Journal of the Korean Data & Information Science Society
2011, **22**(4), 811–817

한국데이터정보과학회지

Estimating reliability in discrete distributions

Yeung-Gil Moon[1] · Chang-Soo Lee[2]

[1]Department of Tourism Quality Management, Kangwon Tourism College
[2]Department of Flight Operation, Kyungwoon University

Received 26 April 2011, revised 20 May 2011, accepted 11 June 2011

Abstract

We shall introduce a general probability mass function which includes several discrete probability mass functions. Especially, when the random variable X is Poisson, binomial, and negative binomial random variables as some special cases of the introduced distribution, the maximum likelihood estimator (MLE) and the uniformly minimum variance unbiased estimator (UMVUE) of the probability $P(X \leq t)$ are considered. And the efficiencies of the MLE and the UMVUE of the reliability are compared each other.

Keywords: Binomial, maximum likelihood estimator, negative binomial, Poisson, uniformly minimum variance unbiased estimator.

1. Introduction

Many authors have considered a right tail probability in continuous distributions for the reliability theory. Lee and Won (2006) considered inference on reliability in an exponentiated uniform distribution. Woo (2007, 2008) studied a reliability in two independent half normal distributions and Levy-uniform distributions. Moon and Lee (2009) considered an inference on the reliability in two independent gamma random variables. Lee and Lee (2010) considered reliability in two independent right truncated Rayleigh distributions. Ali and Woo (2010) studied estimation of tail probability and reliability in exponentiated Pareto case.

Since the curtate future lifetime random variable X has non-negative integer values, it is natural in actuarial studies to consider the right tail probability of discrete random variables to apply a reliability to evaluation of life insurance premiums (see, Bowers *et al.*, 1997).

Because reliability $R(t) = P(X > t) = 1 - P(X \leq t)$ is a monotone function of $P(X \leq t)$, an inference on the reliability is equivalent to an inference on $P(X \leq t)$ in McCool (1991), and hence instead of an inference on the reliability, it's sufficient for us to consider an inference on $P(X \leq t)$.

In this paper, We introduce a general probability mass function which includes several discrete probability mass functions. Especially, when the random variable X is Poisson, binomial, and negative binomial random variables as some special cases of the introduced distribution, the MLE and the UMVUE of the probability $P(X \leq t)$ are considered. And the efficiencies of the MLE and the UMVUE of the reliability are compared each other.

[1] Associate professor, Department of Tourism Quality Management, Kangwon Tourism College, Taebaek 235-711, Korea.
[2] Corresponding author: Associate professor, Department of Flight Operation, Kyungwoon University, Gumi 730-850, Korea. E-mail: cslee@ikw.ac.kr

2. Reliability of discrete random variables

We propose the probability mass function (pmf) of a discrete random variable X as follows:

$$P(X=j)=\alpha_j\cdot\theta^j/f(\theta),\quad j=0,1,2,\cdots,a,\ \theta>0 \text{ and } \alpha_j>0, \tag{2.1}$$

where a is a infinite or a positive integer, $f(\theta)\equiv\sum_{j=0}^{a}\alpha_j\cdot\theta^j$ of which $\sum_{j=0}^{\infty}\alpha_j\cdot\theta^j$ converges surely for $0<\theta<1$.

From the pmf (2.1) of a random variable X, we get the following special cases.

Example 2.1 (i) (binomial) If $a=n$ and $\alpha_j=\binom{n}{j}$ in (2.1), then $f(\theta)=(1+\theta)^n$ and hence X follows a binomial distribution with parameters n and $p=\theta/(1+\theta)$ for $\theta>0$.

(ii) (Poisson) If $a=\infty$ and $\alpha_j=1/j!$ in (2.1), then $f(\theta)=e^\theta$ and hence X follows a Poisson distribution with mean $\theta>0$.

(iii) (geometric) If $a=\infty$ and $\alpha_j=1$ in (2.1), then $f(\theta)=1/(1-\theta)$ for $0<\theta<1$ and hence X follows a geometric distribution with $p=1-\theta$.

(iv) (negative binomial) If $a=\infty$ and $\alpha_j=\binom{r+j-1}{j}$ in (2.1), then $f(\theta)=(1-\theta)^{-r}$ for $0<\theta<1$ and hence X follows a negative binomial distribution with a positive integer r and $p=1-\theta$.

From the pmf (2.1), the moment generating function (mgf) $M_X(t)=E(e^{tX})$ and the factorial moment generating function (fmgf) $p(t)=E(t^X)$ are obtained as following:

Proposition 2.1 (a) The mgf and the fmgf of a random variable X having the pmf (2.1) are

$$M_X(t)=\frac{f(\theta e^t)}{f(\theta)} \text{ and } p(t)=f(\theta\cdot t)/f(\theta).$$

(b) The kth factorial moment $\eta_X(k)=E[X(X-1)\cdots(X-k+1)]$ of a random variable X is

$$\eta_X(k)=\frac{\theta^k}{f(\theta)}\cdot\frac{d^kf(\theta)}{d\theta^k},\ k=1,2,3,\cdots, \tag{2.2}$$

where $\sum_{j=0}^{\infty}P(X=j)\cdot t^j$ converges surely for $|t|\le 1$.

Proof: (a) $M_X(t)=\sum_{j=0}^{a}e^{tj}\alpha_j\theta^j/f(\theta)=\sum_{j=0}^{a}\alpha_j(e^t\theta)^j/f(\theta)=f(\theta e^t)/f(\theta)$, where the last equality comes from the definition of $f(\theta)$. Hence, $p(t)=E(t^X)=M_X(\ln t)=f(\theta\cdot t)/f(\theta)$.

(b) The kth factorial moment of a random variable X is

$$\begin{aligned}\eta_X(t) = p^{(k)}(1) &= \left.\frac{d^k p(t)}{dt^k}\right|_{t=1} \\ &= \sum_{j=0}^{a} j(j-1)(j-k+1)\alpha_j\theta^j/f(\theta) = \frac{\theta^k}{f(\theta)}\sum_{j=0}^{a}\alpha_j \cdot j(j-1)(j-k+1)\theta^{j-k} \\ &= \frac{\theta^k}{f(\theta)}\frac{d^k}{d\theta^k}\left(\sum_{j=0}^{a}\alpha_j\theta^j\right) = \frac{\theta^k}{f(\theta)}\frac{d^k f(\theta)}{d\theta^k},\end{aligned}$$

which completes the proof. □

From the result (2.2), the mean and the variance of a random variable X having the pmf (2.1) are obtained as follows:

$$E(X) = \theta \cdot \frac{f'(\theta)}{f(\theta)} \text{ and } Var(X) = \theta^2 \cdot \frac{d}{d\theta}\left(\frac{f'(\theta)}{f(\theta)}\right) + E(X), \tag{2.3}$$

where $f'(\theta)$ denotes a derivative of $f(\theta)$.

By using the result (2.3) and $f(\theta)$ in (i)-(iv) of Example 2.1, we get the following means and variances of several discrete random variables in Example 2.2.

Example 2.2 (i) (binomial) Since $f(\theta) = (1+\theta)^n$, $E(X) = n \cdot \theta/(1+\theta)$ and $Var(X) = n\theta/(1+\theta)^2$.

(ii) (Poisson) Since $f(\theta) = e^\theta$, $E(X) = \theta$ and $Var(X) = \theta$ for $\theta > 0$.

(iii) (geometric) Since $f(\theta) = 1/(1-\theta)$, $E(X) = \theta/(1-\theta)$ and $Var(X) = \theta/(1-\theta)^2$ for $0 < \theta < 1$.

(iv) (negative binomial) Since $f(\theta) = (1-\theta)^{-r}$, $E(X) = r \cdot \theta/(1-\theta)$ and $Var(X) = r \cdot \theta/(1-\theta)^2$ for $0 < r < 1$ and r is a positive integer.

For a positive random variable X, we define the negative moment of order n by $E(X^{-n})$, where n is a positive integer, which its moment can be applied to evaluation of life insurance premiums in Bowers *et al.* (1997). And hence we obtain negative moment of order 1.

Proposition 2.2 If a random variable X has the pmf (2.1), then

$$E\left(\frac{1}{X+1}\right) = \frac{1}{\theta \cdot f(\theta)} \cdot \int f(\theta)d\theta.$$

Proof: $$\begin{aligned}E\left(\frac{1}{X+1}\right) &= \sum_{i=0}^{a}\frac{1}{j+1}\alpha_j\theta^j/f(\theta) = \frac{1}{\theta f(\theta)}\sum_{j=0}^{a}\alpha_j \cdot \frac{\theta^{j+1}}{j+1} \\ &= \frac{1}{\theta f(\theta)}\sum_{j=0}^{a}\alpha_j \cdot \int \theta^j d\theta = \frac{1}{\theta f(\theta)}\int\left(\sum_{j=0}^{a}\alpha_j\theta^j\right)d\theta \\ &= \frac{1}{\theta \cdot f(\theta)} \cdot \int f(\theta)d\theta.\end{aligned}$$

Therefore, we have done. □

From Proposition 2.2 and Example 2.2 we obtain the following Example 2.3.

Example 2.3 (i) (binomial) Since $f(\theta) = (1+\theta)^n$, $E(1/(1+X)) = (1+\theta)/((n+1)\theta)$.
(ii) (Poisson) Since $f(\theta) = e^\theta$, $E(1/(1+X)) = 1/\theta$.
(iii) (geometric) Since $f(\theta) = 1/(1-\theta)$, $E(1/(1+X)) = -(1-\theta)\ln(1-\theta)/\theta$ for $0 < \theta < 1$.
(iv) (negative binomial) Since $f(\theta) = (1-\theta)^{-r}$, $E(1/(1+X)) = (1-\theta)/(\theta(r-1))$ for $r > 1$.

Next, we consider estimations of the reliability of several discrete random variables as special cases of the pmf (2.1).

2.1. Poisson distribution

Let $X_1, X_2, \cdots, X_n$ be independent Poisson random variables with the mean parameter $\lambda > 0$. Then $S = \sum_{i=1}^n X_i$ is the complete sufficient statistics for $\lambda > 0$.

Define the following statistic $u(X)$ as the following: For given $t_0 > 0$,

$$u(X) = \begin{cases} 0 \text{ if } & X > t_0 \\ 1 \text{ if } & X \le t_0 \end{cases}.$$

Then $u(X)$ is an unbiased estimator of $P(X \le t_0)$ and hence from Lehmann-Scheffe Theorem, $E(U(X)|S)$ is the UMVUE of $P(X \le t_0)$. Since the conditional density of X_1 given $S = s$ can be obtained as

$$P(X_1 = t|S = s) = \binom{s}{t}\left(1 - \frac{1}{n}\right)^{s-t}\left(\frac{1}{n}\right)^t, \ t = 0, 1, 2, \cdots, \ s,$$

the UMVUE of $P(X \le t_0)$ is obtained as

$$\sum_{t-0}^{t_0} \binom{S}{t}(1 - 1/n)^{S-t}(1/n)^t. \tag{2.4}$$

Especially if $0 < t_0 < 1$, then $P(X \le t_0) = e^-\lambda$ and hence, from (2.4) it's a well-known that the UMVUE of $P(X \le t_0) = e^{-\lambda}$ is $(1 - \dfrac{1}{n})^{\sum_{i=1}^n X_i}$.

From the mgf of a Poisson random variable, the variance of the UMVUE is obtained by

$$e^{-2\lambda}[e^{\lambda/n} - 1]. \tag{2.5}$$

If $0 < t_0 < 1$, the MLE of $P(X \le t_0) = e^{-\lambda}$ is

$$\widehat{P}(X \le t_0) = \exp(-\widehat{\lambda}) = \exp\left(-\sum_{i=1}^n X_i/n\right).$$

Hence, the first and 2nd moments of the MLE $\widehat{P}(X \le t_0)$ are obtained by

$$E[\widehat{P}(X \le t_0)] = \exp(n\lambda(\exp(-1/n) - 1))$$

and

$$E[\hat{P}^2(X \le t_0)] = \exp(n\lambda(\exp(-2/n) - 1)) \tag{2.6}$$

Table 2.1 MSE of the UMVUE and the MLE of $P(X \leq t_0)$ in a Poisson case when $\lambda = 1$ and $0 < t_0 < 1$

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

From the results (2.5) and (2.6), Table 2.1 provides mean squared errors (MSE) of the UMVUE and the MLE of $P(X \leq t_0) = e^{-\lambda}$ for $\lambda = 1$ and $0 < t_0 < 1$.

From Table 2.1 and an equivalence between inferences of the probability $P(X \leq t)$ and the reliability in McCool (1991), in the Poisson model with $\lambda = 1$ when $0 < t_0 < 1$, the UMVUE of the reliability performs better than the MLE in the sense of MSE.

2.2. Binomial distribution

Let $X_1, X_2, \cdots, X_n$ be independent random variables each having distributed as the binomial with parameter (k, p), where k is assumed known positive integer. Then it's a well known that $S = \sum_{i=1}^{n} X_i$ is complete and sufficient statistics for p.

By the similar manner as like in Section 2.1, we obtain the UMVUE of $P(X \leq t_0)$ by

$$\sum_{i=1}^{t_0} \binom{k}{t} \cdot \binom{(n-1)k}{S-t} \Big/ \binom{nk}{S}, \; n > 1. \tag{2.7}$$

The MLE of $P(X \leq t_0)$ is given by

$$\hat{P}(X \leq t_0) = \sum_{t=0}^{t_0} \binom{k}{t} \cdot \hat{p}^t (1-\hat{p})^{k-t}, \; \hat{p} = \frac{1}{kn} \sum_{i=1}^{n} X_i. \tag{2.8}$$

Especially if $0 < t_0 < 1$, then $P(X \leq t_0) = (1-p)^k$, and hence from (2.7), the UMVUE of $P(X \leq t_0) = (1-p)^k$ is given as

$$\binom{(n-1)k}{S} \Big/ \binom{nk}{S} \text{ if } S \leq (n-1)k. \tag{2.9}$$

The MLE of $P(X \leq t_0) = (1-p)^k$ is given by

$$\widehat{P}(X \leq t_0) = (1-\widehat{p})^k. \tag{2.10}$$

From (2.9) and (2.10), if $k = 1$, that is, each X_i is a Bernoulli random variable, it's a well-known that the UMVUE and the MLE of $P(X \leq t_0)$ are the same when $0 < t_0 < 1$.

From the results (2.9) and (2.10), Table 2 provides the simulated MSE of the UMVUE and the MLE of $P(X \leq t_0)$ for k =5 and $0 < t_0 < 1$.

From Table 2.2 and an equivalence between inferences of probability $P(X \leq t)$ and the reliability in McCool (1991), in the binomial model with p =0.25, 0.5 and 0.75 for k =5 and $0 < t_0 < 1$, the MLE of the reliability performs better than the UMVUE in the sense of MSE.

Table 2.2 Simulated MSE of the UMVUE and the MLE of $P(X \leq t_0)$ in the binomial case when k=5 and $0 < t_0 < 1$

n		p		
		0.25	0.5	0.75
10	MLE	.01034	.000743	.0000048
	UMVUE	.01327	.001396	.0000181
20	MLE	.00494	.000305	.0000014
	UMVUE	.00563	.000438	.0000033
30	MLE	.00324	.000159	.0000008
	UMVUE	.00413	.000273	.0000016

2.3. Negative binomial distribution

The probability mass function of a negative binomial random variable with parameters r and p is given by

$$f(x) = \binom{x+r-1}{x} p^r (1-p)^x,\ x = 0, 1, 2, \cdots .$$

Let $X_1, X_2, \cdots, X_n$ be independent random variables each X_i having distributed as the negative binomial with parameter (r_i, p) for each i, where r_i is assumed known positive integer and $0 < p < 1$. Then it's a well known that $S = \sum_{i=1}^n X_i$ is a complete sufficient statistics for p.

By the similar manner as like in Section 2.1, the UMVUE of $P(X \leq t_0)$ in the negative binomial case is obtained by

$$\sum_{t=0}^{t_0} \binom{r_1+t-1}{t} \cdot \binom{S-t+\sum_{i=2}^n r_i - 1}{S-t} \Big/ \binom{S+\sum_{i=1}^n r_i - 1}{S}. \tag{2.11}$$

The MLE of $P(X \leq t_0)$ is given by

$$\widehat{P}(X \leq t_0) = \sum_{t=0}^{t_0} \binom{t+r-1}{t} \hat{p}^r (1-\hat{p})^t,$$

where $\widehat{p} = n/(n + \sum_{i=1}^n X_i)$.

Especially if $0 < t_0 < 1$ and each $r_i = 1$, then $P(X \leq t_0) = p$ and hence from (2.11), the UMVUE of $P(X \leq t_0) = p$ is

$$(n-1)/\left(n-1+\sum_{i=1}^n X_i\right) \tag{2.12}$$

and the MLE of $P(X \leq t_0) = p$ is given by

$$\hat{p}(X \leq t_0) = \hat{p} = n/\left(n+\sum_{i=1}^n X_i\right). \tag{2.13}$$

Table 2.3 Simulated MSE of the UMVUE and the MLE of $P(X \leq t_0)$ in the negative binomial case when $0<t_0<1$ and each r_i=1

n		p		
		0.25	0.5	0.75
10	MLE	.00635	.014149	.013521
	UMVUE	.00548	.013805	.014745
20	MLE	.00265	.006693	.006917
	UMVUE	.00242	.006569	.007205
30	MLE	.00123	.004387	.004640
	UMVUE	.00101	.004307	.004765

From the results (2.12) and (2.13), Table 2.3 provides simulated MSE of the UMVUE and the MLE of $P(X \leq t_0) = p$ in the negative binomial case when $0< t_0 <1$ and each r_i =1.

From Table 2.3 and an equivalence between inferences of probability $P(X \leq t)$ and the reliability in McCool (1991), in the negative binomial model with p =0.25 and 0.5 for $0 < t_0 < 1$, the UMVUE of the reliability performs better than the MLE in the sense of MSE. And vice versa when the negative binomial model has p =0.75 for $0 < t_0 < 1$.

3. Conclusion

In this paper, we introduce a general probability mass function which includes several discrete probability mass functions. Especially, when the random variable X is Poisson, binomial, and negative binomial random variables as some special cases of the introduced distribution, we obtain the MLE and the UMVUE of the probability $P(X \leq t)$.

Through the numerical or simulated MSE, we can observe as follows;

In the Poisson model with $\lambda = 1$ when $0 < t_0 < 1$, the UMVUE of the reliability performs better than the MLE in the sense of MSE. In the binomial model with p =0.25, 0.5 and 0.75 for k =5 and $0 < t_0 < 1$, the MLE of the reliability performs better than the UMVUE in the sense of MSE. And in the negative binomial model with p =0.25 and 0.5 for $0 < t_0 < 1$, the UMVUE of the reliability performs better than the MLE in the sense of MSE. And vice versa when the negative binomial model has p =0.75 for $0 < t_0 < 1$.

References

Ali, M. M. and Woo, J. (2010). Estimation of tail probability and reliability in exponentiated Pareto case. *Pakistan Journal of Statistics,* **26**, 39-47.

Bowers, N. J., Gerber, H. U., Hickman, J. C., Jones, D. A. and Nesbitts, C. J. (1997). *Actuarial mathematics*, The Society of Actuaries, Schaumburg, Illinois.

McCool, J. I. (1991). Inference on $P(X < Y)$ in the Weibull case. *Communications in Statistics-Simulation and Computation,* **20**, 129-148.

Moon, Y. G. and Lee, C. S. (2009). Inference on reliability $P(X < Y)$ in the gamma case. *Journal of the Korean Data & Information Science Society*, **20**, 219-223.

Lee, C. C. and Lee, C. S. (2010). Reliability and ratio in a right truncated Rayleigh distribution. *Journal of the Korean Data & Information Science Society*, **21**, 195-200.

Lee, C. S. and Won, H. Y. (2006). Inference on reliability in an exponentiated uniform distribution. *Journal of the Korean Data & Information Sciences Society*, **17**, 507-513.

Woo, J. (2007). On reliability and UMVUE of right-tail probability in a half-normal variable. *Journal of the Korean Data & Information Science Society*, **18**, 259-267.

Woo, J. (2008). Reliability in two independent Levy and uniform-exponential distribution. *Journal of the Korean Data & Information Science Society,* **19**, 635-644.