset{0}{\omega}_{pq}: |\tau_{pq}| \leqq \Delta, \qquad \overset{1}{\omega}_{pq}: \tau_{pq} > \Delta, \qquad \overset{2}{\omega}_{pq}: \tau_{pq} < -\Delta. \tag{4.4}$$

The joint density of the t_{pq}'s is the one that would result, for example, from the common assumptions (a) $\hat{\mu}_1, \cdots, \hat{\mu}_n$ are normal independent variables with means $\mu_1, \cdots, \mu_n$ and the same but unknown variance $\sigma^2_{\hat{\mu}}$ and (b) $s_{\hat{\mu}}$ is an estimator of $\sigma_{\hat{\mu}}$ with v degrees of freedom such that $u = s_{\hat{\mu}}/\sigma_{\hat{\mu}}$ has the χ^2-related density (2.13) independently of $\hat{\mu}_1, \cdots, \hat{\mu}_n$, from which $\sigma_{\hat{\mu}_p - \hat{\mu}_q} = 2^{\frac{1}{2}}\sigma_{\hat{\mu}}$ and $s_{\hat{\mu}_p - \hat{\mu}_q} = 2^{\frac{1}{2}}s_{\hat{\mu}}$. If we put $\mathbf{y}$ for any vector of $(n - 1)$ orthogonal normalized comparisons among the estimates $\hat{\mu}_i$, that is, $\mathbf{y} = \mathbf{A}\hat{\boldsymbol{\mu}}$ where $\mathbf{AA}' = I_{n-1}$, $\mathbf{Aj} = \mathbf{0}$ ($\mathbf{j}$ being a vector of ones), and $\hat{\boldsymbol{\mu}} = (\hat{\mu}_1 \cdots \hat{\mu}_n)'$, then the density of the t_{pq}'s can be represented conveniently in terms of that of the $(n - 1)$-element t vector

$$\mathbf{t} = \mathbf{y}/s_{\hat{\mu}} \tag{4.5}$$

depending on the corresponding $(n - 1)$-element non-centrality parameter vector

$$\boldsymbol{\tau} = \boldsymbol{\eta}/\sigma_{\hat{\mu}} \tag{4.6}$$

where $\boldsymbol{\tau} = \mathbf{A}\boldsymbol{\mu}$ and $\boldsymbol{\mu} = (\mu_1 \cdots \mu_n)'$. This can readily be expressed as

$$f_n(\mathbf{t} \mid \boldsymbol{\tau}, v) = \int_0^{\infty} \exp \frac{\{-\frac{1}{2}(u\mathbf{t} - \boldsymbol{\tau})'(u\mathbf{t} - \boldsymbol{\tau})\}}{(2\pi)^{\frac{1}{2}(n-1)}} u^{n-1}\psi(u \mid v)\, du, \qquad -\infty < \mathbf{t} < \infty \tag{4.7}$$

where $\psi(u \mid v)$ is the χ^2-related density (2.13).

Our main result is a Bayes rule for this problem with respect to generalizations of the linear additive loss functions and normal prior density used in the previous sections.

More explicitly, the subsets $\omega_0, \omega_1, \cdots, \omega_{M-1}$ of the multiple comparisons problem are the non-empty intersections

$$\omega_i = \bigcap_{pq \varepsilon N} \omega_{pq}^{j_{pq}}, \qquad j_{pq} = 0, 1 \text{ or } 2, \qquad i = 0, \cdots, M-1, \tag{4.8}$$

of the subsets of all the component three-decision problems involved. The decision system consists of the M corresponding decisions

$$d_i{:}\, \boldsymbol{\tau} \varepsilon \omega_i, \qquad i = 0, \cdots, M-1. \tag{4.9}$$

For example, thinking of the component subset systems (4.4) in the form

$$\omega_{pq}^{0}{:}\ (|\mu_p - \mu_q| \leqq \Delta'), \qquad \omega_{pq}^{2}{:}\ (\mu_q < \mu_p - \Delta'), \qquad \omega_{pq}^{2}{:}\ (\mu_p < \mu_q - \Delta'), \tag{4.10}$$

(where $\Delta' = \Delta 2^{\frac{1}{2}}\sigma_{\hat{\mu}}$) and using the corresponding more graphic notation

$$(\underline{p, q}), (q, p), (p, q), \tag{4.11}$$

in place of $\omega_{pq}^{0}, \omega_{pq}^{1}, \omega_{pq}^{2}$ respectively, the $M = 19$ multiple comparisons subsets

(4.12)

		$(\underline{1, 2})$	$(2, 1)$	$(1, 2)$
	$(\underline{2, 3})$	$\omega_0 = (\underline{1, 2, 3})$	$\omega_1 = (\underline{2, 3, 1})$	$\omega_2 = (\underline{1, 3, 2})$
$(\underline{1, 3})$	$(3, 2)$	$\omega_3 = (\underline{3, 1, 2})$	$\cdot$	$\omega_4 = (\underline{1, 3}, 2)$
	$(2, 3)$	$\omega_5 = (\underline{2, 1, 3})$	$\omega_6 = (2, \underline{1, 3})$	$\cdot$
	$(\underline{2, 3})$	$\omega_7 = (\underline{3, 2, 1})$	$\omega_8 = (\underline{3, 2}, 1)$	$\cdot$
$(3, 1)$	$(3, 2)$	$\omega_9 = (3, \underline{1, 2})$	$\omega_{10} = (3, 2, 1)$	$\omega_{11} = (3, 1, 2)$
	$(2, 3)$		$\omega_{12} = (2, 3, 1)$	$\cdot$
	$(\underline{2, 3})$	$\omega_{13} = (\underline{1, 2, 3})$	$\cdot$	$\omega_{14} = (1, \underline{2, 3})$
$(1, 3)$	$(3, 2)$	$\cdot$	$\cdot$	$\omega_{15} = (1, 3, 2)$
	$(2, 3)$	$\omega_{16} = (\underline{1, 2}, 3)$	$\omega_{17} = (2, 1, 3)$	$\omega_{18} = (1, 2, 3)$

in the case $n = 3$ may be developed as in (4.12). In general, following Duncan [3], the notation $(\underline{i, j, k,} \cdots)$ may be used to denote subsets in which the corresponding means $\mu_i, \mu_j, \mu_k, \cdots$ are ranked in significantly ascending order from left to right (i.e. with differences $|\delta| > \Delta'$) except that subscripts underscored by a common line denote pairs of means for which the difference is not significant (i.e. $|\delta| \leqq \Delta'$). Thus $\omega_2 = (\underline{1, \underline{3}, 2})$ is the subset in which μ_1 is significantly less than μ_2 but μ_1 and μ_2 each do not differ significantly from μ_3. The remaining $2^3 - 19 = 8$ intersections, e.g., $(2, 1)(\underline{1, 3})(3, 2)$, of the component subsets are not included in the multiple comparisons system since they are empty.

Choosing the elements of $\boldsymbol{\tau}$ as $\tau_1 = (\mu_1 - \mu_2)/2^{\frac{1}{2}}\sigma_{\bar{\mu}}$ for abscissa and $\tau_2 = (\mu_1 + \mu_2 - 2\mu_3)/6^{\frac{1}{2}}\sigma_{\bar{\mu}}$ for ordinate the parameter subsets, in the case $n = 3$, may be represented as shown in Figure 2 of [3], where the vertical lines are $\tau_{12} = \pm\Delta$, the lines from top left to bottom right are $\tau_{13} = \pm\Delta$ and those from bottom left to top right are $\tau_{23} = \pm\Delta$.

Referring back for a moment, the need for the definition (2.1) instead of (2.2) for the decisions of our initial subcomponent problem can now be seen more clearly in the formation of the subsets (4.12). If (2.2) were used the six subsets ω_1, ω_2, ω_3, ω_5, ω_7 and ω_{13} of the form $(\underline{i, \underline{j}, k})$ would be eliminated. These however are useful members of the system, hence the need for some such definition as (2.1) to retain them.

The size M of the multiple comparisons subset system increases rapidly with n the number of means involved. In the next case $n = 4$, for example, the numbers of the various forms of subsets, using the same notation, are

$$
\begin{aligned}
(4.13)\qquad & (\underline{1, 2, 3, 4}) \cdots 1,\ (\underline{i, j, k}, l) \cdots 4,\ (i, \underline{j, k, l}) \cdots 4,\ (\underline{i, j}, k, l) \\
& \cdots 12,\ (i, \underline{j, k}, l) \cdots 12,\ (i, j, \underline{k, l}) \cdots 12,\ (i, j, k, l) \cdots 24, \\
& (\underline{i, \underline{j}, k}, l) \cdots 12,\ (\underline{i, j, \underline{k}, l}) \cdots 12,\ (\underline{i, \underline{j}, k, l}) \cdots 12,\ (\underline{i, \underline{j}, k}, l) \\
& \cdots 24,\ (\underline{i, j}, \underline{k, l}) \cdots 6,\ (i, \underline{j, \underline{k}, l}) \cdots 24,\ (\underline{i, \underline{j}, k, l}) \cdots 24,
\end{aligned}
$$

making $M = 183$ in all.

The losses are defined as the sum of the losses (3.2) for each of the component decisions involved; that is

$$
(4.14)\qquad L_i^{(n)}(\boldsymbol{\tau}) = L^{(n)}(\boldsymbol{\tau}, d_i) = \sum_{pq \in N} L_{j_{pq}^i}^{(2)}(\tau_{pq}), \qquad j_{pq}^i = 0, 1 \text{ or } 2;\ i = 0, \cdots, M-1.
$$

Thus, suppose for example in the case $n = 3$, $\boldsymbol{\mu}/2^{\frac{1}{2}}\sigma = (\mu_1\mu_2\mu_3)'/2^{\frac{1}{2}}\sigma$ represents the expected standardized yields of three manurial treatments on a particular agricultural crop. Then the loss $L_{14}^{(3)}(\boldsymbol{\tau})$ at $\boldsymbol{\mu}/2^{\frac{1}{2}}\sigma = (10, 12, 8)'$ incurred by the decision $d_{14}: \boldsymbol{\tau}\ \varepsilon\ (1, \underline{2, 3})$ is

$$
\begin{aligned}
L_{14}^{(3)}(\boldsymbol{\tau}) &= L_2^{(2)}(\tau_{12}) + L_2^{(2)}(\tau_{13}) + L_0^{(2)}(\tau_{23}) \\
(4.15)\qquad &= L_2^{(2)}(-2) + L_2^{(2)}(2) + L_0^{(2)}(4) \\
&= 0 + 2c_1 + 4c_0 = 2k_1 + 6k_0 .
\end{aligned}
$$

The third contribution $4c_0 = 4k_0$ enters, it may be said, because the decision d_{14} has failed to recognize the 4-unit superiority of treatment two over treatment three. The second contribution $2c_1 = 2k_1 + 2k_0$ enters, because, in similar terms, d_{14} not only fails to recognize the 2-unit superiority of treatment one over treatment three, incurring a loss of $2k_0$, it also commits the more serious error of ranking treatment three above treatment one, incurring an additional loss of $2k_1$. No loss is incurred by the first component decision because d_{14} ranks treatment two correctly above treatment one.

The prior density for averaging risks over all points in the $(n-1)$-dimensional space for $\boldsymbol{\tau}$ is the simple normal density function

$$\xi_n(\boldsymbol{\tau} \mid \gamma^2) = (2\pi\gamma^2)^{-\frac{1}{2}(n-1)} e^{-\boldsymbol{\tau}'\boldsymbol{\tau}/2\gamma^2}, \qquad -\infty < \boldsymbol{\tau} < \infty. \tag{4.16}$$

This is one which would result, for example, from assuming that the means $\mu_1, \cdots, \mu_n$ have independent normal prior distributions with the same mean θ and same variance $\sigma_\mu^2 = \gamma^2\sigma_{\hat{\mu}}^2$.

From the additive losses assumption it follows as before that the average risk for any decision rule $\boldsymbol{\phi}^{(n)}(\mathbf{t}) = (\phi_0^{(n)}(\mathbf{t}) \cdots \phi_{M-1}^{(n)}(\mathbf{t}))$ may be expressed as the sum of average risks for component three-decision rules

$$\boldsymbol{\phi}^{pq}(\mathbf{t}) = (\phi_0^{pq}(\mathbf{t})\ \phi_1^{pq}(\mathbf{t})\ \phi_2^{pq}(\mathbf{t}))$$

provided again that the component rules are compatible. The steps may be written

$$\begin{aligned} A(\xi_n, \boldsymbol{\phi}^{(n)}) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sum_{i=0}^{M-1} L_i^{(n)}(\boldsymbol{\tau})\phi_i^{(n)}(\mathbf{t})f_n(\mathbf{t} \mid \boldsymbol{\tau})d\mathbf{t}\, \xi_n(\boldsymbol{\tau})\, d\boldsymbol{\tau} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sum_{j_1=0}^{2} \cdots \sum_{j_N=0}^{2} [L_{j_1}^{(2)}(\tau_{12}) + \cdots + L_{j_N}^{(2)}(\tau_{(n-1)n})] \\ &\qquad \cdot \phi_{j_1}^{12}(\mathbf{t}) \cdots \phi_{j_N}^{(n-1)n}(\mathbf{t})\, f_n(\mathbf{t} \mid \boldsymbol{\tau})\, d\mathbf{t}\, \xi_n(\boldsymbol{\tau})\, d\tau \\ &= \sum_{pq\varepsilon N} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sum_{j=0}^{2} L_j^{(2)}(\tau_{pq})\phi_j^{pq}(\mathbf{t})\, f_n(\mathbf{t} \mid \boldsymbol{\tau})\, d\mathbf{t}\, \xi_n(\boldsymbol{\tau})\, d\boldsymbol{\tau} \\ &= \sum_{pq\varepsilon N} A(\xi_n\, \phi^{pq}(\mathbf{t})). \end{aligned} \tag{4.17}$$

The compatibility condition may be written as

$$\prod_{pq\varepsilon N} \phi_{j_{pq}}^{pq}(\mathbf{t}) = 0, \quad j_{pq} = 0, 1 \text{ or } 2, -\infty < \mathbf{t} < \infty, \tag{4.18}$$

for all products leading to incompatible decisions not included in

$$\{d_i\,;\, i = 0, \cdots, M-1\}$$

and is required in proceeding from the first to the second line of (4.17).

It then follows as before that the Bayes rule

$$\phi_*^{(n)}(\mathbf{t}) = (\phi_{0*}^{(n)}(\mathbf{t}) \cdots \phi_{(M-1)*}^{(n)}(\mathbf{t})), \tag{4.19}$$

for the multiple comparisons problem is formed by the products

$$\phi_{i*}^{(n)}(\mathbf{t}) = \prod_{pq\varepsilon N} \phi_{j_{pq}^i *}^{pq}(\mathbf{t}), \qquad j_{pq}^i = 0, 1 \text{ or } 2, \tag{4.20}$$

of the elements of the Bayes rules $\boldsymbol{\phi}_*^{pq}(\mathbf{t})$ minimizing

$$A(\xi_n, \boldsymbol{\phi}^{pq}(\mathbf{t})), \qquad pq \, \varepsilon \, N, \tag{4.21}$$

provided these are compatible.

The first step in deriving $\boldsymbol{\phi}_*^{pq}(\mathbf{t})$ is practically identical with that ((3.12) to (3.14)) in Section 3. Thus

$$\boldsymbol{\phi}_*^{pq}(\mathbf{t}) = (\phi_{0*}^{pq+}(\mathbf{t})\phi_{0*}^{pq-}(\mathbf{t}) \quad \phi_{1*}^{pq+}(\mathbf{t})\phi_{0*}^{pq-}(\mathbf{t}) \quad \phi_{0*}^{pq+}(\mathbf{t})\phi_{1*}^{pq-}(\mathbf{t})) \tag{4.22}$$

where, dropping the superscripts $pq+$ or $pq-$, $\boldsymbol{\phi}_*(\mathbf{t}) = (\phi_{0*}(\mathbf{t}) \quad \phi_{1*}(\mathbf{t}))$ now minimizes a subcomponent average risk of the form

$$A(\xi_n \, \boldsymbol{\phi}(\mathbf{t})) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sum_{i=0}^{1} L_i(\tau_{pq})\phi_i(\mathbf{t})f_n(\mathbf{t} \mid \boldsymbol{\tau}) \, d\mathbf{t} \, \xi_n(\boldsymbol{\tau}) \, d\boldsymbol{\tau}. \tag{4.23}$$

The elements of $\boldsymbol{\tau}$ may be chosen so that τ_{pq} is the first element 1_1 of $\boldsymbol{\tau}$. The compatibility condition

$$\phi_{1*}^{pq+}(\mathbf{t})\phi_{1*}^{pq-}(\mathbf{t}) = 0 \tag{4.24}$$

must again be met.

The work of minimizing (4.23) follows closely that of minimizing (2.6) in Section 2 except that now the sample and parameter spaces have additional dimensions, $(n - 2)$ each. Dropping the subscript pq from τ_{pq} the steps follow through with obvious changes till we get to

$$\begin{aligned} &h_3^{(n)}(\tau, \mathbf{t}, v) \\ &\propto \tau \int_{-\infty}^{\infty}\int_0^{\infty} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n=1} [(ut_i - \tau_i)^2 + \tau_i^2/\gamma^2]\right\} u^{n-1}\psi(u \mid v) \, du \, d\tau_2 \end{aligned} \tag{4.25}$$

where the first integration is with respect to $\boldsymbol{\tau}_2$ defined as the last $n - 2$ elements of $\boldsymbol{\tau} \doteq (t \, \boldsymbol{\tau}_2')'$. This appears in place of $h_3(\tau, t)$ as in (2.14) before which may now be denoted by $h_3^{(2)}(\tau, \mathbf{t}, v)$. On integrating with respect to $\boldsymbol{\tau}_2$ we get

$$\begin{aligned} h_3^{(n)}(\tau, \mathbf{t}, v) &\propto \tau \int_0^{\infty} \exp\left\{ -\frac{1}{2}\left[(ut - \tau)^2 + \tau^2/\gamma^2 \right.\right. \\ &\qquad \left.\left. - \sum_{i=2}^{n-1} u^2 t_i^2/(1 + \gamma^2)\right]\right\} u^{n-1}u^{v-1}e^{-\frac{1}{2}vu^2} \, du \\ &\propto \tau \int_0^{\infty} \exp\{-\tfrac{1}{2}[(u't' - \tau)^2 + \tau^2/\gamma^2]\} u'u'^{(v'-1)}e^{-\frac{1}{2}v'u'^2} \, du', \end{aligned} \tag{4.26}$$

where $t = t_1$ is the first element of $\boldsymbol{\tau}$ and is thus $t = t_{pq}$,

$$u' = u/R, \qquad t' = Rt, \qquad R^2 = v' \Big/ \left[v + \sum_{i=2}^{n-1} t_i^2/(1 + \gamma^2)\right]. \tag{4 27}$$

and $v' = v + n - 2$. Thus

$$h_3^{(n)}(\tau, \mathbf{t}, v) \propto \tau \int_0^\infty \exp\{-\tfrac{1}{2}[(ut' - \tau)^2 + \tau^2/\gamma^2]\}u\psi(u \mid v')\, du \tag{4.28}$$
$$= h_3^{(2)}(\tau, t', v').$$

Making a direct application of the remaining derivation in Section 2 it now follows that

$$\phi_*(t) = \begin{cases} (1, 0), & (t' < t_*) = (t < t_*/R) \\ (0, 1), & (t' > t_*) = (t > t_*/R) \end{cases} \tag{4.29}$$

where $t_* = t_*(k, v', \gamma^2)$ is the same significant t value as before except that its degrees of freedom are now $v' = v + n - 2$.

Since $t_*/R > 0$, the compatibility condition (4.24) is met and applying (4.22) the component average risk (4.21) is minimized by

$$\phi_*^{pq}(\mathbf{t}) = \begin{cases} (1\ 0\ 0), & (t'_{pq} < t_*) = (t_{pq} < t_*/R), \\ (0\ 1\ 0), & (t'_{pq} > t_*) = (t_{pq} > t_*/R), \\ (0\ 0\ 1), & (t'_{pq} < -t_*) = (t_{pq} < -t_*/R), \end{cases} \tag{4.30}$$

where $t'_{pq} = Rt_{pq}$ for all $pq \varepsilon N$. Again since $t_*/R > 0$, the preceding compatibility conditions (4.18) are met and applying (4.19) the Bayes rule for the multiple comparisons problem is given by the simultaneous application (4.20) of all $N = n(n - 1)/2$ of the three-decision Bayes rules (4.30).

EXAMPLE 3. Suppose that n samples of yields like those of Example 2 have been obtained for n new treatments. For each and every pair (a, b) of treatments it is required to decide whether a can be recommended as the superior, whether b can or whether to withhold recommendations on both. Losses are scorable with respect to each pair of treatments as in Example 2, the loss ratio c being the same for all pairs, and are additive in giving the losses for each of the joint decisions to which they contribute. Risks are to be averaged with respect to a normal independent prior density for each of the means $\mu_1, \cdots, \mu_n$ each with the same mean and same variance $\gamma^2\sigma^2/r$. An invariant rule is required as before.

Because of sufficiency and invariance considerations the required rule can be restricted to depend on the observations through only the t vector

$$\mathbf{t} = \mathbf{A}\bar{\mathbf{x}}/(s^2/r)^{\frac{1}{2}} \tag{4.31}$$

where A is as defined before for (4.5), $\bar{\mathbf{x}} = (\bar{x}_1, \cdots, \bar{x}_n)'$ is the vector of sample means and s^2 is the pooled within-sample variance estimate

$$s^2 = \sum_{i=1}^{n} \sum_{j=1}^{r} (x_{ij} - \bar{x}_i)^2/n(r - 1) \tag{4.32}$$

with $v = n(r - 1)$ degrees of freedom.

The required decision rule is then given by the simultaneous application (4.20) of (4.30) where $t_* = t_*(k, v', \gamma^2)$ with $k = c - 1$ and $v' = v + n - 2 =$

$n(r - 1) + n - 2$ and where t'_{pq} can be obtained by analysis-of-variance type steps as follow: Put S_t, S_{pq}, S'_{pq} and S_e for the treatment sum of squares, the sum of squares for the pq difference, the residual sum of squares for the pq difference and the error sum of squares

$$S_t = r \sum_{i=1}^{n} \bar{x}_i - C, \qquad S_{pq} = r(\bar{x}_p - \bar{x}_q)^2/2, \tag{4.33}$$

$$S'_{pq} = S_t - S_{pq}, \qquad S_e = \left(\sum_{i=1}^{n} \sum_{j=1}^{r} x_{ij}^2 - C\right) - S_t,$$

respectively, where C is the correction term $(\sum_{i=1}^{n} \sum_{j=1}^{r} x_{ij})^2/nr$. Let s'^2_{pq} denote the pooled estimate of σ^2 obtained as

$$s'^2_{pq} = [S_e + S'_{pq}/(1 + \gamma^2)]/v'. \tag{4.34}$$

Then t'_{pq} may be obtained as the square root of the variance ratio

$$t'^2_{pq} = S_{pq}/s'^2_{pq}, \tag{4.35}$$

t'_{pq} is given the same sign as $\bar{x}_p - \bar{x}_q$.

A more convenient rule for application can be obtained by expressing the inequalities $t'^2_{pq} \lessgtr t^2_*$ in the form $d^2_{pq} \gtrless d^2_*$ where d_* is a least significant value for the difference $d_{pq} = \bar{x}_p - \bar{x}_q$. From $t'^2_{pq} = t^2_*$ we get

$$t^2_* = S_{pq}/s'^2_{pq} = v'S_{pq}/[S_e + (S_t - S_{pq})/(1 + \gamma^2)]$$

$$t^2_*[S_e + (S_t - S_{pq})/(1 + \gamma^2)] = v'S_{pq} \tag{4.36}$$

$$S_{pq}[v' + t^2_*/(1 + \gamma^2)] = t^2_*[S_e + S_t/(1 + \gamma^2)].$$

But $S_{pq} = \frac{1}{2}rd^2_{pq}$, hence this gives $d^2_{pq} = d^2_*$ where

$$d_* = \left\{\frac{2}{r} t^2_*[S_e + S_t/(1 + \gamma^2)]/[v' + t^2_*/(1 + \gamma^2)]\right\}^{\frac{1}{2}} \tag{4.37}$$

From this and a check on signs it follows that the multiple comparisons Bayes rule is given by the simultaneous application of the rules

$$\boldsymbol{\phi}^{pq}_*(\mathbf{t}) = \begin{cases} (1\ 0\ 0), & |d_{pq}| < d_*, \\ (0\ 1\ 0), & d_{pq} > d_*, \\ (0\ 0\ 1), & d_{pq} < -d_*. \end{cases} \tag{4.38}$$

5. Discussion.

5.1. On the additional error degrees of freedom. The emergence of $t'_{pq} = d_{pq}/(2rs'^2)^{\frac{1}{2}}$ with $v' = v + n - 2$ degrees of freedom as the component test statistic in the multiple comparisons solution may be surprising at first but less so after due consideration. In giving $\mu_1, \cdots, \mu_n$ identical independent normal distributions with variance $\gamma^2\sigma^2/r$ for risk-weighting purposes, the residual sum of squares between treatments, $S_{pq} = r\sum_{i=2}^{n-1} y_i^2 = rs^2\sum_{i=2}^{n=1} t_i^2$ in Example 3 for instance, is given the distribution of $(1 + \gamma^2)\sigma^2\chi^2_{n-2}$. On this basis

$$s'^2_{pq} = [S_e + S_{pq}/(1 + \gamma^2)]/v' \tag{5.1}$$

becomes the appropriate estimator for σ^2 in place of $s^2 = S_e/v$ and the result is as might be expected. In practice a user might sometimes be reluctant to depend on the prior distribution assumption to this extent however, and might even want, see Subsection 5.3, to use S'_{pq}, or better S_t, to decide on an appropriate value $\hat{\gamma}^2$ for γ^2. In such cases it would seem good sense to use a modified rule based on t_{pq} instead of t'_{pq}. This would consist of simultaneous applications of

$$\phi_*^{pq}(\mathbf{t}) = \begin{cases} (1\,0\,0), & d_{pq} < d_*, \\ (0\,1\,0), & d_{pq} > d_*, \\ (0\,0\,1), & d_{pq} < -d_*, \end{cases} \tag{5.2}$$

where d_* is the least significant difference

$$d_* = (2s^2/r)^{\frac{1}{2}} t_*, \tag{5.3}$$

with $t_* = t_*(k, v, \hat{\gamma}^2)$ based on v degrees of freedom.

5.2. On the independence of the least significant difference and n. By far the most striking feature of the multiple comparisons Bayes rule is the practically complete lack of dependence of the least significant difference d_* on n, the number of means involved. (The dependence of d_* on n via the estimation of error as discussed in the previous subsection is relatively trivial in this context and decreases to zero as v increases to ∞.) This is a direct consequence of the additive losses assumption similar results of which have also been treated by Duncan [3] and Thompson [15] and, in a more general form and context, by Lehmann [7]. In the past, a rule of this type, with d_* not increasing with n, has been considered more or less unacceptable. The main basis of objection has been the rapid increase in its so-called n-treatment significance level (Duncan [1] and [3]) or its experimentwise error rate (Tukey [17] and [18]),

$$\alpha_n = P[\text{rejecting } H_n \mid H_n], \qquad H_n = H_n\colon \mu_1 \cdots = \mu_n, \tag{5.4}$$

with respect to n.

To illustrate, a non-increasing least significant difference of

$$d_* = 1.960\sqrt{2}\sigma_{\hat{\mu}} = 2.77\sigma_{\hat{\mu}}$$

in the case $v = \infty$ gives the experiment-wise error rates (found as upper-tail probabilities $P[q_n > 2.77]$ of the range q_n)

$$\alpha_2 = .0500, \quad \alpha_3 = .1223, \quad \alpha_4 = .2034, \cdots, \quad \alpha_{20} = .9183. \tag{5.5}$$

The possibility in this case of wrongly rejecting the homogeneity hypothesis for 20 means, for example, with a probability of 91.83 per cent, may at first appear to be unacceptably high. As a result, procedures have been proposed with increasing significant differences aimed at suppressing the rapid increases in α_n. These have varied considerably from rapidly increasing significant differences such as (in comparable cases, dropping the factor $\sigma_{\hat{\mu}}$)

$$2.77, 3.32, 3.63, \cdots, 5.01, \tag{5.6}$$

to more slowly increasing ones such as

$$2.77, 2.92, 3.02, \cdots, 3.47, \tag{5.7}$$

depending on the relative importance attached to experimentwise error rates and the degree to which they should be suppressed. The differences (5.6) termed honest significant differences by Tukey [17] [18] suppress all of the experimentwise error rates to .0500. The differences (5.7) proposed by the author [1] and [3] suppress them to less conservative so-called levels based on degrees of freedom

$$(5.8)\quad \alpha_2' = .0500, \qquad \alpha_3' = .0975, \qquad \alpha_4' = .1426, \cdots, \qquad \alpha_{20} = .6415,$$

obtained as $\alpha_n' = 1 - (1 - \alpha_2)^{n-1}$. A less conservative procedure yet is the one due to Fisher [5] that uses the same least significant difference for all n (2.77 in the case above) provided that first the homogeneity hypothesis H_n can be rejected by an F ratio test.

Now it appears that, in a Bayes sense, provided the losses are additive and other things (e.g., k and γ^2) are equal, the same least significant difference is optimum no matter how large the number n of treatments involved. (From Lehmann's work [7] it is clear that this would also apply under other optimality criteria such as, for example, minimax.) The high experimentwise error-rate of 91.83 per cent in the case quoted might well be worth tolerating for instance, because, it might be said, of the relatively low prior probability of the hypothesis H_{20} involved and its relative unimportance among so many others.

The inverse form of dependence of the least significant difference d_* on the risk-weighting variance ratio γ^2 may do much to reconcile its independence of n with at least some of the common almost instinctive urge to make it increase with n. In the case $v = \infty$ it is directly proportional to $(1 + 1/\gamma^2)^{\frac{1}{2}}$ and approximately so for smaller values of v. If, in the conduct of a large experiment, the treatments under study have a lower anticipated heterogeneity than those which would have been studied in a more limited experiment with fewer populations, a lower risk-weighting variance ratio γ^2 would be appropriate and hence would be a larger least significant difference. Such a situation could often arise in practice, and, if γ^2 is varying in an interval of small values this could make a substantial increase in the significant difference with an increase in n. On the other hand, however, the reverse situation could also arise. In selection experiments, for example, the treatments under consideration may be the top n performers as assessed by experiences in previous trials. Here, the larger the number of treatments it has been possible to include in an experiment, the larger will be the appropriate γ^2 and hence, the smaller the least significant difference.

5.3. A practical adaptation of the Bayes rule. In the complete absence of prior criteria for choosing γ^2, the user might sensibly, it would seem, obtain an estimate of it from the variance ratio

$$(5.9)\qquad F = \frac{1}{n-1}\sum_{i=1}^{n-1}\frac{y_i^2}{s_{\bar{\mu}}^2} \quad \left(= \frac{1}{n-1}\frac{S_t}{s^2} \text{ in Example 3}\right),$$

employed in the preliminary F test of Fisher's least-significant-difference procedure. Since the ratio of the corresponding expectations is $1 + \gamma^2$ he might for example put $\hat{\gamma}^2 = F - 1$, enter Table 1 with $\gamma^2 = \hat{\gamma}^2$, and use the simple direct rule as given (5.2) in Subsection 5.1. It is of considerable interest to note the closeness of Fisher's earlier procedure to this type of adaptation. Both rules depend on a preliminary inspection of the F ratio. In Fisher's rule a big or small F ratio leads to the use of an independently chosen least significant difference on a go-no basis. In the new rule a big or small F ratio leads to the use of a small or big least significant difference on a continuously related basis.

5.4 Concluding remarks. As presented the model is limited to a class of symmetric problems in which the loss and prior probabilities are invariant with respect to all $n!$ permutations of the means involved. Many problems in practice however are naturally taken to be symmetric in this way. Within this class, the assumptions of linear losses and normal prior densities would seem directly appropriate for some problems and useful at least as good approximations for many others.

From the given development it is fairly clear that similar rules for a wider class of less symmetrical problems can be obtained leading to the use of different significant t ratios for each of the component or even each of the subcomponent problems involved. Development and discussion of these are deferred to a further paper.

A most interesting point is one raised (in private correspondence) by Professor F. J. Anscombe following from the type of discussion in Subsection 5.3. In addition to providing an estimate of the prior variance γ^2 and therefore a means of rejecting an assumed value for this parameter, the data may provide evidence for reasonably rejecting other assumed features of the prior density. Further developments are needed for handling problems of this type.

In conclusion, it is worth repeating, the most important result discussed in Subsection 5.2 namely the independence of the least-significant-difference d_* and the number of components problems involved, depends only on the additivity assumption for the losses. It is independent of the form of the component loss functions and of prior density assumed. It appears further that the same result would follow even if the class of component problems were extended to include all contrasts among the means as considered by Scheffé [11]. Thus in a symmetric situation, for example, the same significant t ratio would be appropriate whether it be desired to test just one comparison chosen *a priori*, the set of all $\frac{1}{2}n(n-1)$ comparisons in the multiple comparisons problem or the set of all contrasts. The additivity-of-losses assumption on which this critically depends appears to be a reasonable one, and appropriate to many practical situations.

6. Computation of significant t ratios. In computing the values in Table 1 the ratios $g(y)/g(-y)$ in (2.22) may be simplified first to $g_1(y)/g_1(-y)$ where

$$(6.1)\quad g_1(y) = \begin{cases} (1-y^2)^{\frac{1}{2}} + y\sin^{-1} y + \pi y/2, & v = 1, \\ (1+y)^2 & v = 2, \\ 2(1-y^2)^{\frac{3}{2}} + 3y^2(1-y^2)^{\frac{1}{2}} + 3y\sin^{-1} y + 3\pi y/2, & v = 3, \\ (1+y)^3/(3+y), & v = 4, \\ (1+y)^4/(5+4y+y^2), & v = 6, \\ (1+y)^8/(429 + 1384y + 2063y^2 + 1776y^3 & \\ \qquad + 915y^4 + 264y^5 + 33y^6), & v = 14, \end{cases}$$

or to $g_2(z)/g_2(-z)$ where

$$g_2(z) = f(z) + zF(z), \qquad v = \infty, \tag{6.2}$$

and f and F are the standard normal density and cumulative distribution functions respectively. These follow readily from (2.23) except for (6.2) in the case $v = \infty$ which can be obtained as follows.

Putting $y = z/(v + z^2)^{\frac{1}{2}}$ and thus $(1 - y^2) = 1/(1 + z^2/v)$ and $dy = dz/[v^{\frac{1}{2}}(1 + z^2/v)^{\frac{3}{2}}]$ in $g(y)$ in (2.23) we get

$$g(y) = \frac{1}{(1 + z^2/v)^{v/2}} + \frac{z}{(1 + z^2/v)^{\frac{1}{2}}} \int_0^z \frac{du}{(1 + u^2/v)^{(v+1)/2}} + \frac{(\pi v)^{\frac{1}{2}}[(v - 2)/2]!\, z}{[(v - 1)/2]!\, 2(1 + z^2/v)^{\frac{1}{2}}}. \tag{6.3}$$

Recalling that the probability density function of the Student t distribution is

$$h(t \mid v) = \frac{[(v - 1)/2]!}{(\pi v)[(v - 2)/2]!} \cdot \frac{1}{(1 + t^2/v)^{(v+1)/2}} \tag{6.4}$$

we may write

$$\begin{aligned} g(y) &\propto (1 + z^2/v)^{\frac{1}{2}} h(z \mid v) + \frac{z[H(z \mid v) - \frac{1}{2}]}{(1 + z^2/v)^{\frac{1}{2}}} + \frac{z/2}{(1 + z^2/v)^{\frac{1}{2}}} \\ &\propto (1 + z^2/v) h(z \mid v) + zH(z \mid v) = g_3(z) \text{ say,} \end{aligned} \tag{6.5}$$

where $H(z \mid v)$ is the cumulative distribution $\int_{-\infty}^{z} h(t \mid v)\, dt$. Treating $g(-y)$ in the same way we reach the result

$$g(y)/g(-y) = g_3(z)/g_3(-z), \tag{6.6}$$

which reduces to $g_2(z)/g_2(-z)$ as $v \to \infty$.

Next, $y_*(k, v)$ or $z_*(k)$ is found as the solution of $g_1(y)/g_1(-y) = k$ or $g_2(z)/g_2(-z) = k$ respectively. Finally t_* is found as the positive square root of

$$t_*^2 = v/(\beta^{-2} y_*^{-2} - 1) \quad \text{or from} \quad t_* = z_*/\beta. \tag{6.7}$$

Acknowledgments. Help from Professor Harold Hotelling and other colleagues is gratefully acknowledged.

REFERENCES

[1] David B. Duncan, "A significance test for differences between ranked treatments in an analysis of variance," *Virginia J. Sci.*, Vol. 2 (1951), pp. 171–189.
[2] David B. Duncan, "On the properties of the multiple comparisons test," *Virginia J. Sci.*, Vol. 3 (1952), pp. 50–67.
[3] David B. Duncan, "Multiple range and multiple F tests," *Biometrics*, Vol. 11 (1955), pp. 1–42.
[4] Walter T. Federer, *Experimental Design*, MacMillan Co., New York, 1955.
[5] R. A. Fisher, *The Design of Experiments*, six eds. Oliver and Boyd, London (1935–1951).
[6] E. L. Lehmann, "Some principles of the theory of testing hypotheses," *Ann. Math. Stat.*, Vol. 21 (1950), pp. 1–26.

[7] E. L. Lehmann, "A theory of some multiple decision problems, I," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 1–25.
[8] E. L. Lehmann, "A theory of some multiple decision problems, II," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 547–572.
[9] Jerome C. R. Li, *Introduction to Statistical Inference*, Edwards Bros., Ann Arbor, 1957.
[10] D. Newman, "The distribution of the range in samples from a normal population, expressed in terms of an independent estimate of the standard deviation," *Biometrika*, Vol. 31 (1939), pp. 20–30.
[11] Henry Scheffé, "A method for judging all contrasts in the analysis of variance," *Biometrika*, Vol. 40 (1953), pp. 87–104.
[12] Henry Scheffé, "*The Analysis of Variance*", John Wiley and Sons, New York, 1959.
[13] G. W. Snedecor, *Statistical Methods*, 5th ed., Iowa State College Press, Ames, 1957.
[14] R. G. D. Steel and J. H. Torrie, "*Principles and Procedures of Statistics*," McGraw-Hill Book Co., New York, 1960.
[15] W. A. Thompson, "A note on ranking means," *Ann. Math. Stat.*, Abstract 13, Vol. 27 (1956).
[16] J. W. Tukey, "Comparing individual means in the analysis of variance," *Biometrics*, Vol. 5 (1949), pp. 99–114.
[17] J. W. Tukey, "Quick and dirty methods in statistics, Part II, Simple analyses for standard designs," *Proc. Fifth Ann. Convention, Amer. Soc. for Quality Control*, (1951), pp. 189–197.
[18] J. W. Tukey, "The problem of multiple comparisons," unpublished dittoed notes, 396 pp., Princeton University (1953).