# Log Pearson 3형 분포매개변수의 분산 · 공분산 비교

전시영*, 김민구**, 이종국**

## 1. 서 론

Rao(1980, 1983)는 표본왜곡도의 직접이용을 제거하기 위하여 혼합모멘트법 (Method of Moment, MXM)이라 불리우는 새로운 방법을 제안하여 LP3(Log Pearson type 3)분포의 매개변수를 2가지 MXM 방법과 다른 2가지 방법으로 추정하였다. Hoshi와 Burges(1981)는 실측자료를 LP3 분포에 적용하여 매개변수를 추정하여 이 들의 표본 분산값들로부터 홍수사상의 크기와 분산을 추정하는데 이용하였다. Phien과 Hira(1983)는 LP3 분포의 매개변수추정을 위한 4가지 MXM 방법과 다른 2가지 방법을 소개하였다. Phien과 Hsu(1985)는 LP3 분포에 의하여 추정된 T년 사상의 분산 계산과정을 5가지인 직접모멘트방법과 4가지 MXM 방법으로 기술하고 이 결과로부터 형상계수 b를 추정하는데 가장 큰 오차가 발생한다고 보고한 바 있다. Arora와 Singh(1989)는 LP3 분포의 매개변수 추정과정을 2가지 MXM 방법 및 다른 방법으로 분석하였다. 국내에서도 양 등(1993)이 한국 하천에서의 LP3 분포의 적용성에 대한 연구를 시작하였고 전(1995)은 4가지 MXM에 의한 확률분포형의 T-년 quantiles를 비교하였으며 김(1999)은 MXM과 L-Moments법에 의한 확률분포형의 T-년 quantiles를 비교·분석한 바 있다.

본 연구는 실측된 자료를 LP3 확률분포형에 적합하여 이 분포의 특성을 나타내는 매개변수를 직접모멘트법과 혼합모멘트법으로 추정하여 그 결과로부터 분포매개 변수 $a$, $b$ 및 $c$의 분산과 공분산의 추정치를 비교 · 분석하고져 한다.

## 2. LP3 확률분포형 이론

LP3 분포의 변량 $x$의 확률밀도함수(probability density function, PDF)는 다음과 같다.

$$f(x) = \frac{1}{|a|\Gamma(b)x}\left[\frac{\ln x - c}{a}\right]^{b-1}\exp\left[-\frac{\ln x - c}{a}\right] \tag{1}$$

여기서 $a$, $b$ 및 $c$는 각각 축척(scale), 형상(shape) 및 위치(location)변수이고 $\Gamma$는 gamma함수이다.

LP3 분포의 원점에 관한 $x$의 $r$차 모멘트, $\mu_r'$는 다음 식으로 주어진다.

$$\mu'_r = M'_r = \frac{\exp(rc)}{(1-ra)^b}\ , \quad 1 - ra > 0\ , \quad r = 1, 2, 3 \tag{2}$$

* 원광대학교 토목환경공학과 교 수
** 원광대학교 토목환경공학과 석사과정
** 원광대학교 산업대학원 석사과정

평균에 관한 $r$차 모멘트 $\mu_r$는 다음과 같다.

$$\mu_r = \sum_{j=0}^{r} \binom{r}{j} \mu'_{r-j} (-\mu_1')^j \tag{3}$$

P3(Pearson type 3) 분포의 평균에 관한 $Y = \ln X$의 모멘트는 식(1)을 이용하여 약간의 수학적인 변환을 하면 다음과 같이 나타낼 수 있다.

$$\mu_1(Y) = E(Y) = ab + c \tag{4}$$

$$\mu_2(Y) = a^2 b \tag{5}$$

$$\mu_3(Y) = 2a^3 b \tag{6}$$

매개변수 추정을 위한 표본크기 $N$인 $X$와 $Y$의 모멘트는 다음 식에 의하여 추정된다.

$$M_r' = \frac{1}{N} \sum x^r, \quad r = 1, 2, 3 \tag{7}$$

$$m_1' = \overline{Y} = \frac{1}{N} \sum y \tag{8}$$

$$m_2 = \frac{1}{N-1} \sum (y - \overline{Y})^2 \tag{9}$$

3. 분포매개변수 추정치의 분산 및 공분산

본 연구에서 매개변수 추정에 대한 5가지 방법은 $X$의 모멘트 또는 $X$와 $Y$의 조합에 의한 모멘트에 근거로 하여 이 들의 모멘트나 매개변수 추정치의 분산(variance) 및 공분산(covariance)과의 관계를 확립해야 한다. 우선 두 개의 함수 $M_p'(a_1, a_2, a_3)$와 $M_q'(a_1, a_2, a_3)$를 고려할 때 이 들 함수의 공분산은 다음과 같이 나타낼 수 있다(Kendall과 Stuart, 1977).

$$\begin{aligned} Cov(M_p', M_q') = & \sum_{i=1}^{3} \left( \frac{\partial M_p'}{\partial a_i} \right) \left( \frac{\partial M_q'}{\partial a_i} \right) Var\, a_i \\ & + \sum_{i=1}^{3} \left( \sum_{\substack{j=1 \\ i \neq j}}^{3} \right) \left( \frac{\partial M_p'}{\partial a_i} \right) \left( \frac{\partial M_q'}{\partial a_i} \right) Cov(a_i, a_j) \end{aligned} \tag{10}$$

여기서 $p, q = 1, 2, 3$ 그리고 $a_1 = a$, $a_2 = b$ 및 $a_3 = c$이고 Var( · )와 Cov( · )는 각각 분산과 공분산을 나타낸다. 식(10)에서 $p = q$일 때 $Cov(M_p', M_p') = Var\, M_p'$의 관계를 얻을 수 있다. 식(10)은 다음과 같이 행렬식의 형태로 쓸 수 있다.

$$\begin{bmatrix} Var\, M_1' \\ Var\, M_2' \\ Var\, M_3' \\ Cov(M_1', M_2') \\ Cov(M_2', M_3') \\ Cov(M_1', M_3') \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \\ d_{41} & d_{42} & d_{43} & d_{44} & d_{45} & d_{46} \\ d_{51} & d_{52} & d_{53} & d_{54} & d_{55} & d_{56} \\ d_{61} & d_{62} & d_{63} & d_{64} & d_{65} & d_{66} \end{bmatrix} \begin{bmatrix} Var\ a \\ Var\ b \\ Var\ c \\ Cov(a, b) \\ Cov(b, c) \\ Cov(a, c) \end{bmatrix}$$

또는

$$V_M = D \cdot V_p \tag{11}$$

여기서 $V_M$와 $V_p$는 각각 주어진 자료 $X$의 모멘트와 분포매개변수들의 추정치의 분산 및 공분산 벡터이고 $D$는 요소 $d_{ij}$를 갖는 행렬식이며 이는 매개변수들에 관한 모멘트의 편도함수이다..

식(11)에서 벡터 $V_M$의 요소는 임의의 $p$와 $q$, 임의의 표본자료수 $N$에 대하여 Kendall과 Stuart(1977)의 다음과 같은 식으로부터 쉽게 결정할 수 있다.

$$Var\,\mu_r' = \frac{1}{N}(\mu_{2r}' - \mu_r'^2) \tag{12}$$

$$Cov(\mu_q', \mu_r') = \frac{1}{N}(\mu_{q+r}' - \mu_q'\mu_r') \tag{13}$$

식(11)에서 행렬식 $D$의 요소와 $V_M$의 요소는 쉽게 추정될 수 있으며, 벡터 $V_p$는 다음과 같은 관계로부터 결정될 수 있다.

$$V_p = D^{-1}V_M \tag{14}$$

여기서 $D^{-1}$은 $D$의 역행렬식이고 대부분의 경우에 존재하는 것으로 알려져있다. 다음은 $d_{ij}$와 $V_M$의 추정과 이 들로부터 분포매개변수의 분산에 관한 5가지 방법을 기술하였다.

### 3.1 Bobee 방법 (직접모멘트법)

이 방법은 모멘트 $M_1'$, $M_2'$ 및 $M_3'$를 이용하여 이 들의 분산 및 공분산을 추정하는 것이다. 이 방법에서 $Var M_1'$는 식(12)에 $r=1$를 대입하고 식(2)를 이용하여 정리하면 다음과 같다.

$$Var M_1' = \frac{1}{N}e^{2c}[(1-2a)^{-b}-(1-a)^{2b}] \tag{15}$$

같은 방법으로 $Var M_2'$와 $Var M_3'$는 다음과 같이 얻을 수 있다.

$$Var M_2' = \frac{1}{N}e^{4c}[(1-4a)^{-b}-(1-2a)^{-2b}] \tag{16}$$

$$Var M_3' = \frac{1}{N}e^{6c}[(1-6a)^{-b}-(1-3a)^{-2b}] \tag{17}$$

$Cov(M_1', M_2')$는 식(13)에 $q=1$, $r=2$를 대입하고 식(2)를 이용하여 정리하면 다음과 같다.

$$Cov(M_1', M_2') = \frac{1}{N}e^{3c}[(1-3a)^{-b}-(1-a)^{-b}(1-2a)^{-b}] \tag{18}$$

$Cov(M_2', M_3')$와 $Cov(M_1', M_3')$도 $Cov(M_1', M_2')$를 구하는 방법과 같이 식(13)에 $q=2$, $r=3$ 및 $q=1$, $r=3$을 대입하고 식(2)를 이용하여 정리하면 얻을 수 있다.

$$Cov(M_2', M_3') = \frac{1}{N}e^{5c}[(1-5a)^{-b}-(1-2a)^{-b}(1-3a)^{-b}] \tag{19}$$

$$Cov(M_1', M_3') = \frac{1}{N}e^{4c}[(1-4a)^{-b}-(1-a)^{-b}(1-3a)^{-b}] \tag{20}$$

식(11)의 $Var M_1'$는 행렬식 $D$의 첫 번째 행과 $V_p$를 곱하므로서 다음과 같이 표현된다.

$$\begin{aligned} Var M_1' = \; & d_{11}Var\,a + d_{12}Var\,b + d_{13}Var\,c + \\ & d_{14}Cov(a, b) + d_{15}Cov(b, c) + d_{16}Cov(a, c) \end{aligned} \tag{21}$$

식(21)에서 분산과 공분산들의 계수는 식(10)에 $p=q=1$을 대입하여 정리한 식과 같게 놓으므로서 다음과 같이 구할 수 있다.

$$d_{11} = \left(\frac{\partial M_1'}{\partial a}\right)^2, \qquad d_{12} = \left(\frac{\partial M_1'}{\partial b}\right)^2, \qquad d_{13} = \left(\frac{\partial M_1'}{\partial c}\right)^2$$
$$d_{14} = 2\left(\frac{\partial M_1'}{\partial a}\right)\left(\frac{\partial M_1'}{\partial b}\right), \qquad d_{15} = 2\left(\frac{\partial M_1'}{\partial b}\right)\left(\frac{\partial M_1'}{\partial c}\right), \tag{22}$$
$$d_{16} = 2\left(\frac{\partial M_1'}{\partial a}\right)\left(\frac{\partial M_1'}{\partial c}\right)$$

$Var\, M_2'$와 $Var\, M_3'$도 $Var\, M_1'$구하는 방법과 같은 방법으로 $M_1'$대신 각각 $M_2'$, $M_3'$를 $d_{ij}$의 요소에 대입하면 얻을 수 있다. $Cov(M_1', M_2')$는 $Var\, M_1'$구하는 방법과 같이 식(11)과 (10)에 $p=1$, $q=2$를 대입하여 정리하면 구 할 수 있다. 같은 방법으로 $Cov(M_2', M_3')$와 $Cov(M_1', M_3')$도 얻을 수 있다.

전술한 바와같이 $D$의 요소 $d_{ij}$는 다음과 같이 쉽게 얻을 수 있다. 식(3)에 $r=1$을 대입한 $M_1'$을 각각 $a$, $b$ 및 $c$에 대하여 편미분하면 다음과 같다.

$$d_{11} = A_{11} = \frac{\partial M_1'}{\partial a} = \frac{M_1' b}{(1-a)} \tag{23}$$

$$d_{12} = A_{12} = \frac{\partial M_1'}{\partial b} = -M_1' \ln(1-a) \tag{24}$$

$$d_{13} = A_{13} = \frac{\partial M_1'}{\partial c} = M_1' \tag{25}$$

$M_2'$와 $M_3'$도 $M_1'$와 같은 방법으로 식(2)를 이용하면 구할 수 있다. 다른 $d_{ij}$요소들도 위의 결과를 조합하면 구할 수 있다. $D$와 $V_M$이 결정되면 $V_p$는 식(14)를 이용하면 쉽게 추정된다.

### 3.2 MXM1 방법

이 방법은 전술한 방법을 이용하여 $M_3'$대신에 $m_1'$로 대치하면 분포매개변수의 분산과 공분산 $V_p$를 구할 수 있다.

### 3.3 MXM2 방법

$V_p$는 MXM1 방법의 $M_3'$ 대신에 $m_2$로 대입하여 정리하면 얻을 수 있다.

### 3.4 MXM3 방법

이 방법에서도 $V_p$는 앞의 MXM2 방법의 $M_2'$대신에 $m_1'$로 대치하면 결정할 수 있다.

### 3.5 MXM4 방법

이 방법에서 $V_p$는 MXM3 방법의 $M_1'$ 대신 $M_2'$로 대치하면 산정된다.

## 4. 적용 예

본 연구에서 이용한 자료는 전북 남원측후소의 1973-1996년 동안의 1, 2, 4, 6 및 12시간 지속

강우량이다.

4.1 분포매개변수 추정치의 분산·공분산

이미 3장에서 기술한 5가지 추정방법에 의한 분포매개변수 $a$, $b$ 및 $c$의 분산 · 공분산의 추정치는 표 1에 나타내었다.

표 1. 분포매개변수의 분산 · 공분산 추정치

| 강우지속기간 (hr) | 방 법 | $Var\ a$ | $Var\ b$ | $Var\ c$ | $Cov(a,b)$ | $Cov(b,c)$ | $Cov(a,c)$ |
|---|---|---|---|---|---|---|---|
| 1 | Bobee | .11910E+00 | .59663E+12 | .30713E+07 | -.26657E+06 | -.13537E+10 | .60480E+03 |
| | MXM1 | .75914E-02 | .17687E+04 | .24539E+01 | .36336E+01 | .65272E+02 | .65272E+02 |
| | MXM2 | .18611E-01 | .21683E+03 | .69797E+00 | .19900E+01 | .12164E+02 | .10963E+00 |
| | MXM3 | .43478E-01 | .67974E+08 | .58526E+04 | -.17190E+04 | -.63068E+06 | .15946E+02 |
| | MXM4 | .25037E-01 | .11131E+04 | .25868E+01 | .52550E+01 | .53391E+02 | .24987E+00 |
| 2 | Bobee | .14005E-01 | .64442E+10 | .67910E+05 | -.94997E+04 | -.20919E+08 | .30835E+02 |
| | MXM1 | .82956E-02 | .23029E+03 | .55573E+00 | .13618E+01 | .11115E+02 | .63789E-01 |
| | MXM2 | .22471E-01 | .74259E+02 | .32097E+00 | .12764E+01 | .48141E+01 | .81012E-01 |
| | MXM3 | .42992E-01 | .86121E+08 | .65506E+04 | -.19240E+04 | -.75104E+06 | .16776E+02 |
| | MXM4 | .30690E-01 | .27303E+03 | .97657E+00 | .28761E+01 | .16217E+02 | .16888E+00 |
| 4 | Bobee | .10412E-01 | .59771E+07 | .60657E+03 | .24936E+03 | .60186E+05 | .25087E+01 |
| | MXM1 | .87165E-02 | .12363E+08 | .90023E+03 | -.32816E+03 | -.10546E+06 | .27973E+01 |
| | MXM2 | .16972E-01 | .25684E+03 | .70395E+00 | .20699E+01 | .13309E+02 | .10548E+00 |
| | MXM3 | .42428E-01 | .21274E+09 | .11313E+05 | -.30042E+04 | -.15513E+07 | .21903E+02 |
| | MXM4 | .20081E-01 | .64285E+04 | .75191E+01 | .11332E+02 | .21922E+03 | .38438E+00 |
| 6 | Bobee | .15085E-01 | .62413E+08 | .30882E+04 | -.97016E+03 | -.43897E+06 | .68214E+01 |
| | MXM1 | .80718E-02 | .67728E+07 | .55775E+03 | -.23371E+03 | -.61435E+05 | .21181E+01 |
| | MXM2 | .13377E-01 | .98278E+03 | .15854E+01 | .36044E+01 | .39204E+02 | .14209E+00 |
| | MXM3 | .41684E-01 | .77893E+07 | .11447E+04 | -.56972E+03 | -.94412E+05 | .69034E+01 |
| | MXM4 | .16679E-01 | .11153E+06 | .46002E+02 | .43083E+02 | .22624E+04 | .87206E+00 |
| 12 | Bobee | .10858E-01 | .85077E+07 | .78771E+03 | .30382E+03 | .81832E+05 | .29201E+01 |
| | MXM1 | .97216E-02 | .12114E+07 | .19781E+03 | -.10844E+03 | -.15469E+05 | .13828E+01 |
| | MXM2 | .11096E-01 | .43722E+05 | .22006E+02 | .21980E+02 | .97869E+03 | .48993E+00 |
| | MXM3 | .42094E-01 | .61901E+07 | .10278E+04 | -.51037E+03 | -.79748E+05 | .65728E+01 |
| | MXM4 | .16015E-01 | .11713E+08 | .10939E+04 | .43299E+03 | .11316E+06 | .41812E+01 |

표 1에 의하면 각 지속시간별에 따라 분산과 공분산들의 차이가 매우 크게 나타남을 알 수 있다. 이 표에서 $Var\ a$는 MXM1 방법이 모든 지속시간에서 가장 작은 값으로 산정되었으며 이를 제외한 지속기간에서는 MXM2 방법이 분산과 공분산 추정치에서 작은 값으로 나타났다. 그 다음으로 분산과 공분산 작게 추정된 것이 MXM1 방법이다. Bobee나 MXM3 방법은 전반적으로 크게 산정되었다. 각각의 분산과 공분산별로 비교하면 $Var\ a$는 5가지 방법과 모든 지속기간중에서 가장 작게 추정되었고 이에 반하여 $Var\ b$ 가장 크게 추정되었다. 이 표에서 분포매개변수의 분산

들만 비교하여 보면 전 지점에서 $Var\ b$가 $Var\ a$나 $Var\ c$에 비하여 매우 크게, $Var\ a$는 가장 작게 추정되었다. 또한 공분산도 모든 지점에서 $Cov(b,c)$이 $Cov(a,b)$나 $Cov(a,c)$에 비하여 매우 크게, $Cov(a,c)$가 가장 작게 추정되었다. 이 결과로부터 LP3 분포의 매개변수 $a$, $b$ 및 $c$를 추정하는데 발생된 최대의 분산은 형상계수 $b$를 결정하는데 발생함을 알 수 있었다.

## 5. 결 론

본 연구에서는 직접모멘트와 혼합모멘트법으로 LP3 분포에 대한 분포매개변수 추정치의 분산과 공분산을 해석적인 방법으로 추정하기 위한 자세한 과정이 기술되었다. 본 연구에 적용한 예는 전북 남원측후소의 강우량자료로부터 분산과 공분산이 추정되었으며 얻어진 결과는 다음과 같다.

분포매개변수 $a$, $b$ 및 $c$의 분산과 공분산은 $x$의 처음 두 모멘트와 $y$의 분산을 이용한 MXM2 방법이 MXM1 방법의 추정된 $Var\ a$ 값을 제외한 모든 지속기간에서 작게 추정되었다. 추정된 분산 $Var\ b$ 값은 모든 지속기간에 대한 5가지 방법에서 가장 작게 $Var\ c$ 값은 가장 크게, 공분산 $Cov(b,c)$ 값도 $Cov(a,b)$와 $Cov(a,c)$ 값에 비하여 매우 크게 추정되었다. 따라서 LP3 분포의 분포매개변수를 추정하는데 형상계수 $b$는 분산 $Var\ b$와 공분산 $Cov(b,c)$추정에 상당히 큰 영향을 주는 것을 알 수 있었다.

## 6. 참고문헌


1. 김문기, 혼합모멘트법과 L-Moments법에 의한 확률분포형의 T-년 quantiles를 비교·분석. 원광대학교 산업대학원, 1999. 2.
2. 양정석, 선우중호, 1993. 한국하천에서의 Log-Pearson Type 3 분포의 적용성에 대한 연구. 1993년도 대한토목학회 학술발표회 개요집(Ⅱ), pp. 177-180.
3. 전시영, 1995. 혼합모멘트방법에 의한 제3형 Log Pearson 분포의 T년 Quantiles의 비교. 한국수자원학회 학술발표회 논문집, 한국수자원학회., pp. 309-316.
4. Arora, K. and Singh, V.P., 1989. A Comparative of the Estimators of the Log Pearson Type(LP) 3 Distribution. J. Hydrol., 105, pp. 19-37.
5. Hoshi, K. and Burges, S.J., 1981. Sampling Properties of Parameter Estimayes for the Log Pearson Type 3 Distribution, Using Moments in Real Space. J. Hydrol., 53, pp. 305-316.
6. Kendall, M.G. and Stuart, A., 1977. The Advanced Theory of Statistics, Vol. 1. Charles Griffin, London, 4th ed., pp. 243-262.
7. Rao, D.V., 1980. Log Pearson Type 3 Distribution : Method of Mixed Moments. J. Hydr. Div., ASCE, 106(6), pp. 999-1019.
8. Rao, D.V., 1983. Estimating Log Pearson Parameters by Mixed Moments. J. Hydr. Div., ASCE, 109(8), pp. 1118-1131.
9. Phien, H.N. and Hira, M.A., 1983. Log Pearson Typr-3 Distribution : Parameter Estimatio. J. Hydrol., 64, pp. 25-37.
10. Phien, H.N. and Hsu, L.C., 1985. Variance of the T-Year Event in the Log Pearson Type-3 Distribution. J. Hydrol., 77, pp. 141-158.