Order Statistics for Negative Binomial Distribution

Tabark Ayad Ali **Hussain Ahmed Ali**
University of Babylon /College of Education for Pure Sciences
sdarser@gmail.com

Abstract

In this paper, we considers the order statistics from negative binomial distribution .which have moment of all positive orders.
We take as example for $x_3$to find the mean and variance. We define aformula for B . The incomplete beta function of Negative binomial distribution is $I_{1-B(x)}(i, M)$, But The incomplete beta function of binomial distribution is $I_{B(x)}(i, M-i+1)$.

The purpose of this paper is to study the distribution of x_i and $R = x_j - x_i$ $(x_j, x_i, i > j)$. The results are obtained here can be compared to those for a continuous variate.
Keywords : Order Statistics, Negative Binomial, Joint Distribution, Mean, Variance.

الخلاصة

في هذا البحث ناقشنا الاحصاءات المرتبة لتوزيع ثنائي الحدين السالب . الذي يمتلك عزوم لكل الرتب الموجبة . حيث تم وضع صيغ لحساب الوسط والتباين عندما $P(N) = 1$ وكذلك وضعت صيغ لحساب دالة بيتا غير الكاملة لتوزيعي ثنائي الحدين السالب وتوزيع ذو الحدين . اخذنا مثال بالنسبة الى x_3 لحساب الوسط والتباين . هدف هذا البحث هو دراسة توزيع i و $R = x_j - x_i$ $(x_j, x_i, i > j)$.

النتائج التي توصلنا لها يمكن ان تحسب لهذا الموضوع للحالة المستمرة .

الكلمات المفتاحية:الاحصاءات المرتبة، ثنائي الحدين السالب، التوزيع المشترك، التوقع، التباين.

1. Introduction

Let x_i $(i = 0,1,2, \dots, M)$ be the number of successes in N independent trials from anegative binomial distribution with p as the probability of success in each trial .

Let x_i be arranged in order to give

$$x_1 \le x_2 \le \cdots \le x_i \le \cdots \le x_M$$

The discrete variable x_i *or* $x_{(i)}$ take values $1,2, \dots$ $(i = 1,2, \dots, M)$.
Table are provided giving the cumulative distribution and the expected value and the variance of $x_{(1)}$ and x_M .joint distribution of $x_{(i)}$ and $x_{(j)}$ $(i < j)$ is obtained .
The distribution of $x_j - x_i$ $(j > i)$ is also derived .Given observations, we can see that there are $i\,(M - i + 1)$ different and mutually exclusive cases of the type $i - 1 - k$ observations below x_i ,
$(k + m + 1)$ observations of x_i and $n - i - m$ observations above x_i , with respective probabilities $P(x_i - l), p(x_i)$ *and* $1 - P(x_i)$,
$(k = 0,1,2, \dots i - 1$ *and* $m = 0,1,2, \dots, M - i)$.

2.Literature Review

The negative binomial distribution (NBD) is one of the most useful probability distributions and has been successfully employed to model a variety of natural phenomena. It has been used to construct useful models in many substantive fields: in accident statistics (Greenwood and Yule(1920) , Arbous and Kerrich(1951), Weber(1971)) . in birth-death processes (Furry,1937) ; Kendall 1949) .in psychology (Sichel, 1951).

The first work in actuarial literature that has come to my attention involving the negative binomial was by Keffer in 1929 in connection with a group life experience rating plan. He developed the theory in relationship to the relative dispersion of loss ratios about their true mean.

A. L. Bailey first utilized the negative binomial in the Proceedings of the Casualty Actuarial Society in 1950. The distribution resulting from
repeated trials of this experiment will be in thc form of a negative binomial distribution . This model is suggested by the mathematical development of Feller (1957) .

Creel and Loomis (1990, 1991) estimated truncated and untruncated Poisson and negative binomial models for California deer hunting and derived welfare measures from these models see (Miaou, 1994; ; Hadi et al., 1995; Shankar *et al*.(1995) Poch and Mannering, (1996).

In addition, zero-inflated Poisson and zeroinflated negative binomial models were also employed to analyze accident frequencies by Shankar *et al.* (1997) . also employed to analyze accident frequencies by Lee and Mannering (2002) and Lee *et al.* (2002) .

In biology (Anscombe(1950); Bliss and Fisher (1953); Anderson (1965); Boswell and Patil (1970); Elliot (1977) . In ecology (Martin and Katti (1965); Binn (1986); White and Bennetle (1996)). in entomology (Wilson and Room(1983) . In epidemiology (Byers et al. (2003) . in information sciences (Bookstein(1997) . In meterology (Sakamato (1972) .

In addition, many other physical and biological applications have been described by (Biggeri (1998) and Eke et al. (2001) .

3. Moment of order statistics

Gupta ,S.Shanti .and S. Panchspakesan(1967) "Department of Statistics Division of Mathematical Sciences Mimeograph Series No. 120"

$$Nb(x) = P(x) \equiv P(N; p, x) = \binom{N-1}{x-1} p^x (1-p)^{N-x} \ ,$$
$$N = x, x+1, \dots \ , x > 0$$

Where $Nb(x)$ probability density function of Negative binomial distribution.

$$B(s) = \sum_{x=r}^{s} P(x) \, .$$

$$B(p,q) = (\Gamma(p)\Gamma(q))/(\Gamma(p+q))$$

$$I_p(a,b) = \frac{1}{B(a,b)} \int_0^p u^{a-1} (1-u)^{b-1} du \, .$$

Let $p_i(x)$ be the probability that the ith order statistic $x_{(i)}$ is equal to x and let $P_i(x) = p\{x_{(i)} \le x\}$ be the $c.d.f. of \ x_{(i)}$.

$$p_i(x) = \sum_{k=0}^{i-1} \sum_{m=0}^{M-i} \frac{M! \ \{B(x-1)\}^{i-1-k} \{p(x)\}^{k+m+1} \{1-B(x)\}^{M-i-m}}{(i-1-k)! \, (k+m+1)! \, (M-i-m)!} \qquad (3.1)$$

Where $B(x-1) = 0 \ for \ x = 0$. This can be rewritten as

$$p_i(x) = i \binom{M}{i} \int_{B(x-1)}^{B(x)} \omega^{i-1} (1-\omega)^{M-1} d\omega \qquad (3.2)$$

$$= I_{1-B(x)}(i, M) - I_{1-B(x-1)}(i, M).$$

Further

$$p_i(x) = i\binom{M}{i}\int_0^{B(x)} \omega^{i-1}(1-\omega)^{M-1}d\omega = I_{1-B(x)}(i, M) \quad (3.3)$$

$$E(x_{(i)}) = \sum_{x=0}^{N-1}[1 - p_i(x)] \quad (3.4)$$

$$E(x^2{}_{(i)}) = 2\sum_{x=0}^{N-1} x[1 - p_i(x)] + \sum_{x=0}^{N-1}[1 - p_i(x)] \quad (3.5)$$

$$var(x_i) = E(x^2{}_{(i)}) - [E(x_{(i)})]^2 \quad (3.6)$$

Example

let (x_0, x_1, x_2, x_3) be order statistic , i=3 , $p = q = \frac{1}{2}$, $x_3 = 3$, find the mean and veriance of x_3 ? when $M = 3$, $N = 4$, $Z = 2$.
where $\omega = ZP(x_i) + (x_i - 1)$

$$P(x_3 = 3) = \binom{2}{0}\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^2 = \frac{1}{8}$$

$$P(x_3 - 1) = P(2) = \binom{1}{0}\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4}$$

$$\omega = 2 \cdot \frac{1}{8} + \frac{1}{4} = \frac{1}{2}$$

$$P(x) = i\binom{M}{i}\int_0^{B(x)} \omega^i (1-\omega)^{M-i}d\omega$$

$$B(x) = \sum_{r=i}^{x} p^i(1-p)^{i-r} \quad , i = 0,1,2,3$$

$$B(0) = 0 , B(1) = \frac{1}{2}, B(2) = \frac{5}{8}, B(3) = \frac{13}{8}$$

$$P_3(0) = 0$$

$$P_3(1) = 3\binom{3}{3}\int_0^{\frac{1}{2}} \left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^0 d\omega = \frac{3}{16}$$

$$P_3(2) = 3\binom{3}{3}\int_0^{\frac{5}{8}} \left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^0 d\omega = \frac{15}{64}$$

$$P_3(3) = 3\binom{3}{3}\int_0^{\frac{13}{8}} \left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^0 d\omega = \frac{39}{64}$$

$$E(x_3) = \sum_{x=0}^{4-1} [1 - P_3(x)].$$

$$= [1 - P_3(0)] + [1 - P_3(1)] + [1 - P_3(2)] + [1 - P_3(3)] \ .$$

$$= 1 + \left(1 - \frac{3}{16}\right) + \left(1 - \frac{15}{64}\right) + \left(1 - \frac{39}{64}\right) = 1 + \left(\frac{13}{16}\right) + \left(\frac{49}{64}\right) + \left(\frac{25}{64}\right)$$

$$\frac{64 + 52 + 74}{64} = \frac{190}{64}$$

$$E\left(\ {x^2}_{(3)}\right) = 2 \sum_{x=0}^{N-1} x\,[1 - P_3(x)] + \sum_{x=0}^{N-1} [1 - P_3(x)]$$

$$\sum_{x=0}^{3} x\,[1 - P_3(x)] = 0.[1] + 1.[\frac{13}{64} + \frac{49}{32} + \frac{75}{64}] = \frac{52 + 98 + 75}{64} = \frac{225}{64}$$

$$\boldsymbol{E}\left(\boldsymbol{x}^2(\mathbf{3})\right) = 2.\frac{225}{64} + \frac{190}{64} = \frac{640}{64}$$

$$\boldsymbol{var}\ (\boldsymbol{x}_3) = E\left(x^2(3)\right) - [E(x_3)]^2$$

$$= \frac{640}{64} - \frac{36100}{4096} = \frac{4860}{4096}$$

4 . Joint distribution of $\boldsymbol{x_{(i)}}$ *and* $\boldsymbol{x_{(j)}}$, $\boldsymbol{i < j}$

BIGGERI, A. (1998), FISHER,R.A. (1941), Gupta ,S.Shanti .and S. Panchspakesan(1967)

Let $p_{i,j}(x, y)(i < j)$ be the probability that $x_{(i)}$ is equal to x and $x_{(j)}$ is equal to y and let $p_{i,j}(x, y) = p\left(x_{(i)} \leq x, x_{(j)} \leq y\right)$. If $x \geq y$,

$$p_{i,j}(x, y) = p\left(x_{(j)} \leq y\right)$$

$$= j\binom{M}{j} \int_0^{B(y)} u^{i-1}(1 - u)^{M-1} du\ ) \quad (4.1)$$

By repeated application of the results

$$\sum_{s=a}^{n} \binom{n + s - 1}{s} p^n (k - p)^s = \frac{1}{B(a, n)} \int_0^{p} u^{a-1}(1 - u)^{n-1} du$$

and

$$\sum_{s=0}^{b} \binom{n + s - 1}{s} p^n (k - p)^s = \frac{1}{B(b + 1, n)} \int_0^{p} u^{b}(1 - u)^{n-1} du$$

Where $0 \le p < 1$ and $p < k$; we obtain

$\boldsymbol{p_{i,j}(x,y)}$

$$= i\binom{M}{i}\int_0^{B(x)} u^{i-1}(1-u)^{M-1}du - \frac{M!}{(i-1)!\,(j-i-1)!\,(M-1)!}\int_{B(y)}^{1} dv \int_0^{B(x)} \omega^{i-1}\,(v-\omega)^{j-i-1} (1-v)^{M-1}d\omega \qquad (\mathbf{4.2})$$

$$\omega = zP(x_i) + P(x_i - 1) \quad , \quad v = P(x_j) - y\,P(x_j)$$

we get the relation

$$\frac{M!}{(i-1)!(j-i-1)!(M-j)!}\int_{B(x)}^{1} dv \int_0^{B(x)} \omega^{i-1}(v-\omega)^{j-i-1}(1-v)^{M-1}d\omega \quad (4.3)$$

$$= i\binom{M}{i}\int_0^{B(x)} u^{i-1}(1-u)^{M-1}du - i\binom{M}{i}\int_0^{B(x)} u^{j-1}(1-u)^{M-1}du$$

$$= I_{1-B(x)}(i,M) - I_{1-B(x)}(j,M) \qquad (4.4)$$

$$E\left(x_{(i)}x_{(j)}\right) = \sum_{x=0}^{N}\sum_{y=0}^{N} xy\,p_{i,j}(x,y)$$

$$= \sum_{x=0}^{N} x^2\,p_{i,j}(x,x) + \sum_{x=0}^{N-1}\sum_{y=x+1}^{N} xy\,p_{i,j}(x,y) \qquad (4.5)$$

so

$$cov\left(x_{(i)},x_{(j)}\right) = \sum_{x=0}^{N} x^2\,p_{i,j}(x,x) + \sum_{x=0}^{N-1}\sum_{y=x+1}^{N} xy\,p_{i,j}(x,y) - \left\{\sum_{x=0}^{N-1}[1-p_i(x)]\right\}\left\{\sum_{x=0}^{N-1}[1-p_j(x)]\right\} \quad (\mathbf{4.6})$$

5 . Distribution of $\boldsymbol{Y_{i,j} = X_{(j)} - X_{(i)} (j > i)}$

BIGGERI, A. (1998), FISHER,R.A. (1941), Gupta ,S.Shanti .and S. Panchspakesan(1967)

$Y_{i,j}$ represents a generalized range and can take values $0,1,\ldots\ldots,N$ .for $r \ge 0$,

$$P\left(Y_{i,j} = r\right) = \sum_{k=0}^{n-r}\frac{M!}{(i-1)!(j-i-1)!(M-1)!}\int\int_A \omega^{i-1}(V-\omega)^{j-i-1}\,(1-V)^{M-1}dvd\omega \quad (5.1)$$

Where A is the region given by

$$\begin{cases} v \ge \omega \\ B(k) \ge \omega \ge B(k-1) \\ B(k+r) \ge v \ge B(k+r-1) \end{cases}$$

This can be rewritten as

$$p\{Y_{i,j}=r\}$$
$$=\begin{cases}\sum_{k=0}^{N}\frac{M!}{(i-1)!\,(j-i-1)!\,(M-1)!}\int_{B(k-1)}^{B(k)}d\omega\int_{\omega}^{B(k)}\omega^{i-1}(V-\omega)^{j-i-1}(1-V)^{M-1}dv & ,r=0 \quad (\mathbf{5.2}) \\ \sum_{k=0}^{N-r}\frac{M!}{(i-1)!\,(j-i-1)!\,(M-1)!}\int_{B(k-1)}^{B(k)}d\omega\int_{B(k+r-1)}^{B(k+r)}\omega^{i-1}(V-\omega)^{j-i-1}(1-V)^{M-1}dv & ,r>0\ ) \\ & (\mathbf{5.3})\end{cases}$$

so

$$E(Y_{i,j})=\sum_{r=1}^{N}r\sum_{k=0}^{N-r}\frac{M!}{(i-1)!\,(j-i-1)!\,(M-1)!} \qquad (5.4)$$
$$\int_{B(k-1)}^{B(k)}d\omega\int_{B(k+r-1)}^{B(k+r)}\omega^{i-1}\,(V-\omega)^{j-i-1}(1-V)^{M-1}dv \qquad (5.5)$$

But we also know that

$$E(Y_{i,j})=E(X_{(j)})-E(X_{(i)})$$
$$=\sum_{x=0}^{N-1}i\binom{M}{i}\int_{0}^{B(x)}\omega^{i-1}(1-\omega)^{M-1}d\omega-\sum_{x=0}^{N-1}j\binom{M}{j}\int_{0}^{B(x)}\omega^{j-1}(1-\omega)^{M-1}d\omega$$
$$=\sum_{x=0}^{N-1}[I_{1-B(x)}(i,M)-[I_{1-B(x)}(j,M)] \qquad (5.6)$$

(5.5) and (5.6) lead to the relation

$$\sum_{r=1}^{N}\sum_{k=0}^{n-r}r\,\frac{M!}{(i-1)!\,(j-i-1)!\,(M-1)!}.$$
$$\int_{B(k-1)}^{B(k)}d\omega\int_{B(k+r-1)}^{B(k+r)}\omega^{i-1}(V-\omega)^{j-i-1}(1-V)^{M-1}dv$$
$$=\sum_{x=0}^{N-1}[I_{1-B(x)}\,(i,M)-[I_{1-B(x)}(j,M)] \qquad (5.7)$$

6. Conclusions

A . Since $P(N)=1$,

$$E(x_i)=\sum_{x_i=0}^{N}x_i p*(x_i)=N-\sum_{x_i=0}^{N-1}P*(x_i)$$

where $P*(x_i)$ is defined by
$p_i(x) = i\binom{M}{i}\int_0^{B(x)} \omega^{i-1}(1-\omega)^{M-1}d\omega = I_{1-B(x)}(i,M)$.

Hence, we have

$$E(x_i) = \sum_{x_i=0}^{N-1}\{1 - p*(x_i)\}$$

For the variance of x_i , we have

$$E(x_i{}^2) = \sum_{x_{i=0}}^{N} x_i{}^2\, p*(x_i) = N^2 - \sum_{x_i=0}^{N-1}(2x_i+1)P*(x_i)$$

That is,

$$E(x_i{}^2) = 2\sum_{x_{i=0}}^{N-1} x_i\{1 - P*(x_i)\} + E(x_i)$$

Hence, the variance of x_i, is

$$V(x_i) = 2\sum_{x_{i=0}}^{N-1} x_i\{1 - P*(x_i)\} + E(x_i)\{1 - E(x_i)\}$$

B . The incomplete beta function of Negative binomial distribution is

$$\boldsymbol{I_{1-B(x)}(i,M)}$$

But

The incomplete beta function of binomial distribution is

$$I_{B(x)}(i, M-i+1)$$

7. References

ANDERSON, F.S. 1965: “The negative binomial distribution and the sampling of insect population”, Proceedings of XII International Conference on Entomology, 395-402.

ANSCOMBE, F.J. 1950: “Sampling theory of negative binomial and logarithmic series distributions”, Biometrika, 37, 358-382.

BIGGERI, A. 1998: “Negative binomial distribution”, In: Encyclopedia of Biostatistics, 4, 2962-2967. (Eds. P. Armitage and T. Colton). John Wiley: New York.

BINNS, M. R. 1986: “Behavioral dynamics and the negative binomial distribution”, Oikos, 47, 315-318.

BLISS, C. I. and R.A. FISHER 1953: “Fitting the negative binomial distribution to biological data and note on the efficient fitting of negative binomial”, Biometrics, 9, 176-200. JOHNSON, N.L.; S. KOTZ and A. KEMP 1992: Discrete Distributions, Second edition. John Wiley: New York.

Griffiths, R. C. and R. K. Milne 1987. A class of infinitely divisible multivariate negative binomial distributions. J. Multivariate Anal. 22, 13–23.

Gupta ,S.Shanti.and S. Panchspakesan1967 ,"order statistics Arising from Independent Binomial populations " , Department of Statistics Division of Mathematical Sciences Mimeograph Series No. 120 .