

GEV와 GLO 분포의 유출량 교차상관과 L-moment 추정값의 교차상관의 관계 유도

Derivation of relationship between cross-site correlation among flows and among estimators of L-moments for GEV and GLO distribution

정대일* / 제리 스테던저**
Dae Il Jeong, Jerry R. Stedinger

요 지

3개의 매개변수(location, scale, shape)로 이루어진 GEV와 GLO 분포는, 미국의 공식적인 홍수빈도 분포인 Log Pearson Type III와 함께 수문분야에서 중요한 위치를 차지하고 있다. 본 연구에서는 Monte Carlo 실험을 이용하여 GEV와 GLO 분포에서 서로 다른 두 지점의 유출량 자료를 생성하여 L-CV(L-moment Coefficient of Variation; τ_2)와 L-CS(L-moment Coefficient of Skewness; τ_3)를 추정하였으며, 이 L-moment 추정값들 간의 교차상관($\tau_2-\tau_2$, $\tau_3-\tau_3$, $\tau_2-\tau_3$)과 유출량 자료간의 교차상관의 관계를 Simple Power 함수를 이용하여 유도하였다. 실험 과정에서 GEV와 GLO 분포가 비현실적인 음수 유출량을 생성하여, 실험 결과에 큰 영향이 있음을 확인하여, 두 분포에서 생성된 유출량 자료에서 음수값을 제외한 GEV+와 GLO+ 분포를 이용하여 관계식을 유도하고 이를 GEV와 GLO 분포의 결과와도 비교하였다. 본 연구에서 도출된 관계식은 향후 Generalized Least Square 회귀식을 이용하여 홍수분포의 지역 매개변수를 추정하기 위해 활용성이 클 것으로 기대한다.

핵심용어 : Generalized Extreme Value, Generalized Logistic, L-CS, L-CV, 교차상관, 홍수빈도분석

1. 서론

Jenkinson(1955)에 의해 1950년대에 제안된 Generalized Extreme Value(GEV) 분포는 수문학 분야에서 홍수빈도, 확률강우 등의 추정에 가장 일반적으로 사용되는 분포 중 하나이다. 최근 영국에서는 L-moment법을 이용한 매개변수 추정과 함께 Generalized LOGistic(GLO) 분포를 이용한 홍수빈도분석을 공인 하였다(Robson and Reed, 1999). 따라서 위치(location) 규모(scale) 형상(shape) 매개변수로 이루어진 GEV와 GLO 분포는, 미국의 공식적인 홍수빈도 분포인 Log Pearson Type III 분포와 함께 수문분야에서 중요한 위치를 차지하고 있다.

본 연구에서는 먼저 수학적 표현이 서로 유사한 GEV와 GLO 분포에 대해 이론적으로 비교하였다. 다음으로, Monte Carlo 실험을 통해 GEV와 GLO 분포에서 서로 다른 두 지점(x , y)의 유출량 자료를 생성하여 L-CV(L-moment Coefficient of Variation; τ_2)와 L-CS(L-moment Coefficient of Skewness; τ_3)를 추정하였으며, 이 L-moment 추정값들 간의 교차상관($\tau_{2x}-\tau_{2y}$, $\tau_{3x}-\tau_{3y}$, $\tau_{2x}-\tau_{3y}$)과 유출량 자료간의 교차상관의 관계를 Simple Power 함수를 이용하여 유도하였다. 이 과정에서 GEV와 GLO 분포에서 생성된 유출량 자료에, 비현실적인 음수의 유출량이 생성되며, 그 양이 크지 않을지라도 실험 결과에 크게 영향을 미치고 있음을 확인하였다. 따라서 두 분포에서 생성된 유출량 자료에서 음수값을 제외한 GEV+와 GLO+ 분포를 이용한 관계식을 제시하였다.

* Post Doc, School of Civil and Environmental Engineering, Cornell University · E-mail: dj64@cornell.edu

** Professor, School of Civil and Environmental Engineering, Cornell University · E-mail: jrs5@cornell.edu

본 연구를 통해 도출된 두 지점 유출량의 L-moment 추정값들 간의 교차상관과 유출량 자료간의 교차상관 관계는 Generalized Least Square(GLS) 회귀식을 이용한 홍수분포의 지역 매개변수 추정에 활용가치가 크다. Stedinger and Tasker(1985)에 의해 제안된 GLS 회귀식은 홍수빈도 분석의 분포로 사용되는 매개변수를 지점(at-site) 추정값의 정확성과 유역의 물리적 특성을 반영하여 미계측 유역이나 자료의 길이가 짧은 유역에 지역화 할 수 있다.

2. GEV와 GLO 분포의 비교

GEV 분포의 누가분포함수(cumulative distribution function)는 식(1)과 같다(Stedinger et al., 1993).

$$F(x) = \exp\left[-\left(1 - \frac{\kappa}{\alpha}(x - \xi)\right)^{1/k}\right] \quad \kappa \neq 0 \quad (1a)$$

$$F(x) = \exp\left[-\exp\left(-\frac{(x - \xi)}{\alpha}\right)\right] \quad \kappa = 0 \quad (1b)$$

여기서, ξ , α , κ 는 각각 위치(location), 규모(scale), 형상(shape)에 관한 매개변수이며, $\kappa = 0$ ($\tau_3 = 0.17$)인 경우 잘 알려진 Gumbel 분포가 된다. 식(1)에서 x 의 범위는 κ 에 따라, $\kappa < 0$ 일 때는 $\xi + \alpha/\kappa \leq x < \infty$, $\kappa = 0$ 일 때는 $-\infty < x < \infty$, $\kappa > 0$ 일 때는 $-\infty < x \leq \xi + \alpha/\kappa$ 가 된다.

GLO 분포의 누가분포함수는 식(2)와 같다(Hosking and Wallis, 1997).

$$F(x) = [1 + (1 - \frac{\kappa}{\alpha}(x - \xi))^{1/k}]^{-1} \quad \kappa \neq 0 \quad (2a)$$

$$F(x) = [1 + \exp(-\frac{(x - \xi)}{\alpha})]^{-1} \quad \kappa = 0 \quad (2b)$$

GEV분포에서와 마찬가지로, ξ , α , κ 는 각각 위치, 규모, 형상 매개변수이며, 형상계수 $\kappa = 0$ ($\tau_3 = 0.0$)인 경우 Logistic 분포가 된다. 식(2)에서 x 의 범위 역시 κ 에 따라 GEV 분포와 같은 범위를 갖는다.

식(1)과 (2)에서 살펴본 바와 같이, GEV와 GLO 분포는 수학적으로 매우 유사한 형태를 가진다. 식(1a)에서 x 가 커지면 $[1 - \kappa(x - \xi)/\alpha]^{1/k}$ 의 값이 작아져 결국 식(2a)로 정리되며, 마찬가지로 식(1b)에서는 x 가 커지면 $\exp(-(x - \xi)/\alpha)$ 이 작아져 식(2b)로 정리된다. 따라서 x 가 커질수록 같은 값의 매개변수 가지는 GEV와 GLO 분포의 우측 꼬리 부분이 일치됨을 짐작할 수 있다.

그림1은 GEV와 GLO 분포의 κ 와 L-CS의 관계를 그래프로 그린 것이다. 같은 L-CS에서 GLO 분포의 κ 가 GEV 분포에 비해 작아 GLO 분포의 우측 꼬리부분이 더 두터움을 짐작할 수 있다. 그림 2는 같은 값의 평균($\mu = 10$), L-CV($\tau_2 = 0.3$), L-CS($\tau_3 = 0.5, 0.3, 0.1, -0.1$)를 갖는 GEV와 GLO 분포를 비교한 것이다. 그림에서 x-축은 누가확률(P)을 Gumbel 스케일로 변환한 것으로 분포의 좌우 극한을 자세히 관찰하기 위한 것이다. L-CS가 큰 경우(예; $\tau_3 = 0.5$), 임의의 누가확률 P 에 대해, 두 분포의 좌우 극한 값에 큰 차이가 없었으나, 그렇지 않은 경우(예; $\tau_3 = 0.3, 0.1, -0.1$)에는 GLO 분포의 우 극한 값이 더 컸으며, 좌 극한 값은 더 작음을 확인할 수 있었다.

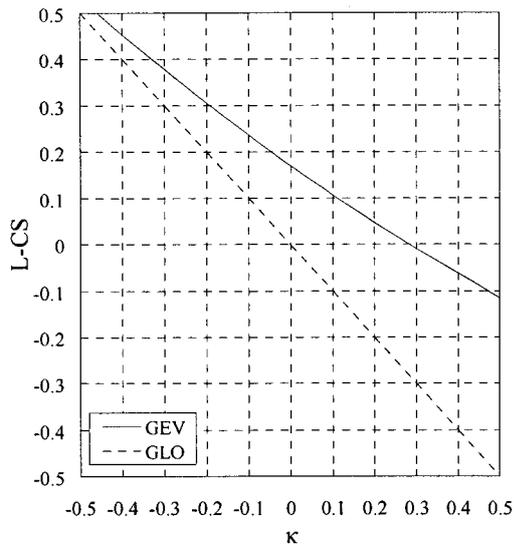


그림 1 GEV와 GLO 분포의 κ 와 τ_3 관계

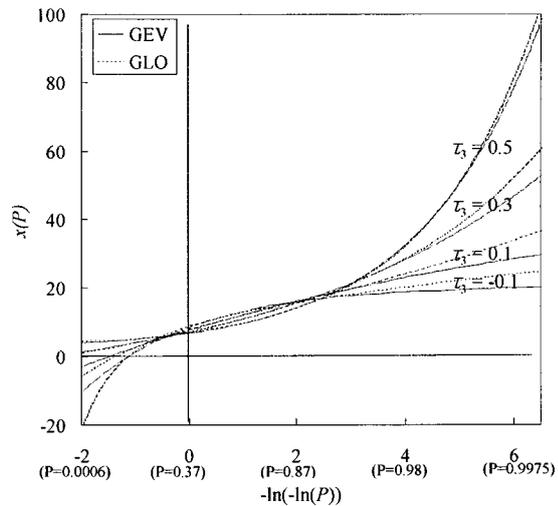


그림 2 동일한 μ, τ_2, τ_3 에 대한 GEV와 GLO 분포의 누가분포함수 비교

3. Monte Carlo 실험 및 결과

3.1 Monte Carlo 실험

Martins and Stedinger(2002)는 Monte Carlo 실험을 이용하여 Log Person Type III 분포의 외도(skewness)계수와 GEV와 Generalized Pareto 분포의 형상계수에 대한, 두 지점간의 상관관계를 유도하였다. 본 연구의 Monte Carlo 실험은 기존의 Martins and Stedinger(2002)가 사용한 방법과 같은 맥락에서 이루어졌다. 연최대 홍수량을 대변하는 두 지점(x, y)의 유출량 자료를 표 1의 길이를 갖도록 생성하였다. 표 1에서 n_{xy} 는 두 지점의 유출량 자료가 겹치는 길이를 의미하며, n_x 와 n_y 는 서로 겹치지 않는 자료의 길이를 의미한다.

GEV분포의 L-CV와 κ 의 범위는 $0.1 \leq \tau_2 \leq 0.5$ 와 $-0.5 \leq \kappa \leq 0.4$ 로, GLO 분포는 $0.1 \leq \tau_2 \leq 0.5$ 와 $-0.1 \leq \kappa \leq 0.4$ 로 정하였다. 두 분포의 τ_2 와 κ 의 범위는 일반적인 홍수량 자료의 통상적인 τ_3 범위를 포함할 수 있도록 고려하였다. 그러나 선정된 L-CV와 κ 의 범위를 모집단으로 하는 GEV와 GLO 분포는 비현실적인 음수의 유출량이 생성되어 결과에 큰 영향을 미치므로, 본 실험에서는 생성된 유출량 중에서 음수 값을 배제한 GEV+와 GLO+ 분포를 사용하였다. Monte Carlo 분포의 중요한 3가지 절차를 설명하면 다음과 같다.

표 1. 두 지점의 자료 길이

Series	n_x	n_{xy}	n_y
1	0	40	0
2	20	40	0
3	20	40	20

1. 2변수 정규분포에서 상관계수 $-0.75 \leq \rho(z_x, z_y) \leq 0.95$ 를 갖는 z_x 와 z_y 자료를 생성한 후, 표 1의 자료 길이를 갖는 x 와 y 지점의 유출량 시나리오로 GEV와 GLO 분포로 변환한다. 이때 음수의 유출량 시나리오는 배제한다.
2. 생성된 자료로부터 x 와 y 지점의 L-moment 추정값($\hat{\tau}_2, \hat{\tau}_3$)을 추정한다. 위 과정을 10,000번 반복 수행하여 L-moment 추정값 간의 상관계수($\rho(\hat{\tau}_{2x}, \hat{\tau}_{2y}), \rho(\hat{\tau}_{3x}, \hat{\tau}_{3y}), \rho(\hat{\tau}_{2x}, \hat{\tau}_{3y})$)를 계산한다.
3. 생성된 x 와 y 지점의 모든 유출량 시나리오($10,000 \times n_{xy}$)의 상관계수 ρ_{xy} 를 계산한다.

이와 같은 실험과정을 통해 얻어진 $\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{2y})$, $\rho(\widehat{\tau}_{3x}, \widehat{\tau}_{3y})$, $\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{3y})$ 와 ρ_{xy} 의 관계는 Simple Power 함수를 이용하여 정의하였으며, 이를 수식으로 표현하면 다음과 같다.

$$\rho(\widehat{\tau}_{kx}, \widehat{\tau}_{hy}) = acf_{xy} |\rho_{xy}|^b \quad (3)$$

여기서, k 와 h 는 2또는 3을 의미하며, a 와 b 는 자료로부터 추정되는 값이다. cf_{xy} 는 두 지점의 유출량 자료간의 길이가 같지 않을 경우 이를 보정해주는 인자로서 다음과 같다.

$$cf_{xy} = \frac{n_{xy}}{\sqrt{(n_{xy} + n_x)(n_{xy} + n_y)}} \quad (4)$$

3.2 결과 분석

그림 3은 $\kappa_x = \kappa_y = -0.3$, $\tau_{2x} = \tau_{2y} = 0.3$ 인 경우 GEV+ 분포에 대한 $\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{2y})$ 와 ρ_{xy} 의 관계를 그림으로 그린 것이다. 그림에서 실선으로 표시된 Simple Power 함수가 $\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{2y})$ 와 ρ_{xy} 의 관계를 잘 표현함을 확인하였으며, 자료의 길이가 다를 경우 이를 보정하는 cf_{xy} 역시 좋은 효과를 보였다. 그림 5는 $\kappa_x = \kappa_y = -0.3$, $\tau_{2x} = \tau_{2y} = 0.3$ 인 경우 GEV+ 분포에 대한 $\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{3y})$ 와 ρ_{xy} 의 관계를 그림으로 나타낸 것으로, 상관계수 ρ_{xy} 가 1일 지라도 $\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{3y})$ 는 1이 되지 않는 특징을 보였으며, Simple Power 함수가 두 자료간의 관계를 잘 표현함을 다시 확인하였다. 그림 6은 $\kappa_x = \kappa_y = -0.1$, $\tau_{2x} = \tau_{2y} = 0.5$ 인 경우 GEV 분포와 와 음수값 유출량을 제외한 GEV+ 분포의 $\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{3y})$ 와 ρ_{xy} 의 관계를 비교한 것이다. 이때 GEV 분포의 음수확률은 8.9 %이며, 두 분포간의 상관관계 결과에 큰 차이가 있음을 확인할 수 있다.

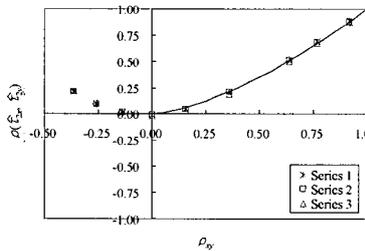


그림 3. $\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{2y})$ 와 ρ_{xy} 의 관계

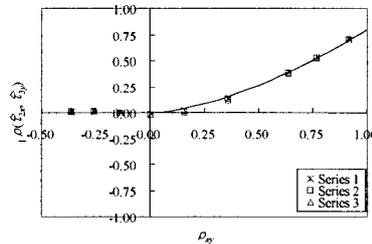


그림 4. $\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{3y})$ 와 ρ_{xy} 의 관계

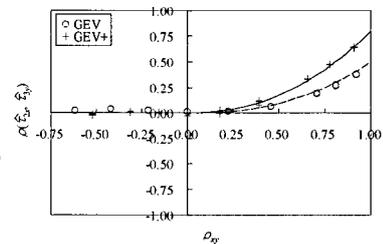


그림 5. GEV와 GEV+ 결과 비교($\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{3y})$)

표 2는 식(10)에서 설명한 $\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{2y})$, $\rho(\widehat{\tau}_{3x}, \widehat{\tau}_{3y})$, $\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{3y})$ 와 ρ_{xy} 의 관계를 나타내는 Simple Power 함수의 a 와 b 값을 정리한 것이다.

4. 결론

본 연구에서는 Monte Carlo 실험을 통해 GEV+와 GLO+ 분포의 서로 다른 두 지점에 대한 L-moment 추정값 간의 교차상관계수와 두지점 사이의 교차상관계수의 관계를 Simple Power 함수를 통해 유도하였다. 음수의 유출량 값을 배제한 GEV+와 GLO+ 분포의 결과와, 그렇지 않은 GEV와 GLO 분포의 결과를 서로

비교한 결과 큰 차이가 있음을 확인하였다. 본 연구에서 유도된 관계식은 미계측 유역이나 계측이 이루어지고 있으나 자료의 길이가 길지 않은 유역의 홍수분포 매개변수를 추정을 위해, 지점(at-site)추정된 매개변수의 정확성과 유역의 물리적 특성을 반영할 수 GLS(Generalized Least Square) 지역회기모형을 구축을 위해 기초 자료로 활용성이 클 것으로 기대한다.

표 2. GEV와 GLO 분포의 Simple Power 함수의 a 와 b 값

		GEV+ (a/b)								
		$\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{2y})$			$\rho(\widehat{\tau}_{3x}, \widehat{\tau}_{3y})$			$\rho(\widehat{\tau}_{2x}, \widehat{\tau}_{3y})$		
K_x, K_y	τ_3	$\tau_{2x}, \tau_{2y} =$ 0.1	$\tau_{2x}, \tau_{2y} =$ 0.3	$\tau_{2x}, \tau_{2y} =$ 0.5	$\tau_{2x}, \tau_{2y} =$ 0.1	$\tau_{2x}, \tau_{2y} =$ 0.3	$\tau_{2x}, \tau_{2y} =$ 0.5	$\tau_{2x}, \tau_{2y} =$ 0.1	$\tau_{2x}, \tau_{2y} =$ 0.3	$\tau_{2x}, \tau_{2y} =$ 0.5
-0.4	0.453	1.0/1.0	1.0/1.0	1.0/1.1	1.0/1.3	1.0/1.3	1.0/1.3	0.8/1.1	0.9/1.2	0.9/1.2
-0.3	0.379	1.0/1.4	1.0/1.5	1.0/1.6	1.0/2.0	1.0/2.0	1.0/2.0	0.8/1.7	0.8/1.6	0.9/1.8
-0.2	0.305	1.0/1.6	1.0/1.8	1.0/2.0	1.0/2.5	1.0/2.5	1.0/2.4	0.7/2.0	0.7/1.9	0.8/1.9
-0.1	0.236	1.0/1.8	1.0/2.0	1.0/2.2	1.0/2.8	1.0/2.9	1.0/2.7	0.6/2.3	0.6/2.3	0.7/1.9
0	0.170	1.0/1.9	1.0/2.1	1.0/2.3	1.0/3.1	1.0/3.2	1.0/2.9	0.4/2.0	0.5/2.5	0.7/2.2
0.1	0.107	1.0/2.1	1.0/2.1	1.0/2.4	1.0/3.3	1.0/3.4	1.0/3.0	0.3/3.0	0.4/2.2	0.6/2.0
0.2	0.048	1.0/2.1	1.0/2.1	1.0/2.4	1.0/3.4	1.0/3.5	1.0/3.1	0.1/2.8	0.3/1.5	0.5/1.6
0.3	-0.090	1.0/2.0	1.0/2.1	1.0/2.3	1.0/3.4	1.0/3.5	1.0/3.1	0.0/ *	0.3/1.3	0.5/1.6
0.4	-0.063	1.0/1.9	1.0/2.1	1.0/2.2	1.0/3.4	1.0/3.4	1.0/3.0	0.0/ *	0.3/1.2	0.5/1.5
0.5	-0.114	1.0/1.8	1.0/2.0	1.0/2.2	1.0/3.3	1.0/3.3	1.0/2.9	0.0/ *	0.3/1.0	0.5/1.4
		GLO+ (a/b)								
-0.4	0.4	1.0/1.1	1.0/1.1	1.0/1.2	1.0/1.4	1.0/1.4	1.0/1.4	0.8/1.1	0.9/1.3	0.9/1.3
-0.3	0.3	1.0/1.6	1.0/1.6	1.0/1.7	1.0/2.1	1.0/2.1	1.0/2.1	0.8/1.9	0.8/1.9	0.9/1.9
-0.2	0.2	1.0/1.9	1.0/2.0	1.0/2.1	1.0/2.7	1.0/2.7	1.0/2.5	0.6/2.0	0.6/2.1	0.8/2.1
-0.1	0.1	1.0/2.0	1.0/2.2	1.0/2.3	1.0/3.0	1.0/3.2	1.0/2.8	0.3/2.1	0.5/2.4	0.7/2.2
0	0	1.0/2.0	1.0/2.3	1.0/2.5	1.0/3.1	1.0/3.5	1.0/3.1	0.0/ *	0.3/1.6	0.6/2.3
0.1	-0.1	1.0/1.9	1.0/2.3	1.0/2.5	1.0/3.4	1.0/3.7	1.0/3.3	0.0/ *	0.2/1.2	0.4/1.5

여기서 *는 식(5)에서 제한한 Simple Power 함수의 a 값이 0이므로 b 값을 추정할 수 없음

감 사 의 글

이 논문은 2006년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임.” (KRF-2006-214-D00174)

참 고 문 헌

1. Jenkinson, A. F.(1995). “The frequency distribution of annual maximum (or minimum) values of Meteorological elements.” *Quart. J. Roy. Met. Soc.* Vol. 81, pp. 158-171.
2. Robson, A. J. and Reed, D.(1999). *The Flood Estimation Handbook*, Wallingford, UK.
3. Stedinger, J. R and Tasker, G.(1985) “Regional hydrologic analysis, 1, Ordinary, weighted and generalized least squares compared.” *Water Resour. Res.*, Vol. 22(5), pp. 1421-1432, 1985.
4. Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E. (1993). *Frequency Analysis of Extreme Events*, Chapter 18, *Handbook of Hydrology*, D. Maidment (ed.), McGraw-Hill, Inc., NY, USA.
5. Hosking, J. R. M. and Wallis, J. R. (1997). *Regional Frequency Analysis: An Approach Based on L-moments*, Cambridge University Press, Cambridge, UK.
6. Martins, E. S. and Stedinger, J. R.(2002). “Cross correlations among estimators of shape.” *Water Resour. Res.* Vol. 38(11), 1252, doi:10.1029/2002WR001589.