J. Appl. Prob. **20**, 822–834 (1983)
Printed in Israel

PERMUTATION PROBABILITIES FOR GAMMA RANDOM VARIABLES

ROBERT J. HENERY,* *University of Strathclyde*

Abstract

The order statistics of a set of independent gamma variables, in general not identically distributed, may serve as a basis for ordering players in a hypothetical game. An alternative formulation in terms of negative binomial variables leads to an expression for the probability that the random gammas are in a given order. This expression may contain rather many terms and some approximations are discussed — firstly as the gamma parameters α_i tend to equality with all n_i the same, and secondly when the probability of an inversion is small. In another interpretation the probabilities discussed arise in the statement of confidence limits for the ratios of population variances, and here the inversion probability is small enough usually that lower and upper bounds may be given for the probability that the sample variances occur in their expected order. These bounds are calculated from the probability that two variables are in expected order, and for gamma variables this probability is obtained from the F-distribution.

GAMMA DISTRIBUTION; NEGATIVE BINOMIAL; APPROXIMATION; BOUNDS; TREND TESTS

1. Introduction

Permutations and their probabilities arise in fields as diverse as consumer preference studies, voting bias and rank order test. One of the simplest ways to define the probability of a permutation $\pi = \{\pi_1, \pi_2, \cdots, \pi_m\}$ is to take a set of underlying independent random variables $X_1, \cdots, X_m$ and let the probability of π be

$$\Pr\{\pi\} = \Pr\{X_{\pi_1} < X_{\pi_2} < \cdots < X_{\pi_m}\},$$

and this is the model used throughout. The simplest model has the random variables mutually independent with distribution functions $F(x, \theta_i)$ which depend on a single parameter θ_i, and with no loss of generality we take $\theta_1 > \theta_2 > \cdots > \theta_m$. The case of identically distributed variables is not so interesting since then all permutations are equally likely.

Received 8 February 1982; revision received 28 January 1983.
* Postal address: Department of Mathematics, University of Strathclyde, Livingstone Tower, 26 Richmond St, Glasgow G1 1XH, U.K.

This paper deals with the two-parameter gamma distribution with density

$$f(x, r, \alpha) = x^{r-1}\alpha^r \exp(-\alpha x)/\Gamma(r) \tag{1.1}$$

for $\alpha > 0$, $x > 0$ and mostly for integer $r > 0$. If X has density (1.1) we say that X is gamma(r, α).

Savage (1957) considers several classes of distribution function $F(x, \theta)$, two of which are of interest here. The first is the exponential family with densities of the form

$$f(x, \theta_i) = g(\theta_i)h(x)\exp(-\theta_i x), \tag{1.2}$$

where g and h are non-negative functions, and the second class of distributions is a subset of (1.2) being the class whose distribution functions $F(x, \theta_i)$ are of the form

$$1 - F(x, \theta_i) = \{1 - H(x)\}^{\theta_i}, \tag{1.3}$$

for some absolutely continuous distribution function $H(x)$. Since permutation probabilities are unaffected if the X_i variables are transformed by a monotonic increasing function, we might as well choose $H(x) = 1 - \exp(-x)$ for $x \geqq 0$, $H(x) = 0$ for $x < 0$. In discussing the admissibility of rank order tests against trends in θ, Savage derives inequalities for permutation probabilities. These inequalities are based on a partial ordering of permutations and on order-preserving functions of permutations and a unified theory of such inequalities is presented by Marshall and Olkin (1979).

For type (1.3) the probability of permutation π is $P(\pi)$,

$$P(\pi) = \prod_{i=1}^{m} \theta_{\pi_i} \prod_{j=1}^{m} \left(\sum_{k=j}^{m} \theta_{\pi_k}\right)^{-1}, \tag{1.4}$$

and the probability that X_i is the smallest of $X_1, \cdots, X_m$ is just $\theta_i / \Sigma_{k=1}^m \theta_k$. The place probability $P_i(k)$, the probability that X_i is the kth smallest of $X_1, \cdots, X_m$, can be found by summing (1.4) over permutations with $\pi_k = i$, but this may involve prohibitively many terms for large m and for k close to m, so simple approximations for $P_i(k)$ are then desirable.

For distributions which are totally positive (this includes type (1.2)), Savage (1957) shows that $P_i(1)$ decreases as θ_i decreases and Henery (1981a) notes that $P_i(m)$ increases as θ_i decreases. Normal $\mathrm{N}(\theta_i, 1)$ variables are also of type (1.2), and for this distribution Henery (1981b) gives an upper bound for $P_i(1)$ as well as a Taylor series expansion for $P_i(k)$ up to fourth order in θ_i.

This paper deals mostly with gamma variables, and in Section 3 we write down the probability of a permutation π by using the correspondence between gamma(r_i, α_i) and negative binomial variables. The resulting expression (3.1), which reduces to (1.4) when $r_i \equiv 1$, contains too many terms when m is even moderately large, so we develop two approximations to $P(\pi)$. In Section 4 a simple linear approximation is given for gamma(r, α_i) variables whose parameters α_i are nearly equal — a situation that may arise in handicap systems for which all variables are intended to have distributions as close to each other as possible.

Special importance is attached to the case where the independent variables have a natural ordering and we require the probability that the variables occur in this natural order. For instance we may have $E(X_1) < E(X_2) < \cdots < E(X_m)$ and if the random variables occur in the order of their expectations we say that the event CO (correct order) has occurred. Gibbons, Olkin and Sobel (1977) give many examples in which it is desirable to order populations by their means or variances, and for the statement of confidence limits it is necessary to calculate the probability P(CO) that the correct order has occurred. For the common assumption of normal populations the ordering of variance estimates leads, if the degrees of freedom are even, to the type of gamma variables considered here. However the evaluation of (1.4) is laborious if not impossible in many cases, so Bechhofer and Sobel (1954) recommend a logarithmic transformation for the χ^2 distribution, thus transforming the problem to the ordering of approximately normal variables. An important practical case arises when the population means form an arithmetic progression (or in the case of sample variances a geometric progression), and in this 'trend' case we often require the probability P(CO) given equal sample sizes and assuming constant variance. Table P1 of Gibbons, Olkin and Sobel (1977) gives these probabilities for up to ten populations, i.e. they give $P(X_1 < \cdots < X_m)$ where the X_i are independent normals of unit variance with means $E(X_i) = i\theta$, $1 \leqq i \leqq m$, $2 \leqq m \leqq 10$, for values of θ up to 4.4 The case of sample variances for normal populations whose variances are in geometric progression with common ratio θ and with equal degrees of freedom ν, has correct order probabilities tabulated for $m = 3$ to $m = 8$ and for θ up to 3.0 in Table P3 of Gibbons, Olkin and Sobel (1977). However, the entries of Table P3 are approximate, being the result of Monte Carlo calculations due to Schafer and Rutemiller (1975). Bechhofer and Sobel (1954) give exact values of P(CO) for ordering two or three sample variances, again assuming equal degrees of freedom. The last authors suggest the transformation $X_i = \sqrt{\tfrac{1}{2}(\nu - 1)} \cdot \log s_i^2$ which renders the X_i approximately normal with unit variance and with means $E(X_i) = i\sqrt{\tfrac{1}{2}(\nu - 1)} \cdot \log \theta$. Although this normal approximation works quite well even for rather small degrees of freedom, it is restricted to $m \leqq 10$ due to lack of tables and absence of numerical techniques for larger m.

In Section 5 we give upper and lower bounds for P(CO) which may be used when the trend is sufficiently strong that inversions in the expected order are very unlikely. This is the situation that arises when an experimenter wishes to rank populations according to their means and also wishes that the probability of a completely correct ordering be close to 1.

We shall also give, in Section 6, an approximation to $P(\mathrm{CO})$ based on the properties of the negative exponential for which $P(\mathrm{CO})$ takes its simplest form. Both Sections 5 and 6 are based on knowledge of the elementary probability $P(X_1 < X_2)$, so we begin, in Section 2, with the well-known problem of ordering two gamma random variables.

2. Gamma and negative binomial

Let events A occur independently and at random so that the number of events occurring in a given interval is a Poisson random variable with rate α. The waiting time between events has a negative exponential distribution with mean α^{-1}, and the total time T between $r + 1$ events has a gamma(r, α) distribution. We obtain some simple results for order statistics of gamma random variables by using the equivalence of the following two events: (i) the number of Poisson (rate α) arrivals in a fixed time interval t is less than r; and (ii) the sum of r independent negative exponentials (mean $1/\alpha$) is greater than t. Of course the random sum mentioned in (ii) has a gamma(r, α) distribution.

Now consider a second Poisson process, independent of the first, in which events B occur at an average rate β. The total number of events of either type is also Poisson with rate $\alpha + \beta$, and in this combined process the probability that a randomly selected event was type A is $\alpha/(\alpha + \beta)$. Indeed, a sequence of events in the combined process becomes a sequence of Bernoulli trials by noting whether each event is type A or B. Thus, for example, if a fixed number n of events is observed in the combined process, the number of these that are of type A is a binomial random variable with probability of success $p = \alpha/(\alpha + \beta)$ and number of trials n. Similarly, if we start from a specified instant and count the number N_B of B events occurring before the rth A event, then the distribution of N_B is negative binomial with parameters r and $p = \alpha/(\alpha + \beta)$.

Using this equivalence of events for Poisson arrivals and gamma variables, we see that the number N_B of B arrivals in the random time interval T is negative binomial if T has a gamma(r, α) distribution and r is integer. Engel and Zijlstra (1980) have used this property to give a characterisation of the gamma distribution: in other words they have shown that T is a gamma(r, α) if and only if N_B is negative binomial.

Now consider a simple game in which two players have to collect a fixed number of events of a given type, respectively n_A of type A and n_B of type B. Both players start with no events to their credit, and the winner is the first player to reach his target. In terms of gamma random variables, we choose two independent variables T_A and T_B, where T_A is gamma(n_A, α) and T_B is gamma(n_B, α), and the first player wins if T_A is less than T_B.

Since the number N_B of B events occurring before the n_Ath A event is negative binomial,

$$\Pr\{N_B = k\} = p^{n_A}(1-p)^k \binom{n_A + k - 1}{n_A - 1},$$

where $p = \alpha/(\alpha + \beta)$, it is clear that the player who collects A's will win with probability

$$\Pr\{N_B < n_B\} = p^{n_A} \sum_{k=0}^{n_B - 1} (1-p)^k \binom{n_A + k - 1}{n_A - 1}, \tag{2.1}$$

and by observing that one player must win, we obtain the identity

$$\Pr\{N_B < n_B\} + \Pr\{N_A < n_A\} = 1,$$

or

$$p^{n_A} \sum_{k=0}^{n_B - 1} (1-p)^k \binom{n_A + k - 1}{n_A - 1} + (1-p)^{n_B} \sum_{k=0}^{n_A - 1} p^k \binom{n_B + k - 1}{n_B - 1} = 1. \tag{2.2}$$

Alternatively we can express the win probability by means of the well-known relation between the cumulative negative binomial distribution and the binomial distribution (Johnson and Kotz (1969)). Thus, noting that when $n_A + n_B - 1$ events have accumulated one (and only one) of the players must have accumulated his target, the probability that it is the first player is

$$\Pr\{M_A \geqq n_A\} = \sum_{r=n_A}^{n_A + n_B - 1} p^r (1-p)^{n_A + n_B - 1 - r} \binom{n_A + n_B - 1}{r}, \tag{2.3}$$

where M_A is a binomial variable, being the number of A's among the first $n_A + n_B - 1$ events of either type. Thus the probability that A wins is given by the right-hand side of either (2.1) or (2.3), although (2.3) is the more useful since tables of the binomial distribution are more readily available.

The equivalence of the events $\{T_A < T_B\}$ and $\{N_B < n_B\}$ shows that the right-hand side of Equation (2.1) is also the probability that T_A is smaller than T_B, as could be shown with very little difficulty, by integrating the joint density of T_A and T_B over the region $T_A < T_B$:

$$\begin{aligned} \Pr\{T_A < T_B\} &= \int_0^\infty \frac{\alpha^{n_A}}{\Gamma(n_A)} y^{n_A - 1} e^{-\alpha y} dy \int_y^\infty \frac{\beta^{n_B}}{\Gamma(n_B)} x^{n_B - 1} e^{-\beta x} dx \\ &= \int_0^\infty \frac{\alpha^{n_A}}{\Gamma(n_A)} y^{n_A - 1} e^{-\alpha y} \sum_{i=0}^{n_B - 1} \frac{(\beta y)^i}{\Gamma(i)} e^{-\beta y} dy. \end{aligned} \tag{2.4}$$

The probability (2.1) is also of interest in connection with the F or variance ratio distribution. In fact if X_1 and X_2 are χ^2 random variables with $2n_1$ and $2n_2$ degrees of freedom respectively, then $T_1 = X_1/2\alpha_1$ is gamma(n_1, α_1) and $T_2 = X_2/2\alpha_2$ is gamma(n_2, α_2). The event $T_1 < T_2$ is then equivalent to the event $F < \alpha_1 n_2/\alpha_2 n_1$ where F has a variance ratio distribution based on $2n_1$, $2n_2$

degrees of freedom. Later we deal mostly with the case of equal sample sizes or equal degrees of freedom $2n_1 = 2n_2$, and we use the notation $R(\theta, 2n)$ for the probability that such an F-statistic is less than $\theta = \alpha_1/\alpha_2$. We use then the well-known result

$$R(\theta, 2n) = p^n \sum_{k=0}^{n-1} (1-p)^k \binom{n+k-1}{k}, \tag{2.5}$$

where $\theta = p/(1-p)$, so that $p = \theta/(1+\theta)$, and the common degrees of freedom are $2n$.

3. m-player game

These results extend quite easily to m players, collecting $1 + n_i$ Poisson events of type A_i and rate α_i, $i = 1$ to m. The players are then ranked according to how soon they achieve their targets (all players start from scratch at time 0). The player who first achieves his target is the winner; the next player to achieve his target is second and so on. In terms of the waiting times $T_1, T_2, \cdots, T_m$ (which all have gamma distributions), the players are ranked $(1, 2, \cdots, m)$ if $T_1 < T_2 < \cdots < T_m$. Again, the probability of this last event is quite easily evaluated from first principles, but we prefer the challenge of formulating the problem as a series of negative binomials. For the sake of clarity we take the case $T_1 < T_2 < \cdots < T_m$, which corresponds to player order $(1, 2, \cdots, m)$.

The negative binomial is the distribution of the number of failures before the rth success in a sequence of Bernoulli trials. In the present problem this arises when we look back from the time T_s at which player s first achieves his target. Up till that point player s competes, if $s < m$, against players $s+1, s+2, \cdots, m$ for the sth place, and events of type $A_s, A_{s+1}, \cdots, A_m$ constitute the Bernoulli trials with probability of success $\alpha_s/(\alpha_s + \cdots + \alpha_m)$. The number of successes is $r = n_s + 1$, and we define the negative binomial variable I_s as the number of events $A_{s+1}, A_{s+2}, \cdots, A_m$ preceding the (n_s+1)th occurrence of event A_s. Of course once the first $m-1$ positions are decided the last position is also decided, so we define I_m to be 0. The negative binomial variables $I_1, \cdots, I_{m-1}$ are mutually independent as we show now. The first step is to show that the number of failures I_s preceding the (n_s+1)th success is independent of the order in which these failures occur. To start with take the case of event A_1, and let I_1 be the number of failures, i.e. events from the set $\{A_2, A_3, \cdots, A_m\}$, preceding the (n_1+1)th event A_1. Also let the observed sequence of failures be $A_{F_1}, A_{F_2}, \cdots, A_{F_M}$, where $F_i \in \{2, 3, \cdots, m\}$ for $1 \leqq i \leqq M$, and M is some fixed integer. Denote by $\mathscr{E}$ the event: $I_1 = i_1$ *and* the sequence of failures is $A_{F_1}, A_{F_2}, \cdots, A_{F_M}$. Then the probability of $\mathscr{E}$ is, writing $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_m$ and taking only the case $M > i_1$ (the case $M \leqq i_1$ is easily dealt with also):

$$\begin{aligned}\Pr\{\mathscr{E}\} &= \binom{n_1+i_1}{n_1}\cdot\left(\frac{\alpha_1}{\alpha}\right)^{n_1+1}\prod_{r=1}^{i_1}\left(\frac{\alpha_{F_r}}{\alpha}\right)\cdot\prod_{s=i_1+1}^{M}\frac{\alpha_{F_s}}{(\alpha-\alpha_1)}\\ &= \binom{n_1+i_1}{n_1}\left(\frac{\alpha_1}{\alpha}\right)^{n_1+1}\left(\frac{\alpha-\alpha_1}{\alpha}\right)^{i_1}\cdot\prod_{s=1}^{M}\frac{\alpha_{F_s}}{(\alpha-\alpha_1)}\\ &= \Pr\{I_1=i_1\}\cdot\Pr\{A_{F_1},A_{F_2},\cdots,A_{F_M}\}.\end{aligned}$$

Thus I_1 is independent of the relative order of events $\{A_2,\cdots,A_m\}$ and is therefore independent of any event described entirely by this relative order, and so I_1 is independent of the set $\{I_2,\cdots,I_{m-1}\}$. Similarly I_2 is independent of the set $\{I_3,\cdots,I_{m-1}\}$ and so on.

To ensure that the players emerge in the given order $(1,2,\cdots,m)$, we put restrictions on the possible values i_s of the I_s variables as follows. During the time interval T_{s-1} to T_s only events $A_s,\cdots,A_m$ are pertinent to the question of deciding sth place, and we have one event A_s occurring at T_s (by definition). The number of A_s events before T_s is n_s, and the number of events $A_s,\cdots,A_m$ before T_{s-1} is i_{s-1}. Hence the number of pertinent events occurring after T_{s-1} and before T_s is $n_s+i_s-i_{s-1}$, and this is a non-negative integer so $0\leqq i_{s-1}\leqq n_s+i_s$ for $1<s\leqq m$. (Define $i_m=0$.)

We can now write down the probability P for the order $(1,2,\cdots,m)$ as the product of the probability densities for the $i_1,\cdots,i_{m-1}$ summed over the permissible ranges of the variables $i_1,\cdots,i_{m-1}$. We find that P is equal to

$$\sum_{i_{m-1}=0}^{n_m+i_m}\sum_{i_{m-2}=0}^{n_{m-1}+i_{m-1}}\cdots\sum_{i_1=0}^{n_2+i_2}\prod_{k=1}^{m-1}\left(\frac{\alpha_k}{\sum_{j=k}^{m}\alpha_j}\right)^{n_k+1}\left(\frac{\sum_{j=k+1}^{m}\alpha_j}{\sum_{j=k}^{m}\alpha_j}\right)^{i_k}\binom{n_k+i_k}{n_k}. \tag{3.1}$$

The number of terms in (3.1) may be very large even for moderate values of n_s, $1\leqq s\leqq m$, and the problem is made worse if we try to get place probabilities $P_i(k)$ by summing (3.1) over the $(m-1)!$ permutations π with $\pi_k=i$. Going back to the original gamma(r_i,α_i) variables, the win probabilities $P_i(1)$ are given by

$$\begin{aligned}P_i(1) &= \int_0^\infty f(x,r_i,\alpha_i)\prod_{j\neq i}^{m}\{1-F(x,r_j,\alpha_j)\}dx\\ &= \int_0^\infty\frac{x^{r_i-1}\alpha_i^{r_i}}{\Gamma(r_i)}\left(\prod_{j\neq i}^{m}\sum_{k=0}^{r_j}\frac{(\alpha_jx)^k}{k!}\right)\exp(-\alpha x)dx,\end{aligned}$$

where $\alpha=\Sigma\alpha_i$, and this integral can be reduced to a multiple sum easily, although there are rather many terms and there appears to be no simple analogue to (2.3). A crude upper bound for $P_i(1)$ is got by noting that if $N=1+\Sigma_1^m n_j$ events have occurred *at least* one player has achieved his target, so the game has been decided at or before the Nth event. Now the first player wins if at the (n_1+1)th occurrence of event A_1 all other players are short of their respective targets, and this implies that the total number of non-A_1 events before the (n_1+1)th A_1 is less than or equal to $\Sigma_2^m n_j = N - n_1 - 1$. Using (2.3) this gives, for example,

$$\text{(3.2)} \qquad \Pr\{\text{player 1 wins}\} \leqq \sum_{k=n_1+1}^{N} \left(\frac{\alpha_1}{\alpha}\right)^k \left(1-\frac{\alpha_1}{\alpha}\right)^{N-k} \binom{N}{k},$$

with equality only if $m = 2$.

Although (3.1) is readily computed for given n_k, α_k, $k = 1, 2, \cdots, m$, little can be said theoretically about its magnitude except in limiting cases. For example consider the following try at constructing a fair game. Suppose the $\alpha_1, \cdots, \alpha_m$ are different and that constants n_k have been chosen so that $\alpha_k^{-1}(n_k+1) = N$. If we consider the limit $N \to \infty$, $\alpha_1, \cdots, \alpha_m$ fixed, it is clear that the game is fair in the sense that all players have waiting times which are normal with the same mean N. However the variances are different and the player with the largest variance has the greatest probability of winning (or of being last). The central limit theorem shows that the transformed variables $Z_k = (T_k - N)/\sqrt{N}$ tend to normality and since the Z_k involve the same linear transformation on all T_k variables, we see that expression (3.1) is equal to the probability that $Z_1 < Z_2 < \cdots < Z_m$ where, in the limit $N \to \infty$, Z_k is $\mathrm{N}(0, \alpha_k^{-1})$. From the symmetry of these normals it follows that the permutations $1, 2, \cdots, m$ and $m, \cdots, 2, 1$ have the same probabilities. Accordingly all players have the same expected place $(m+1)/2$ and the game is fair in that sense. It is also fair in that player i will beat player j with probability $\frac{1}{2}$ (the permutation (i, j) is as likely as (j, i)).

4. The limit $\alpha_i \to 1$

We consider now the probability (3.1) for the given order when all players have the same target, $n_i + 1 = r$ for all i, and the rates α_i are close to unity. We will assume that the differences $|\alpha_i - 1|$ are so small that a Taylor series expansion of (3.1) up to linear terms in $\alpha_i - 1$ is sufficiently accurate. However, to demonstrate the connection between this expansion and the order statistics for gamma$(r, 1)$ random variables, and incidentally to state a slightly more general result, we revert to the formulation of the problem by means of the waiting times $T_1, T_2, \cdots, T_m$.

Let the probability density of T_i be $f(t, \alpha_i)$, with $f(t, \alpha_i)$, $i = 1, 2, \cdots, m$, all belonging to the same family of densities (1.2) written as

$$f(t, \alpha) = \exp\{-\alpha t + C(t) + D(\alpha)\}, \qquad t \in \{a, b\},$$

for given functions $C(t)$, $D(\alpha)$, and for given constants, a, b (which may not depend on α). For example, the T_i's might be independent $\mathrm{N}(\alpha_i, 1)$ distributions or, as in the present discussion, the T_i's might be independent gamma(r, α_i) variables with fixed r.

We wish to evaluate the probability that $T_1 < T_2 < \cdots < T_m$ where the T_i's are independent, and T_i has probability density $f(t_i, \alpha_i)$, $i = 1, 2, \cdots, m$. This is the integral of the joint density over the region $R = \{a < t_1 < t_2 < \cdots < t_m < b\}$:

$$\Pr\{T_1 < T_2 < \cdots < T_m\} = \int_R \prod_{i=1}^{m} f(t_i, \alpha_i) dt_i.$$

We denote this probability by $P(\pi)$, where π is the permutation $(1, 2, \cdots, m)$. $P(\pi)$ will be approximated by the linear terms of a Taylor series expansion in the α_i's about the symmetric point $\alpha_1 = \alpha_2 = \cdots = \alpha_m = \alpha_0$ say, and the value of $P(\pi)$ when all the α_i's are equal to α_0 will be written $P_0(\pi)$. Clearly $P_0(\pi) = 1/m!$ since all permutations are equally likely in the symmetric case. For future reference note that the joint density of the order statistics from a random sample of size m from a population with density $f(t, \alpha_0)$ is

$$m! \prod_{i=1}^{m} f(t_i, \alpha_0), \qquad \text{for } (t_1, t_2, \cdots, t_m) \in R,$$

and is 0 otherwise.

With a slight abuse of notation, the partial derivative of $P(\pi)$ with respect to α_i, evaluated at $\alpha_1 = \alpha_2 = \cdots = \alpha_m = \alpha_0$ will be denoted by $\partial P_0(\pi)/\partial \alpha_i$ and this is easily seen to be

$$\begin{aligned} \frac{\partial P_0(\pi)}{\partial \alpha_i} &= \int_R \frac{\partial}{\partial \alpha_0} \log f(t_i, \alpha_0) \prod_{i=1}^{m} f(t_i, \alpha_0) dt_i. \\ &= \frac{1}{m!}\{-\mu_{i;m} + D'(\alpha_0)\}. \end{aligned}$$

Here $\mu_{i;m}$ is the expected value of the ith order statistic in a random sample of size m from the distribution with density $f(t, \alpha_0)$.

To linear terms in $\alpha_i - \alpha_0$ we have

$$P(\pi) = P_0(\pi) + \frac{1}{m!} \sum_{i=1}^{m} (-\mu_{i;m} + D'(\alpha_0))(\alpha_i - \alpha_0). \tag{4.1}$$

Since $P(\pi)$ is unaltered by multiplying all the T_i by the same positive constant, or if the same constant is added to all the T_i, equivalent to dividing the α_i by a constant or adding a constant to $\mu_{i;m}$ respectively, we may choose origin and scale for T_i so that $\Sigma\, \alpha_i = m$ and $\Sigma\, \mu_{i;m} = rm$, as is appropriate for gamma$(r, 1)$ samples. For the gamma(r, α) density, $D(\alpha) = r \log \alpha$ and (4.1) becomes

$$P(\pi) = \frac{1}{m!} + \frac{1}{m!} \sum_{i=1}^{m} \{r - \mu_{i;m}\}(\alpha_i - 1). \tag{4.2}$$

Summing (4.2) over all $(m-1)!$ permutations with α_j in kth position we obtain

$$P_k(j) = \frac{1}{m} + \frac{1}{(m-1)}\{r - \mu_{k;m}\}(\alpha_j - 1). \tag{4.3}$$

5. Bounds for permutation probabilities

Let $X_1, \cdots, X_m$ be m independent random variables $m > 1$, each with a continuous density $f_i(x)$, $1 \leqq i \leqq m$. The event $X_1 < X_2 < \cdots < X_m$ can be defined as a succession of comparisons $X_1 < X_2$ *and* $\cdots$ *and* $X_{m-1} < X_m$, and if these comparisons were independent, which they are certainly not, the probability $P(X_1 < \cdots < X_m)$ would be

$$\prod_{i=1}^{m-1} P(X_i < X_{i+1}). \tag{5.1}$$

In fact this is an upper bound for $P(X_1 < \cdots < X_m)$ as we proceed to show after stating some preliminary and fairly obvious facts. Firstly, the random variable obtained by truncating X below at t is stochastically greater than X, or $P(X > x \mid X > t) \geqq P(X > x)$ for all t. Secondly, if X is stochastically greater than Y and if $g(x)$ is an increasing function, then $Eg(X) \geqq Eg(Y)$. For a proof of this last, see for example Marshall and Olkin (1979).

Using $g(x) = 1 - P(x < X_2 < \cdots < X_m)$ with stochastically ordered variables $Y = X_1$ and $X = X_1$ truncated below at x, we obtain

$$P(X_1 < X_2 < \cdots < X_m \mid X_1 > x) \leqq P(X_1 < X_2 < \cdots < X_m),$$

with strict inequality if $P(X_1 < x) > 0$ and $P(X_1 < X_2 < \cdots < X_m) > 0$. Now the probability $P(X_1 < X_2 < \cdots < X_m)$ can be expressed as the integral, for $m > 2$,

$$\begin{aligned} &\int f_1(x) \cdot P(x < X_2 < \cdots < X_m) \cdot dx \\ &= \int f_1(x) \cdot P(X_2 < \cdots < X_m \mid X_2 > x) \cdot P(X_2 > x) \cdot dx \\ &\leqq P(X_2 < \cdots < X_m) \cdot \int f_1(x) \cdot P(X_2 > x) \cdot dx \end{aligned}$$

and interpreting this last integral we have

$$P(X_1 < X_2 < \cdots < X_m) \leqq P(X_1 < X_2) \cdot P(X_2 < \cdots < X_m), \tag{5.2}$$

with equality only when $P(X_1 < X_2) = 0$ or 1. Repeated application of (5.2) shows that (5.1) is an upper bound for $P(X_1 < \cdots < X_m)$. It is also easy to obtain an upper bound similar to (5.2) by taking the first k and last $(m-k+1)$ variables instead of the first two and last $(m-1)$.

To obtain a lower bound for $P(X_1 < \cdots < X_m)$ it is only a matter of choosing an event which implies the given order and whose probability is known. For example, choose $m+1$ constants $x_0, x_1, \cdots, x_m$ with $-\infty = x_0 < x_1 < \cdots < x_m = \infty$ and consider the event $x_{i-1} < X_i < x_i$, $1 \leqq i \leqq m$. Then clearly we have

$$P(X_1 < \cdots < X_m) \geqq \prod_{i=1}^{m} P(x_{i-1} < X_i < x_i), \tag{5.3}$$

with equality possible only if $P(X_1 < \cdots < X_m) = 0$ or 1.

A lower bound more useful when the probability of inversion in the natural order is small can be obtained by the method of inclusion and exclusion. Consider the $m-1$ pairs of adjacent variables X_i, X_{i+1} and the associated events $X_i < X_{i+1}$. If $X_i > X_{i+1}$ we shall say that X_i and X_{i+1} are inverted or are out of natural order. Noting that $X_1 < \cdots < X_m$ if and only if no adjacent pair is inverted, the probability of no inversion is greater than

$$1 - \sum_{i=1}^{m-1} P(X_i > X_{i+1}) \tag{5.4}$$

since, in subtracting the sum of the probabilities $P(X_i > X_{i+1})$, double transpositions $X_i > X_{i+1}$ and $X_j > X_{j+1}$, $i \neq j$, have been counted twice instead of once.

The single most important permutation problem arising in applications occurs in the case of trend in the means. And within this class of problems the normal and χ^2 distributions play a central role, so we shall spell out some of the details in applying the bounds (5.1) and (5.4) for these distributions. For our purposes the normal serves only to approximate the χ^2 for large degrees of freedom.

For independent normal variables with common variance σ^2 and with means $E(X_i) = i\sqrt{2}\sigma\delta$ for $i = 1, \cdots, m$ the bound (5.1) reduces to $\Phi(\delta)^{m-1}$, and the lower bound (5.4) reduces to $1-(m-1)\cdot\{1-\Phi(\delta)\}$. Should this lower bound be negative, as is quite possible, the bound (5.3) may be used instead.

In ordering populations according to their variability it is frequently necessary to calculate the probability $P(\mathrm{CO})$ of correctly ordering the variables X_i, $1 \leqq i \leqq m$, where the X_i are independent gamma$(\nu/2, \theta^{-i})$ random variables, where ν is an integer. Gibbons, Olkin and Sobel (1977) discuss statistical inference problems of this type. The probability $R(\theta, \nu)$ that $X_i < X_{i+1}$ is given by (2.5) for $\nu = 2n$, and combining (5.1) and (5.4) we obtain

$$1-(m-1)\cdot\{1-R(\theta,\nu)\} < P(\mathrm{CO}) < R(\theta,\nu)^{m-1}. \tag{5.5}$$

In the typical inference problem, $P(\mathrm{CO}) = P$, say, is used in a confidence statement about the ratio θ of successive variances, and in this case P may be assumed close to unity. Then the bounds in (5.5) must be close since they differ by a quantity of order $(1-P)^2$.

6. Estimate of P(CO)

The fact that the ranks of a set of random variables are unaffected by applying the same monotonic increasing transformation on each variable means that whole classes of distributions share the same permutation probabilities. Thus the class (1.3), the negative exponential class, has permutation probabilities given by (1.4). Also the bounds (5.1) and (5.4) are defined once the adjacent pair probabilities $p_i = P(X_i < X_{i+1})$ are known. Indeed for certain distributions such as the negative exponential it is possible to obtain the probability of any permutation given only the p_i for $1 \leqq i \leqq m-1$. In fact the coefficients α_i, α_{i+1} of an adjacent pair of negative exponentials are related by $\alpha_{i+1} = \alpha_i(1-p_i)/p_i$ if $P(X_i < X_{i+1}) = p_i$.

Alternatively if we know the p_i for an arbitrary set of variables $X_1, X_2, \cdots, X_m$ we can construct a set of exponential variables $Y_1, Y_2, \cdots, Y_m$ with the *same* pair probabilities $p_i = P(X_i < X_{i+1}) = P(Y_i < Y_{i+1})$ by taking $\alpha_{i+1} = \alpha_i(1-p_i)/p_i$ for $1 \leqq i < m$. We may then approximate $P(X_1 < X_2 < \cdots < X_m)$ by $P(Y_1 < Y_2 < \cdots < Y_m)$ since, by construction, these have the same bounds (5.1) and (5.4). In particular, if all the p_i are equal to p say, then we can take $\alpha_i = x^i$ where $x = (1-p)/p$ and (1.4) becomes

$$\prod_{i=1}^{m-1}\left(\sum_{k=0}^{i} x^k\right)^{-1}. \tag{6.1}$$

As a numerical illustration, consider the ordering of four sample variances s_1^2, s_2^2, s_3^2, s_4^2 each based on 60 degrees of freedom, and let us find the probability P(CO) that the estimates are in the correct order when the ratios of successive population variances are all equal to 1.4, i.e. $\sigma_2^2/\sigma_1^2 = \sigma_3^2/\sigma_2^2 = \sigma_4^2/\sigma_3^2 = \theta = 1.4$. For ν degrees of freedom, with an $N(0, \sigma^2)$ sample, a variance estimate s^2 is a gamma$(\nu/2, \nu/2\sigma^2)$ variable; and in the present case we have $\sigma_i^2 = \sigma_1^2\theta^{i-1}$, $i = 1, 2, 3, 4$. Since σ_1^2 is a common factor it drops out of all comparisons so we may as well take $\sigma_1^2 = \nu/2$. Then $P(s_1^2 < s_2^2 < s_3^2 < s_4^2) = P(X_1 < X_2 < X_3 < X_4)$ where X_i is a gamma$(\nu/2, \theta^{1-i})$ variable. Using (3.1) with $\alpha_i = 1.4^{1-i}$ and $n_i = 30$, $1 \leqq i \leqq 4$, we obtain $P(\mathrm{CO}) = 0.71780$ to five decimal places. There are more than 70000 terms in (3.1) here, so it would be necessary to use double precision arithmetic to get much greater accuracy.

The probability $p = R(\theta, \nu)$ that an adjacent pair of variances is correctly ordered is found via (2.5) to be $p = 0.90229$ from which we get lower and upper bounds for P of 0.70687 and 0.73458 from (5.4) and (5.1) respectively.

Using the negative exponential approximation (6.1) with $p = 0.90229$ gives, with $x = (1-p)/p = 0.10829$,

$$P = \frac{1}{(1+x)(1+x+x^2)(1+x+x^2+x^3)} = 0.71846.$$

The suggestion of Bechhofer and Sobel (1954) is to approximate P by $Q = P(Z_1 < Z_2 < Z_3 < Z_4)$ where Z_i is $\mathrm{N}(i\mu, 1)$ with $\mu = \sqrt{\frac{1}{2}(\nu - 1)} \cdot \log \theta = 1.82751$. Q was calculated, using a reduction formula of Plackett (1954), to be 0.71660.

For comparison the Monte Carlo estimate of P given by Schafer and Rutemiller (1975) is 0.713.

Finally we should emphasize that (3.1) gives an exact calculation of $P(\mathrm{CO})$ for any set of population variances $\sigma_1^2, \sigma_2^2, \cdots, \sigma_k^2$ using sample estimates $s_1^2, s_2^2, \cdots, s_k^2$ all with *even* degrees of freedom $\nu_1, \nu_2, \cdots, \nu_k$. Formula (3.1) is inapplicable if one ν_i is odd, although approximation (6.1) may still be useful since we can use the F distribution to get the probabilities $P(s_i^2 < s_{i+1}^2)$ for $i = 1, \cdots, k-1$.

References

BECHHOFER, R. E. AND SOBEL, M. (1954) A single-sample multiple decision procedure for ranking variances of normal populations. *Ann. Math. Statist.* **25**, 273–289.

ENGEL, J. AND ZIJLSTRA, M. (1980) A characterisation of the gamma distribution by the negative binomial distribution. *J. Appl. Prob.* **17**, 1138–1144.

GIBBONS, J. D., OLKIN, I. AND SOBEL, M. (1977) *Selecting and Ordering Populations: A New Statistical Methodology.* Wiley, New York.

HENERY, R. J. (1981a) Permutation probabilities as models for horse races. *J. R. Statist. Soc. B* **43**, 86–91.

HENERY, R. J. (1981b) Place probabilities in normal order statistics models for horse races. *J. Appl. Prob.* **18**, 839–852.

JOHNSON, N. L. AND KOTZ, S. (1969) *Distributions in Statistics.* Wiley, New York.

MARSHALL, A. W. AND OLKIN, I. (1979) *Inequalities: Theory of Majorization and its Applications.* Academic Press, New York.

PLACKETT, R. L. (1954) A reduction formula for normal multivariate integrals. *Biometrika* **41**, 351–360.

SAVAGE, I. R. (1957) Contributions to the theory of rank order statistics — the 'trend' case. *Ann. Math. Statist.* **28**, 968–977.

SCHAFER, R. E. AND RUTEMILLER, H. C. (1975) Some characteristics of a ranking procedure for population parameters based on chi-square statistics. *Technometrics* **17**, 327–331.