Journal of the Korean
Data & Information Science Society
2006, Vol. 17, No. 4, pp. 1405 ~ 1412

The UMVUE and MLE of the Tail Probability in Discrete Model

Jungsoo Woo[1)]

Abstract

We shall derive the UMVUE of the tail probability in Poisson, Binomial, and negative Binomial distributions, and compare means squared errors of the UMVUE and the MLE of the tail probability in each case.

Keywords : Poisson, Binomial, Negative binomial, Tail probability

1. Introduction

So far many authors have considered the UMVUE and the MLE of the reliability in the continuous distributions. And also recently Kim(2006), Lee(2006), and Lee & Won(2006) studied inferences on the reliability in an exponentiated uniform distribution and an exponential distribution.

For an example, since the curtate-future-life-time X has non-negative integer value, it is natural in actuarial studies to consider the tail probability of a discrete random variable to apply the reliability to evaluation of life insurance premium(see Bowers et al(1997)).

Here we shall consider the reliability in such important discrete random variables as Poisson, binomial, and negative binomial distribution, and derive UMVUE of the reliability in each case, and hence consider comparisons of the UMVUE and MLE of the reliability in each distribution for special case.

1.1 Poisson Model

Let X and Y be independent Poisson random variables with parameters λ_1

1) Professor, Department of Statistics, Yeungnam University, Gyeongsan, 712-749, Korea
E-mail: jswoo@ynu.ac.kr

and λ_2 , respectively.

Then, from Theorem 14 on p.212 in Rohatgi(1976) and the formula 3.381(3) in Gradshteyn & Ryzhik(1965), the reliability can be shown as follows:

$$P(Y \leq X) = \sum_{x=0}^{\infty} e^{-\lambda_1} \frac{\lambda_1^x}{(x!)^2} \cdot \Gamma(x+1, \lambda_2) \ ,$$

where $\Gamma(a,z)$ is an incomplete gamma function.

Next we shall consider the UMVUE of the reliability in the Poisson distribution.

Assume $X_1.X_2, \cdot\cdot\cdot, X_n$ be independent Poisson random variables with parameter λ . Then it is well-known that $S = \sum_{i=1}^{n} X_i$ is a complete sufficient statistics for λ ,

Define a statistic by $u(X_1) = \begin{cases} 0, & \text{if} X_1 > t_0 \\ 1, & \text{if} X_1 \leq t_0 \end{cases}$,

then $u(X_1)$ is an unbiased estimator of $P(X \leqq t_0)$ in the Poisson distribution and hence, from Lehmann-Scheffe Theorem in Rohatgi(1976).

$E(u(X_1) \mid S)$ is a UMVUE of $P(X \leq t_0)$ in the Poisson distribution.

Since the conditional pdf of X_1 given S=s can be obtained as:

$$P(X_1 = t \mid S = s) = \binom{s}{t} \cdot (1 - \frac{1}{n})^{s-t} \cdot (\frac{1}{n})^t \ , \text{ t=0.1.2....,s}$$

In fact, it's clear that $\sum_{t=0}^{s} \binom{s}{t} (1 - \frac{1}{n})^{s-t} (\frac{1}{n})^t = 1.$

Therefore,

$$E(u(X_1) \mid S) = \sum_{t=0}^{[t_0]} \binom{S}{t} \cdot (1 - \frac{1}{n})^{S-t} \cdot (\frac{1}{n})^t \text{ is the UMVUE of } P(X \leqq t_0) \qquad (1.1)$$

where $S = \sum_{i=1}^{n} X_i$.

and also we can derive the UMVUE of the reliability $P(X > t_0) = 1 - P(X \leq t_0)$ from the result (1.1).

Especially if $0 < t_0 < 1$, then $P(X \leqq t_0) = e^{-\lambda}$. From the result (1.1), it's well-known in Rohatgi(1976) that the UMVUE of $P(X \leq t_0) = e^{-\lambda}$ is

$(1-\frac{1}{n})^{\sum_{i=1}^{n} X_i}$, from the moment generating function(mgf) of the Poisson random variable in Rohatgi(1976), the variance of the UMVUE is obtained as:

$$\exp(-2\lambda) \cdot [\exp(\frac{\lambda}{n})-1]$$

While, the MLE of $P(X \leq t_0) = e^{-\lambda}$ is

$$P(\widehat{X \leq t_0}) = \exp(-\hat{\lambda}\ ) = \exp(-\frac{1}{n}\sum_{i=1}^{n} X_i) \ ,$$

the expectation and 2nd moment of the MLE are:

$$E[P(\widehat{X \leq t_0})] = \exp(\, n\lambda \cdot [\exp(-1/n)-1]\ ) \text{ and}$$
$$E[P(\widehat{X \leq t_0})^2] = \exp(\, n\lambda \cdot [\exp(-2/n)-1]\ )$$

When $\lambda = 1$ and n=10, 20, and 30, Table 1 shows MSE of the UMVUE and the MLE of $P(X \leqq t_0) = e^{-\lambda}$,

<Table 1> MSE of the UMVUE and the MLE of $P(X \leqq t_0) = e^{-\lambda}$ when $0 < t_0 < 1$ and $\lambda = 1$

n	10	20	30
UMVUE	.01423	.00694	.00459
MLE	.04737	.00701	.00462

Table 1 shows that the UMVUE of $P(X \leqq t_0) = e^{-\lambda}$ is more efficient than the MLE, and hence we can more recommend the UMVUE than the MLE of $P(X \leq t_0)$ in the Poisson distribution.

1.2 Binomial Model

Let X and Y be independent binomial random variables with parameters (m, p_1) and (n, p_2) , respectively,

Then, from Theorem 19 on p.226 in Rohatgi(1976) and the formulas 3.391 & 8.392 in Gradshteyn & Ryzhik(1965), the reliability in the binomial distribution can be shown as follows:

$$P(Y \leq X) = 1 - p_2 + \sum_{y=1}^{n} \binom{n}{y} p_2^y \cdot (1-p_2)^{n-y} \cdot I_{p_1}(y, m-y+1). \quad \text{if} m \geqq n \ ,$$

where $I_p(a,b) = \dfrac{B_p(a,b)}{B(a,b)}$ is an incomplete beta function.

If n>m, then we can obtain it similarly as follow.

$$P(Y \geqq X) = 1 - p_1 + \sum_{y=1}^{m} \binom{m}{y} p_1^y (1-p_1)^{m-y} \cdot I_{p_2}(y, n-y+1). \quad \text{if } n > m$$

Especially if m=n=1, that is, X and Y are independent Bernoulli random variables with parameters $p_1 \text{and} p_2$, respectively, then

$$P(Y \leq X) = (1-p_1)(1-p_2) + p_1 \ .$$

Next we shall consider the UMVUE of the reliability in the binomial distribution.

Assume $X_1 . X_2, \cdot \cdot \cdot X_n$ be independent binomial random variables each having parameter (k, p) when k is known. Then it is well-known that $S = \sum_{i=1}^{n} X_i$ is a complete sufficient statistics for p,

By the same method in the Poisson distribution of section 1.2.

Define a statistic by $u(X_1) = \begin{cases} 0, & \text{if } X_1 > t_0 \\ 1, & \text{if } X_1 \leq t_0 \end{cases}$,

then $u(X_1)$ is an unbiased estimator of $P(X \leq t_0)$ in the binomial distribution and hence, from Lehmann-Scheffe Theorem in Rohatgi(1976), $E(u(X_1) \mid S)$ is a UMVUE of $P(X \leq t_0)$ in the binomial distribution.

Since the conditional pdf of X_1 given S=s can be obtained as:

$$P(X_1 = t \mid S = s) = \binom{k}{t} \cdot \binom{(n-1)k}{s-t} \Big/ \binom{nk}{s} \ , \ \text{t=0.1.2....,k}$$

In fact, it's clear that $\sum_{t=0}^{k} \binom{k}{t} \cdot \binom{(n-1)k}{s-t} \Big/ \binom{nk}{s} = \binom{nk}{s} \Big/ \binom{k}{s} = 1$.

Therefore, a UMVUE of $P(X \leq t_0)$ in the binomial distribution can be obtained as:

$$E(u(X_1) \mid S) = \sum_{t=0}^{t_0} \binom{k}{t} \cdot \binom{(n-1)k}{S-t} \Big/ \binom{n \cdot k}{S} \text{ is a UMVUE of } P(X \leq t_0) \tag{1.2}$$

and also we can derive the UMVUE of the reliability $P(X > t_0) = 1 - P(X \leq t_0)$ from the result (1.2).

The MLE of $P(X \leqq t_0)$ is $P(\widehat{X \leq t_0}) = \sum_{t=0}^{t_0} \binom{k}{t} \hat{p}^t (1-\hat{p})^{k-t}$,

where $\hat{p} = \frac{1}{nk} \sum_{i=1}^{n} X_i$

Especially if $0 < t_0 < 1$, then $P(X_1 \leq t_0) = (1-p)^k$ and hence from the result (1.2)

$$\frac{\binom{(n-1)k}{S}}{\binom{nk}{S}} \text{ is a UMVUE of } P(X \leq t_0) = (1-p)^k \text{ , if } S \leqq (n-1)k \qquad (1.3)$$

$$\text{and the MLE of } P(X \leq t_0) = (1-p)^k \text{ is } P(\widehat{X \leq t_0}) = (1-\hat{p})^k \qquad (1.4)$$

And in addition, if k=1, that is, each X_i is Bernoulli random variable, it's no wonder that the UMVUE equals the MLE of $P(X \leqq t_0) = 1-p, 0 < t_0 < 1$, which is a well-known in Rohatgi(1976).

From the results (1.3) and (1.4), Table 2 shows the numerical values of MSE of the UMVUE and the MLE of $P(X \leq t_0)$ in the binomial distribution when k=5 and $0 < t_0 < 1$.

<Table 2> MSE of the UMVUE and the MLE of $P(X \leqq t_0)$ in the binomial distribution when k=5 and $0 < t_0 < 1$

n	/ p	1/4	1/2	3/4
10	MLE	0.01034	0.000743	0.0000048
	UMVUE	0.01327	0.001396	0.0000181
20	MLE	0.00494	0.000305	0.0000014
	UMVUE	0.00563	0.000438	0.0000033
30	MLE	0.00324	0.000189	0.0000008
	UMVUE	0.00413	0.000273	0.0000016

From Table 2, we can observe that the MLE of the reliability has less MSE than the UMVUE of the reliability in the binomial distribution when k=5 and $0 < t_0 < 1$.

1.3 Negative Binomial Model

Let X and Y be independent negative binomial random variables with parameters (r, p_1) and (r, p_2) ,respectively, The pmf of the negative binomial random variable with parameter r and p is defined as:

$$P(X=x)=\binom{x+r-1}{x}p^r(1-p)^x, \quad x=0,1,2,3,\ldots$$

Then, from the result of p.188 and Theorem 19 on p.216 in Rohatgi(1976) and the formulas 3.391 & 8.392 in Gradshteyn & Ryzhik(1965), the reliability in the negative binomial distribution can be shown as follows:

$$P(Y\le X)=\sum_{y=0}^{\infty}\binom{y+r-1}{y}p_1^r\cdot(1-p_1)^y\cdot I_{p_2}(r,y+1) \ ,$$

where $I_p(a,b)=\dfrac{B_p(a,b)}{B(a,b)}$ is an incomplete beta function.

Especially if r=1, that is, X and Y are geometric random variables each having parameters p_1 and p_2 ,respectively, then the reliability is:

$$P(Y\le X)=1-\frac{p_1-p_1p_2}{p_1+p_2-p_1p_2} \ .$$

Next we shall consider the UMVUE of the reliability in the negative binomial case. Assume $X_1.X_2,\cdot\cdot\cdot X_n$ be independent negative binomial random variables each having parameter (r_i,p) . i=1,2,...,n when the parameters $r_i's$ are known. Then it is well-known that $S=\sum_{i=1}^{n}X_i$ is a complete sufficient statistics for parameter 0<p<1,

By the same method of Poisson model in section (1.1), we can obtain the UMVUE of $P(X\leqq t_0)$ in the negative binomial distribution:

$$\sum_{t=0}^{t_0}\binom{r_1+t-1}{t}\cdot\binom{S-t+\sum_{i=2}^{n}r_i-1}{S-t}\Big/\binom{S+\sum_{i=1}^{n}r_i-1}{S} \text{ is a UMVUE of } P(X\leqq t_0) \quad (1.5)$$

where $S=\sum_{i=1}^{n}X_i$.

and also we can derive the UMVUE of the reliability $P(X>t_0)=1-P(X\le t_0)$ from the result (1.5).

The MLE of $P(X\le t_0)$ is $P(\widehat{X\le t_0}) = \sum_{t=0}^{t_0}\binom{t+r-1}{t}\hat{p}^r(1-\hat{p})^t$, where $\hat{p}=\dfrac{n}{n+\sum_{i=1}^{n}X_i}$.

Especially if $0<t_0<1 \text{ and } each\, r_i=1$, then $P(X\le t_0)=p$ and hence. from the result (1.5), we can observe the following special case (1.6) which is a well-known result in Rohatgi(1976):

$$\frac{n-1}{n-1+\sum_{i=1}^{n}X_i} \text{ is a UMVUE of } P(X\le t_0)=p \text{ ,} \tag{1.6}$$

and the MLE of $P(X\le t_0)=p$ is:

$$P(\widehat{X\le t_0})=\hat{p}=\frac{n}{n+\sum_{i=1}^{n}X_i} \text{ .} \tag{1.7}$$

From the results (1.6) and (1.7), Table 3 shows the numerical values of MSE of the UMVUE and the MLE of $P(X\le t_0)$ in the negative binomial distribution when each $r_i=1$ and $0<t_0<1$.

<Table 3> MSE of the UMVUE and the MLE of $P(X\le t_0)$ in the negative binomial distribution when each $r_i=1$ and $0<t_0<1$

n / p		1/4	1/2	3/4
10	MLE	0.00635	0.014149	0.013521
	UMVUE	0.00548	0.013805	0.014745
20	MLE	0.00265	0.006693	0.006917
	UMVUE	0.00242	0.006569	0.007205
30	MLE	0.00123	0.004387	0.004640
	UMVUE	0.00101	0.004307	0.004765

From Table 3, when each $r_i=1$ and $0<t_0<1$, we can observe that the UMVUE of the reliability has less MSE than the MLE of the reliability in the negative binomial distribution when p=1/4. 1/2 , while the MLE of the reliability has less MSE than the UMVUE of the reliability in the negative binomial distribution when p=3/4 .

Remark. As we have shown that the UMVUE's of $P(X\le t_0)$ for $0<t_0<1$ are well-known results in each discrete distribution. the results have been our purpose to show that the UMVUE should be generalized.

References

1. Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A., & Nesbitts, C. J.(1997), Actuarial Mathematics, The Society of Actuaries, Schaumburg, Illinois
2. Gradshteyn, I. S. & Ryzhik, I. M.(1965), Table of Integrals, Series, and Products, Academic Press, New York
3. Kim, J.(2006), The UMVUE of $[P(Y > X)]^k$ in a two Parameter Exponential Distribution, J. of the Korean Data & Information Science Society 17-2, pp.493-498
4. Lee. J. (2006), Reliability and Ratio of two Independent Exponential Distribution, J. of the Korean Data & Information Science Society 17-2, pp.515-520
5. Lee, C. & Won, H.(2006), Inference on Reliability in an Exponentiated Uniform Distribution, J. of the Korean Data & Information Science Society 17-2, pp.507-514
6. Rohatgi, V. K.(1976), Introduction to Probability Theory and Mathematical Statistics, John Wiley & Sons, New York

[received date : Oct. 2006, accepted date : Nov. 2006]