On Some Distributions Generated by Riff-Shuffle Sampling

M. S. Son*
Department of Math/Stat
University of Vermont, Vermont, USA

H. I. Hamdy
Department of Business & Information System
Kuwait University, Kuwait

ABSTRACT

The work presented in this paper is divided into two parts. The first part presents finite urn problems which generate truncated negative binomial random variables. Some combinatorial identities that arose from the negative binomial sampling and truncated negative binomial sampling are established. These identities are constructed and serve important roles when we deal with these distributions and their characteristics. Other important results including cumulants and moments of the distributions are given in somewhat simple forms. Second, the distributions of the maximum of two chi-square variables and the distributions of the maximum correlated F-variables are then derived within the negative binomial sampling scheme. Although multinomial theory applied to order statistics and standard transformation techniques can be used to derive these distributions, the negative binomial sampling approach provides more information and deeper insight regarding the nature of the relationship between the sampling vehicle and the probability distributions of these functions of chi-square variables. We also provide an algorithm to compute the percentage points of these distributions. We supplement our findings with exact simple computational methods where no interpolations are involved.

***Key words**: Incomplete beta function, Binomial sum, Chi-square distribution, Negative binomial distribution, correlated F-variable, Random walk with absorbed barriers, Truncated Negative Binomial distribution, Poisson Distribution.*

1. INTRODUCTION

There are many practical problems which give rise to the negative binomial probability model. The distribution is used to represent the waiting time to reach a predetermined number of successes with fixed probability of success at any trial. The probability mass function is the general term in the expansion of $p^m q^{-m}$, where $p+q=1$. The distribution is also known as Pascal distribution, Pascal (1679). Kemp (1967) listed most of commonly used types of negative binomial and geometric distributions (see Johnson, Kotz, and Kemp (1992)). The negative binomial arises also from some stochastic models, for example the time-homogenous-birth and immigration process with zero initial population was first studied by McKendrick (1914), Kendall (1949). For detailed historical remarks one refers to Johnson, Kotz, and Kemp (1992), pp 203. Regarding the truncated negative binomial distribution, Boswell and Patil (1970) modeled the sizes of groups by zero-truncated negative binomial where the zeroes are not recorded. Hamdan (1975) has considered the truncated negative binomial to model the data of Reed and Reed (1965). Ahuja (1971) has investigated the n-fold convolution of zero-truncated negative binomial distribution. The concept of truncation in this study, however, is used in the sense that the probability mass function of the regular negative binomial is truncated from right, or truncated from left. This can be achieved through a finite urn model experiment or by utilizing the theory of random walk with absorbed barriers. However, the details concerning these random experiments will be given in subsequent sections.

In the following subsections, we view the negative binomial sampling through a sequence of independent trials generated by either flipping a coin or by urn models. Section 2 presents the right-truncated negative binomial probability mass function, its cumulants, and some useful identities. While section 3 treats the left-truncated negative binomial, its cumulants and some useful identities. Section 4 discusses the distribution of the maximum of two $i.i.d.$ chi–square random variables through the left truncated negative binomial sampling scheme. The distribution of the maximum correlated F-variable and its characteristics is also presented in Section 5 as a distribution that arises by a left-truncated negative binomial sampling procedure. Computational algorithms accompanied with critical values are given for both the distribution of the maximum chi-square and the distribution of the maximum correlated F-variable in Section 5.

*Corresponding author. E-mail : son@cems.uvm.edu
Manuscript received May. 12, 2006 ; accepted Jun. 16, 2006

2. NEGATIVE BINOMIAL SAMPLING

Consider a random experiment where an unbalanced coin is flipped for a sequence of independent trials until the head appears m times, where m is a specific positive integer predetermined beforehand. We also assume that at any trial the probability of observing a head equals p and the probability of observing a tail equals q, where $p + q = 1$. Upon observing the m^{th} head, the experiment is terminated. Let K denotes the number of observed tails preceding the m^{th} head. It is known that the random variable K follows the negative binomial distribution with the following probability mass function,

$$P(K = k) = \binom{m+k-1}{k} q^k p^m, \quad k = 0, 1, 2, ... \tag{2.1}$$

The special case of $p = q = 1/2$ provides the following probability mass function of the negative binomial distribution

$$P(K = k) = \binom{m+k-1}{k} (1/2)^m (1/2)^k, \quad k = 0,1,2, ... \tag{2.2}$$

It follows from (2.2) that

$$\sum_{k=0}^{\infty} \binom{m+k-1}{k} (1/2)^k = 2^m \tag{2.3}$$

In the following sections we present different forms of truncated negative binomial distributions, some characteristics of negative the binomial sampling vehicle.

2.1 Right-Truncated Negative Binomial Distribution (Riff-Shuffle Distribution)

Consider two urns A and B each containing m balls. A sequence of independent trials is performed to select urn A with probability p and urn B with probability q, where p and q are fixed during the sampling course and $p + q = 1$. At any trial, upon selecting an urn, a ball is drawn out from the selected urn. Sequentially, the experiment is continued until one of the two urns is completely exhausted. Let K be a random variable representing the number of balls drawn out from the non-exhausted urn when the other urn is completely exhausted. Clearly, when the experiment is terminated the non-exhausted urn still contains $m - k$ balls. Uppuluri and Bolt (1970), Lingappaiah (1987) (see also Johnson, Kotz, and Kemp (1992), pp 234 equation 5.92, with the correction $x = 0,1,..., m-1$ instead $x = 0,1,..., m$, for details) proposed the following probability mass function for the random variable K.

$$P(K = k) = \binom{m+k-1}{k}(p^m q^k + q^m p^k), \quad k = 0, 1,...,m-1. \tag{2.4}$$

The probability mass function in (2.4) is a proper probability mass function, since

$$\sum_{k=0}^{m-1} \binom{m+k-1}{k} p^m q^k = I_p(m,m),$$

where $I_p(m,m)$ is an incomplete beta integral, and $I_p(m,m) + I_q(m,m) = 1$. Johnson, Kotz, and Kemp (1990), page 235, indicated that the distribution can be thought of as a mixture of a two tailed-truncated negative binomial distribution.

The problem can also be viewed within the context of the theory of random walks with absorbed barriers. That is where a particle moves along the real line one step per unit length at a time, where moves are independent. Starting at zero, the particle makes random move to the right with probability p and to the left with probability q. Moreover, the particle will be absorbed if it reaches the m^{th} steps in either direction. In this case the random variable K represents the number of steps the particle has made in one direction before its absorption (final state) in the other direction. The probability mass function of K is given by (2.4).

Assume further the special case where the particle is equally likely to have a step to the right or a step to the left (where the probability equals ½ in either direction). Then (2.4) leads to the following right-truncated negative binomial distribution of the form

$$P(K = k) = \binom{m+k-1}{k} (1/2)^{m-1} (1/2)^k, \quad k = 0,1,...,m-1. \tag{2.5}$$

Since (2.5) is a proper probability mass function, it follows that

$$\sum_{k=0}^{m-1} \binom{m+k-1}{k} 2^{-k} = 2^{m-1}. \tag{2.6}$$

This result will be used in subsequent development to simplify results.

To see that the right-truncated negative binomial model presented in (2.5) can be obtained by truncating (2.2) from the right. We recall (2.2) and make use of (2.6). The result is

$$P(K \le m-1) = \sum_{k=0}^{m-1} \binom{m+k-1}{k} (1/2)^k (1/2)^m = 1/2.$$

Therefore, the probability mass function in (2.5) is obtained by dividing (2.2) by ½. Other characteristics of the distribution in (2.5), including cumulants, the expected sample size, and the variance are given in Hamdy et. al. (2003). However, in the following subsections we focus our investigation on the left-truncated negative binomial to study some related distributions.

2.2 Riff-Shuffle Sampling and the Left-Truncated Negative Binomial Distribution

Assume that after one of the urns is completely exhausted we continue to observe the remaining $m-k$ balls in the non-exhausted urn in a random fashion. An unbalanced coin is flipped for a sequence of independent trials where, the rule of the game is as follows. If urn A was exhausted first, then the game is continued exhausting urn B in a random fashion. If urn B was exhausted first, then the game is continued exhausting urn A in a random fashion. Assume that urn A is to be exhausted. An unbalanced coin is flipped. If the outcome is heads we select a ball from urn A with probability p, and if the outcome is tails we do nothing with probability q. On the other hand, if urn B is to be exhausted a ball is drawn from urn B with probability q if the coin's outcome is heads, and if the coin's outcome is tails we do nothing with probability p. Let K be a random variable representing the number of times we were not able to select a ball from the urn under exhaustion. Then the probability mass function of the random variable K can be written as

$$P(K=k)=\binom{m+k-1}{k}(p^m q^k+q^m p^k),\quad k=m,m+1,\ldots, \tag{2.7}$$

which is the left-truncated Negative binomial probability mass function. Moreover, in the context of random walk theory we discussed earlier in section 2.1, assume that after the particle has been absorbed in one direction (we have made m-steps in that direction), we want to continue to achieve the remaining $m-k$ steps in the other direction starting from the last state. In this case the particle has two choices either to move forward to a new state with probability p or q (depends on which direction the particle moves) or to stay at the current state with probability q or p. Eventually, the random variable K obeys the probability mass function in (2.7). To show that (2.7) is a proper probability mass function, we need to prove that the infinite sum of the right hand side of (2.7) is one. This is straight forward if we write

$$\sum_{k=m}^{\infty}\binom{m+k-1}{k}(p^m q^k+q^m p^k)=\sum_{k=0}^{\infty}\binom{m+k-1}{k}(p^m q^k+q^m p^k) -\sum_{k=0}^{m-1}\binom{m+k-1}{k}(p^m q^k+q^m p^k).$$

And since

$$\sum_{i=0}^{\infty}\binom{t+i-1}{i}a^i=(1-a)^{-t},\ |a|<1,$$

it is not hard to show that

$$\sum_{k=0}^{\infty}\binom{m+k-1}{k}(p^m q^k+q^m p^k)=2$$

On the other hand the finite sum

$$\sum_{k=0}^{m-1}\binom{m+k-1}{k}(p^m q^k+q^m p^k)=1$$

using Riff-Shuffle distribution in (2.4). This completes the assertion. Considering the special case of $p=q=1/2$ in (2.7), leads to the following Left-truncated negative binomial distribution of the form

$$P(K=k)=\binom{m+k-1}{k}(1/2)^{m-1}(1/2)^k,\quad k=m,m+1,\ldots \tag{2.8}$$

It follows from (2.8) that

$$\sum_{k=m}^{\infty}\binom{m+k-1}{k}2^{-k}=2^{m-1} \tag{2.9}$$

which will be used in simplifying results in subsequent sections.

2.2.1 The r^{th} Cumulant of Left-truncated Negative Binomial Distribution

In this context the r^{th} cumulant of the random variable $m+K$, whenever it exists, is given by the factorial moments

$$\mu'_{(r)}=E_K\left(\frac{(m+K+r-1)!}{(m+K-1)!}\right).$$

For the random variable K, which is distributed according to the left-truncated negative binomial probability mass function given in (2.7), the r^{th} cumulant has the form

$$\mu'_{(r)}=\frac{(m+r-1)!}{(m-1)!}\left[2^r+\sum_{k=m}^{m+r-1}(1/2)^{m+k-1}\binom{m+k+r-1}{k}\right] \tag{2.10}$$

The special cases of $r=1$ and $r=2$ in (2.10) provides the following first two factorial moments from which we obtain the mean and the variance of the random variable k

$$\mu'_{(1)}=E(m+K)=2m\left[1+(1/2)^{2m-1}\binom{2m}{m}\right]$$

It follows that the mean of the random variable K is given by

$$E(K)=m\left[1+(1/2)^{2m-2}\binom{2m}{m}\right]$$

$$\mu'_{(2)}=E\{(m+K+1)(m+K)\}=4m(m+1)\left[1+(1/2)^{2m}\binom{2m+1}{m}\right]$$

Hence, the variance of the left- truncated negative binomial random variable K can be written as

$$Var(K)=2m\left\{1+(1/2)^{2m}\binom{2m}{m}\left(1-2m(1/2)^{2m}\binom{2m}{m}\right)\right\},$$

where we have used (2.10) to obtain the above results.

In Section 3 we modify the random experiment which generates the left-truncated negative binomial distribution discussed in Section 2. First, we proceed to give the distribution of the maximum of two chi-square variables, then we proceed to give the distribution of the maximum of two correlated F-variable in Section 4

3. NEGATIVE BINOMIAL SAMPLING AND THE DISTRIBUTION OF THE MAXIMUM

Let X, Y be independent and identically distributed random variables each chi-square with $2m$ degrees of freedom. Denote the maximum of two chi-square random variables by $U = \max(X, Y)$. The distribution of U or some function of it $Z = (n/m)U/X_0$, where X_0 is a random variable distributed as $\chi^2_{(2n)}$ independent of both X and Y, is known as studentized maximum chi-square variables or the maximum of two correlated F-variables. These distributions arise in applied statistical problems when parent distributions are gamma, exponential, Pareto, Weibull, or Rayleigh. For example, in reliability analysis of parallel systems, a statement of the form $P(U < C_1)$ for some known constant C_1 is encountered. Testing the equality of the scale parameters of three exponential population

$$f(t,\theta_i) = \theta_i^{-1} e^{-t/\theta_i} \quad ,\theta_i, i = 1,2,3, \quad t > 0,$$

give rise to statements similar to $P(Z > C_2)$ for some constant C_2. The fundamental work of Gupta and Sobel (1962 a), (1962, b), Hartely (1983), Finney(1941), Nair(1984), Ramachandran(1958), David(1956), Krishna and Armitage(1964) focused on the distributions of forms of chi-squares, and Hamdy et al. (1987). Most of the work done in these papers, the standard transformation techniques and inverse probabilities, were central to the analysis. In this section, we intend to focus on the role of left-truncated negative binomial sampling techniques to determine the distribution of the random variables U and Z and their relation to the sampling procedure. The distribution of the maximum of the correlated F- variables is also relevant to analysis of variance, (see Johnson, Kotz and Balakrishnan, 1995, pp 352-355, for details)

Recall the random experiment given in Section 2.2 with equal probability of selecting either urn. Assume further that selecting a ball from urn A is associated with taking a realization x_{1i}, $i = 1,2,..., m$ on the random variable X_1, and selecting a ball from urn B is associated with taking a realization $y_{1j}, j = 1,2,..., m$ on the random variables Y_1. X_1 and Y_1 are independent and identically distributed each $\chi^2_{(2)}$. A sequence of independent trials is performed until one of the urns is completely exhausted and consequently the minimum is identified. Obviously, if the random variables X_1 and Y_1 have been observed s and r times respectively, then $X = \sum_{i=1}^{s} x_{1i}$, $Y = \sum_{j=1}^{r} y_{1j}$ are distributed according to $\chi^2_{(2s)}$ and $\chi^2_{(2r)}$ in that order. We stress the point that the urn which is exhausted first provides the minimum of two chi-square variates. In the sense that the random variable which has been completely observed is the minimum. Of course the urn which will be completely observed second will identified as the maximum of two chi-square variables. Naturally observing the maximum of two random variables is conditioned on determining the minimum of the two random variables in advance. Since it takes $m + k$ samples to identify the minimum, we can assume the maximum value k is m - 1. Let us now continue the process with flipping a fair coin. A ball is drawn from the remaining $m - k$ balls in the non-exhausted urn, upon obtaining a head, and at the same time we take a realization on the yet completely unobserved random variable. While if the coin turns up tail no further action is taken. Sequence of independent flipping are conducted until all balls in the yet non-exhausted urn are drawn out. Logically by then, the random variable $U = \max(X, Y)$ is completely observed in a random fashion. Hence the entire experiment is terminated. The random variable K in this case represents the number of times we were not able to draw a ball from the non-exhausted urn. Therefore, the random variable K assumes $k = m, m+1,....$ and has the probability mass function given by (2.8). Moreover, since the random variable K counts the number of times we were not able to select a ball from the non-exhausted urn with average realization $(u/2)$, the conditional probability mass function of the random variable K given U is given by the following Poisson probability mass function

$$P(K = k \mid U) = \frac{(u/2)^k e^{-u/2}}{k!}, \quad k = 0,1,2,... \tag{3.1}$$

Consequently, the joint probability density function of K and U is given by

$$P(k,u) = 2f(u)P(K = k \mid U), \tag{3.2}$$

where, $f(u)$ is the χ^2_{2m} probability density function of the random variable U. While, $P(K = k \mid U)$ is the conditional probability of K given U as defined in (3.1) for the event "no further action is taken". Therefore (3.2) yields

$$P(k,u) = \binom{m+k-1}{k}(1/2)^{m+k-1}\frac{u^{m+k-1}e^{-u}}{\Gamma(m+k)}, \quad k = m, m+1,.... \text{and}\ u > 0. \tag{3.3}$$

The marginal probability mass function of the random variable K is then obtained from (3.3) as

$$P(K = k) = \binom{m+k-1}{k}(1/2)^{m+k-1}\int_0^\infty \frac{u^{m+k-1}e^{-u}}{\Gamma(m+k)}du,$$
$$= \binom{m+k-1}{k}(1/2)^{m+k-1} \quad k = m, m+1,...$$

which is the probability mass function of the left-truncated negative binomial given in (2.8). Similarly the conditional probability density function of U given $K = k$ is

$$h(u \mid K = k) = \frac{u^{m+k-1} e^{-u}}{\Gamma(m+k)}, \quad u > 0. \tag{3.4}$$

which is a gamma probability density function with $m + k$ degrees of freedom. Finally, the marginal density function of U is given by

$$g(u) = \sum_{k=m}^{\infty} \binom{m+k-1}{k} (1/2)^{m+k-1} \frac{u^{m+k-1} e^{-u}}{\Gamma(m+k)} \quad u > 0. \tag{3.5}$$

which is an infinite sum of weighted gamma probability density functions. The identity in (2.6) can be used to check that $g(u)$ is a proper probability density function. We give the following theorem to summarize the above results.

Theorem

Let X and Y be iid $\chi^2_{(2m)}$ random variables. Further, let K be a random variable generated by a negative binomial sampling process to determine the distribution of the random variable $U = \max(X, Y)$. The following results hold

1. *The random variable K follows a left -truncated negative binomial distribution as in (2.8).*
2. *The conditional probability density function of the random variable U given K is gamma with $m + k$ degrees of freedom as in (3.4).*
3. *The marginal probability density function of U is weighted gamma as in (3.5).*
4. *For a real integer $r \geq 1$, whenever one exists,*

$$E(U^r) = E_K E_U(U^r \mid K = k) = E_K\left[\frac{(m+K+r-1)!}{(m+K-1)!}\right]$$

as given in (2.10).

In Section 4, the distribution of the maximum of two correlated F-variables is given; its moments are also given in a somewhat simple form.

4. NEGATIVE BINOMIAL SAMPLING AND THE DISTRIBUTION OF MAXIMUM OF CORRELATED-F

Let X_0 be a $\chi^2_{(2n)}$ random variable independent of both X and Y which were defined previously. The random variable $Z = (n/m)U/X_0$ is known as a studentized maximum chi-square variable or the maximum of two correlated F-variables. The distribution of Z arises in many statistical applications including ranking and selection of exponential distributions and reliability estimation of parallel systems. In this section we proceed to relate the distributions of Z to negative binomial sampling discussed in Section 3. The distributions of the maximum correlated F-variables is obtained by utilizing the representation of the joint probability function in (3.3) and the probability density function of the random variable X_0. Hence, the joint probability function of $K = k$ and Z is given by

$$P(k,z) = \binom{m+k-1}{k} \frac{2\,(m/n)^{m+k} z^{m+k-1}}{\beta(m+k,n)(1+(2m/n)z)^{m+n+k}}, \quad k = m, m+1, \ldots, z \geq 0. \tag{4.1}$$

Therefore, the marginal probability density function of Z is

$$f(z) = \sum_{k=m}^{\infty} \binom{m+k-1}{k} \frac{2\,(m/n)^{m+k} z^{m+k-1}}{\beta(m+k,n)(1+(2m/n)z)^{m+n+k}}, \quad z \geq 0. \tag{4.2}$$

The conditional probability density function of Z given K is also written as

$$f(z \mid K = k) = \frac{2^{m+k}\,(m/n)^{m+k} z^{m+k-1}}{\beta(m+k,n)(1+(2m/n)z)^{m+n+k}}, \quad z \geq 0 \tag{4.3}$$

and finally the r^{th} moment of the random variable Z, whenever exists, is given by

$$E(Z^r) = (n/m)^r \left[\binom{n-1}{r}\right]^{-1} \binom{m+r-1}{r} \left[1 + \sum_{k=m}^{m+r-1} (1/2)^{m+k+r-1} \binom{m+k+r-1}{k}\right],$$

from which we obtain the first two moments which are given as

$$E(Z) = n/(n-1)\left(1 + m(1/2)^{2m-1}\right), \quad n > 1.$$

$$E(Z^2) = n^2(m+1)/(m(n-1)(n-2))\left(1 + (1/2)^{2m-2}(2m+1)(3m+1)\right), \quad n > 2$$

5. COMPUTATIONAL ALGORITHMS AND THE CONSTRUCTION OF TABLES

In the present section we designed computational algorithms to provide critical values of the random variable U given in Section 3 and the random variable Z given in Section 4. First recall the probability mass function of the random variable U given in (3.5) and let C be the solution of the following integral equation for some given values of m and α such that

$$\alpha = \int_0^C g(u)\,du = \sum_{j=m}^{\infty} \binom{m+j-1}{j} (1/2)^{m+j-1} \int_0^C \frac{u^{m+j-1} e^{-u}}{\Gamma(m+j)}\,du \qquad (5.1)$$

We evaluate each term in the above infinite sum to determine the degrees of freedom for which the term is less than or equal 10^{-23}. Once we determine the degrees of freedom we need to include in our search, we use the bisection method to find C which satisfies (5.1) for the given α. Denoted by C_1 and $C2$, the inverse gamma function at m and the already determined degrees of freedom of the last term, call this, m_1, it follows by the monotonicity property of the gamma function that $C_1 \le C \le C_2$.. Hence, an iterative method of root finding and the bisection method can be used to locate the value of C accurate to 12 decimal places for given values of m and α. A Fortran program with the aid of the well known IMSL library routines are used to generate tables for the maximum chi-square random variable U for $m = 1(1)30$ and $35(5)250$ and values of $\alpha = 0.005$, 0.001, 0.025, 0.10, 0.90, 0.975, 0.99, and 0.995. We report some selected values of C with only three decimal places in Table 1. for illustration purpose. Other critical values are available by request.

Table 1.

m	Values of α									
	0.005	0.001	0.025	0.10	0.90	0.95	0.97	0.975	0.99	0.995
1	0.147	0.211	0.344	0.506	0.760	5.939	7.352	8.751	10.592	11.980
2	0.867	1.064	1.413	1.779	2.284	9.425	11.113	12.747	14.855	16.421
3	1.893	2.204	2.730	3.254	3.948	12.520	14.416	16.228	18.541	20.246
4	3.079	3.490	4.166	4.822	5.673	15.429	17.498	19.461	21.948	23.771
5	4.367	4.865	5.674	6.447	7.436	18.223	20.444	22.540	25.181	27.109
6	5.725	6.304	7.232	8.111	9.223	20.937	23.295	25.511	28.292	30.315
7	7.137	7.790	8.828	9.804	11.029	23.590	26.076	28.401	31.312	33.422
8	8.591	9.312	10.454	11.519	12.849	26.197	28.800	31.229	34.259	36.452
9	10.079	10.865	12.103	13.252	14.681	28.766	31.479	34.005	37.148	39.418
10	11.596	12.443	13.772	15.000	16.522	31.303	34.121	36.738	39.988	42.332
11	13.136	14.041	15.457	16.761	18.371	33.813	36.730	39.434	42.787	45.200
12	14.698	15.659	17.157	18.533	20.227	36.300	39.312	42.099	45.549	48.029
13	16.278	17.292	18.870	20.315	22.090	38.766	41.869	44.736	48.281	50.825
14	17.874	18.939	20.593	22.105	23.958	41.215	44.405	47.349	50.984	53.590
15	19.484	20.599	22.327	23.903	25.830	43.647	46.922	49.941	53.662	56.328
16	21.107	22.271	24.070	25.708	27.707	46.066	49.422	52.512	56.318	59.042
17	22.742	23.952	25.820	27.518	29.588	48.471	51.907	55.066	58.954	61.733
18	24.388	25.643	27.578	29.335	31.472	50.864	54.377	57.604	61.571	64.405
19	26.044	27.343	29.343	31.156	33.360	53.246	56.833	60.126	64.171	67.058
20	27.709	29.051	31.114	32.982	35.250	55.618	59.278	62.636	66.755	69.694
21	29.382	30.765	32.891	34.813	37.143	57.981	61.712	65.132	69.325	72.314
22	31.063	32.487	34.673	36.648	39.039	60.335	64.136	67.616	71.881	74.920
23	32.751	34.215	36.460	38.486	40.938	62.681	66.550	70.090	74.425	77.512
24	34.446	35.949	38.252	40.328	42.838	65.019	68.955	72.553	76.957	80.091
25	36.147	37.689	40.048	42.173	44.741	67.351	71.351	75.007	79.478	82.658
26	37.855	39.433	41.848	44.022	46.645	69.676	73.740	77.452	81.989	85.214
27	39.568	41.183	43.652	45.873	48.551	71.994	76.121	79.888	84.490	87.760
28	41.286	42.937	45.460	47.727	50.460	74.307	78.495	82.316	86.981	90.295
29	43.009	44.696	47.271	49.584	52.369	76.614	80.862	84.736	89.464	92.821
30	44.737	46.459	49.085	51.443	54.281	78.916	83.223	87.150	91.939	95.338
35	53.441	55.329	58.201	60.772	63.858	90.354	94.944	99.119	104.202	107.802
40	62.236	64.278	67.379	70.148	73.465	101.692	106.545	110.952	116.307	120.095
45	71.104	73.291	76.607	79.562	83.097	112.948	118.048	122.673	128.284	132.249
50	80.034	82.358	85.877	89.008	92.748	124.136	129.470	134.300	140.154	144.286

Second, to find the critical points $0 < C < \infty$ recall the probability density function of the random variable Z in (4.2). They are found by solving the following integral equation for some given values of n, m and α.

$$\alpha=\int_0^C f(z)\,dz=\sum_{j=m}^{\infty}\binom{m+j-1}{j}\int_0^C \frac{2\,(m/n)^{m+j}\,z^{m+j-1}}{\beta(m+j,n)(1+(2m/n)z)^{m+n+j}}\,dz \qquad (5.2)$$

We then make use of the transformation $W = (1+2m/nZ)^{-1}$ to restrict our search for C in the interval $(0,1)$. Therefore,

$$\alpha=\int_h^1 f(z)\,dz=\sum_{j=m}^{\infty}\binom{m+j-1}{j}(1/2)^{m+j-1}\int_h^1 \frac{w^{n-1}\,(1-w)^{m+j-1}}{\beta(m+j,n)}\,dw \qquad (5.3)$$

It follows that

$$(1-\alpha)=\sum_{j=m}^{\infty}\binom{m+j-1}{j}(1/2)^{m+j-1}\int_0^h \frac{w^{n-1}\,(1-w)^{m+j-1}}{\beta(m+j,n)}\,dw\,, \qquad (5.4)$$

where $C = n(1-h)/2mh$. Where we used (2.9) to reach the above result.

Similar algorithms lead to solutions of using (5.1) iterative methods of root finding and the bisection method to locate h for given values of $\alpha = 0.005, 0.001, 0.025, 0.10, 0.90, 0.95, 0.97, 0.975, 0.99, 0.995$, $m = 1(1)30$ and $35(5)250$ and $n = 1(1)30$ and $35(5)250$. We report some critical points of Z in the following Table 2. Other values are available by request.

Table 2.

m	n									
	1	2	3	4	5	6	7	8	9	10
1	0.206	0.224	0.232	0.236	0.239	0.241	0.243	0.244	0.245	0.246
2	0.296	0.343	0.366	0.381	0.391	0.398	0.404	0.408	0.412	0.414
3	0.326	0.387	0.420	0.441	0.456	0.467	0.475	0.482	0.487	0.492
4	0.340	0.409	0.448	0.473	0.491	0.504	0.514	0.523	0.530	0.536
5	0.348	0.422	0.464	0.492	0.512	0.527	0.539	0.549	0.557	0.563
6	0.352	0.430	0.474	0.504	0.526	0.542	0.556	0.566	0.575	0.583
7	0.355	0.435	0.481	0.513	0.536	0.553	0.567	0.579	0.588	0.597
8	0.356	0.438	0.486	0.519	0.543	0.561	0.576	0.588	0.598	0.607
9	0.357	0.441	0.490	0.524	0.548	0.568	0.583	0.596	0.606	0.615
10	0.358	0.443	0.493	0.527	0.553	0.572	0.588	0.601	0.612	0.622
11	0.359	0.444	0.495	0.530	0.556	0.576	0.592	0.606	0.617	0.627
12	0.359	0.445	0.496	0.532	0.558	0.579	0.596	0.610	0.621	0.631
13	0.359	0.446	0.498	0.534	0.561	0.582	0.599	0.613	0.625	0.635
14	0.359	0.446	0.499	0.535	0.562	0.584	0.601	0.615	0.627	0.638
15	0.359	0.447	0.500	0.536	0.564	0.585	0.603	0.617	0.630	0.640
16	0.359	0.447	0.500	0.537	0.565	0.587	0.605	0.619	0.632	0.643
17	0.359	0.447	0.501	0.538	0.566	0.588	0.606	0.621	0.634	0.645
18	0.359	0.447	0.501	0.539	0.567	0.589	0.607	0.622	0.635	0.646
19	0.359	0.447	0.501	0.539	0.568	0.590	0.608	0.623	0.636	0.648
20	0.358	0.447	0.502	0.540	0.568	0.591	0.609	0.624	0.638	0.649
21	0.358	0.447	0.502	0.540	0.569	0.591	0.610	0.625	0.639	0.650
22	0.358	0.447	0.502	0.540	0.569	0.592	0.611	0.626	0.639	0.651
23	0.358	0.447	0.502	0.540	0.569	0.592	0.611	0.627	0.640	0.652
24	0.358	0.447	0.502	0.541	0.570	0.593	0.612	0.627	0.641	0.653
25	0.357	0.447	0.502	0.541	0.570	0.593	0.612	0.628	0.642	0.653
26	0.357	0.447	0.502	0.541	0.570	0.593	0.613	0.628	0.642	0.654
27	0.357	0.447	0.502	0.541	0.570	0.594	0.613	0.629	0.643	0.655
28	0.357	0.447	0.502	0.541	0.571	0.594	0.613	0.629	0.643	0.655
29	0.357	0.447	0.502	0.541	0.571	0.594	0.613	0.630	0.643	0.655
30	0.357	0.447	0.502	0.541	0.571	0.594	0.614	0.630	0.644	0.656
35	0.356	0.446	0.502	0.541	0.571	0.595	0.615	0.631	0.645	0.657
40	0.355	0.446	0.501	0.541	0.571	0.595	0.615	0.632	0.646	0.658
45	0.354	0.445	0.501	0.541	0.571	0.595	0.615	0.632	0.646	0.659
50	0.353	0.444	0.500	0.540	0.571	0.595	0.615	0.632	0.646	0.659

REFERENCES

[1] Ahuja, J. C. (1971). "Distribution of the sum of independent decapitated negative binomial variables", **Annals of Mathematical Statistics**, 42, 383-384.

[2] Boswell, M. T., and Patil, G. P. (1970). **Chance mechanisms generating the negative binomial distributions, *Random Counts in Scientific Work*, 1: *Random Counts in Models and Structures***, G. P. Patil (editor), 3-22. University Park: Pennsylvania State University Press.

[3] David, H.A. "On the Application to Statistics of an Elementary Theorem in Probability", **Biometrika**, vol. 43 (1956) 85-91.

[4] Finney, D.J. "The Joint Distribution of Variance Ratios based on a Common Error Square", **Annals of Eugenics**, vol. 11 (1941).

[5] Ghosh, M.N. "Simultaneous Tests of Linear Hypothesis", **Biometrika**, vol. 42 (1955) 441-44

[6] Gnanadesikan, R. "Equality of more than Two Variances and of more than Two Dispersion Matrices against Certain Alternatives", **Ann. Math. Statis.**, vol. 30 (1959) 177-184.

[7] Gupta, S.S., and Sobel, M. "On selecting a subset containing the population with the smallest variance", **Biometrika** 49 (1962) 495-508.

[8] Gupta, S.S. "On a Selection and Ranking Procedure for Gamma Populations", **Annals of Institute of Statisitcal Mathematics**, vol. 14 (1963) 199-216.

[9] Hamdan, M. A. (1975). "Correlation between the numbers of two types of children when the family size distribution is zero-truncated negative binomial", **Biometrics**, 31, 765-766.

[10] Hamdy, H. ,Son, M. And AlMahmeed, M (1987). "A note on the distribution of maximum correlated F. ", **Proceedings of ASA Statistial Computing Section**, pp 406-411.

[11] Hamdy, H. I. and Son, M.S. (2005). **The role of negative binomial sampling in determining the distribution of Ch-squre.** Submitted for publication

[12] Hartly, H. O., "Studentization and large sample theory", **Journal of the Royal Statistical Statistical Society (Supplement)** 5 (1938) 80-88.

[13] Hartley, H.O. "The Maximum F-ratios as a Short Cut Test for Hetrogeneity of Varince.", **Biometrika**, vol. 37 (1950) 308-312.

[14] Hartley, H. O. "The use of range in analysis of variance.", **Biometrika**, 37, 271-280.

[15] Johnson, N. L., Kotz, S., and Kemp, A. W. (1992). **Univariate Discrete Distributions.** John Wiley & Sons, Inc.

[16] Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). **Continuous Univariate Distributions.** John Wiley & Sons, Inc.

[17] Kemp, C. D. (1967). 'Stuttering-Poisson' distribution", **Journal of Statistical and Social Enquiry Society of Ireland**, 21(5), 151-157.

[18] Kendall, D. G. (1949). "Stochastic processes and population growth", **Journal of the Royal Statistical Society, Series B**, 11, 230-282.

[19] Krishnaiah, P.K. and Armitage, J.V. "Tables for the Studentized Largest Chi-square Distribution and their *Applications*", **Aerospace Research Laboratories** , Wright Patterson Air Force (1964) 64-188.

[20] McKendrick, A. G. (1914). "Studies on the theory of continuous probabilities, with special reference to its bearing on natural phenomena of progressive nature", **Proceedings of the London Mathematical Society**, 2, 13, 401-416.

[21] Nair, K., "The Studentized form of the extreme mean square test in the analysis of variance", **Biometrika** 35 (1948) 16-31.

[22] Pascal, B. (1679). **Varia opera Mathematica D. Petri de Fermat**: Tolossae.

[23] Ramachandran, K.V. "On the Simultaneous Analysis of Variance Test", **Ann. Math. Stat.** 27 (1956) 521-528.

[24] Ramachandran, K. V., "On the Studentized smallest chi-square", **Journal of American Statistical Statistical Association** 53 (1958) 868-872.

[25] Reed, S. W., and Reed, S. C. (1965). **Mental Retardation: A Family Study**, Philadelphia: W. B. Saunders.

[26] Roy, S.N. "A Note on Some Further Results in Simultaneous Confidence Interval Estimation", **Ann. Math. Stat.**, vol 27(1956) 856-858.

[27] Roy, S.N. and Gnanadesikan, R. "Further Contribution to Multivariate Confidence Bounds", **Biometrika**, vol.44 (1957) 399-410

Mun-Shig Son

He received M.S. and Ph.D. degree in Statistics, Oklahoma State University, U.S.A . He is currently a professor in college of engineering and mathematical sciences at University of Vermont in U.S.A.