GEV 분포형을 이용한 홍수빈도해석에서의 불확실성 해석

이동진[1], 허준행[2]

1. 서론

홍수빈도해석시 추정된 확률홍수량은 여러 종류의 불확실성을 내포하고 있으며 특히 제한된 양의 홍수자료를 이용하여 모집단의 분포형을 추정하는 과정에서 발생하는 불확실성이 대표적이라 할 수 있다. 본 연구에서는 홍수빈도해석을 통하여 추정된 확률홍수량의 초과확률과 재현기간의 불확실성 및 확률홍수량을 통하여 결정되는 수공구조물의 위험도의 불확실성을 추정된 각 매개변수의 분산량을 이용하여 규명하고자 한다. 매개변수 추정치의 기대값과 분산량을 산정하기 위하여 수문자료 해석에 널리 이용되고 있는 GEV 분포형을 적정분포형으로 가정하여 각 매개변수의 불확실성을 정량적으로 규명하였다. 매개변수의 분산량과 초과확률, 재현기간, 위험도, 수공구조물의 내구연한과의 관계를 이용하여 위험도의 분산량을 계산하였으며, 매개변수 추정방법별, 재현기간, 위험도, 수공구조물의 내구연한, 매개변수의 변화에 따른 위험도의 분산량을 비교분석 하였다.

2. GEV(Generalized Extreme Value) 분포형

2변수 GEV 분포형의 누가분포함수(Cumulative Distributed Function; CDF)와 확률밀도함수(Probability Density Function; PDF)는 다음과 같이 정의된다(NERC, 1975).

$$F(x) = \exp\left[-\left\{1-\frac{\beta}{\alpha}x\right\}^{1/\beta}\right] \tag{1}$$

$$f(x) = \frac{1}{\alpha}\left[1-\frac{\beta}{\alpha}x\right]^{1/\beta-1}\exp\left[-\left\{1-\frac{\beta}{\alpha}x\right\}^{1/\beta}\right] \tag{2}$$

여기서, α는 규모매개변수(scale parameter), β는 형상매개변수(shape parameter)를 나타낸다. 규모매개변수 α는 항상 0보다 크며, 또한 $\beta>0$인 경우 $x<\alpha/\beta$ 이고 $\beta<0$인 경우 $x>\alpha/\beta$인 조건을 만족해야 한다. GEV 분포형을 이용한 확률홍수량은 식 (1)로부터 역함수의 형태로 산정되며, 해당 재현기간의 역수로 정의되는 비초과확률에 대한 확률홍수량은 다음 식 (3)과 같다.

$$x = \frac{\alpha}{\beta}[1-(-\ln q)^{\beta}] \tag{3}$$

여기서, q는 주어진 홍수량에 대해 그 값을 넘지 않을 비초과확률을 나타낸다.

한편, 식 (1)로부터 산정된 확률홍수량에 대한 비초과확률을 산정하는 식을 재정렬하면 다음과 같다.

$$\bar{q} = \exp\left[-\left\{1-\frac{\beta}{\alpha}x\right\}^{1/\beta}\right] = \exp\{-\exp(-y)\} \tag{4}$$

여기서, $y = \frac{1}{\beta}\ln\left[1-\frac{\beta}{\alpha}x\right]$와 같다.

3. 위험도의 불확실성

댐, 암거, 제방과 같은 수공구조물의 파괴위험도는 설계에 기준이 되는 재현기간 또는 설

---

1) 정회원, 연세대학교 토목공학과 박사과정수료
2) 정회원, 연세대학교 사회환경시스템공학부 토목공학전공 교수

계홍수량의 초과확률과 직접적으로 관련이 된다. 이러한 경우 파괴위험도는 n 기간동안 설계홍수량보다 큰 홍수량이 발생할 확률이 1보다 크거나 같게 되는 확률로 정의된다. 따라서, 발생하는 연최대홍수량이 독립적이라고 가정하에 파괴위험도는 다음 식 (5)와 같이 정의할 수 있다(Yen 1970).

$$R = 1-(1-p)^n = 1-q^n = 1-\left(1-\frac{1}{T}\right)^n$$
$$= 1-\exp\left[-n\left\{1-\frac{\beta}{\alpha}x\right\}^{1/\beta}\right] = \exp\{-n\exp(-y)\} \tag{5}$$

여기서, R은 파괴위험도, p는 초과확률, q는 비초과확률, T는 재현기간, n은 내구연한을 의미한다.

본 연구에서는 파괴위험도 R의 불확실성을 기대값과 분산량을 산정함으로써 정량적으로 규명하였으며, 우선적으로 2변수 GEV 분포형에 대하여 확률홍수량 x_q에 대한 파괴위험도 R의 기대값을 Taylor 전개를 통한 1차 근사식으로 산정하면 다음 식 (6)과 같다.

$$E(\hat{R}) \approx 1-[E(\hat{q})]^n = 1-\left[\exp\left\{-\left(1-\frac{E(\hat{\beta})}{E(\hat{\alpha})}x_q\right)^{1/E(\hat{\beta})}\right\}\right]^n \tag{6}$$

여기서, 추정된 규모매개변수 α와 β가 불편의량이라 가정하면 $E(q) \approx q$가 되고 따라서, 식 (6)은 다음 식 (7)과 같이 정리될 수 있다.

$$E(\hat{R}) \approx 1-q^n = \exp\left[-\left\{1-\frac{\beta}{\alpha}x_q\right\}^{1/\beta}\right]^n \tag{7}$$

동일한 방법으로 파괴위험도 R의 분산량을 Taylor 급수를 통한 1차 근사화한 식으로 나타내면 다음 식 (8)과 같다.

$$Var(\hat{R}) \approx \left(\frac{\partial\hat{R}}{\partial\hat{\alpha}}\right)_\mu^2 Var(\hat{\alpha}) + 2\left(\frac{\partial\hat{R}}{\partial\hat{\alpha}}\right)_\mu\left(\frac{\partial\hat{R}}{\partial\hat{\beta}}\right)_\mu Cov(\hat{\alpha},\hat{\beta}) + \left(\frac{\partial\hat{R}}{\partial\hat{\beta}}\right)_\mu^2 Var(\hat{\beta}) \tag{8}$$

여기서, 각항 괄호밖 μ는 각 미분항에서 확률변수들을 기대값으로 사용하는 것을 의미한다.

식 (5)로부터 $\left(\frac{\partial\hat{R}}{\partial\hat{\alpha}}\right)_\mu$와 $\left(\frac{\partial\hat{R}}{\partial\hat{\beta}}\right)_\mu$를 유도하여 정리하면 다음과 같다.

$$\frac{\partial\hat{R}}{\partial\hat{\alpha}} = \frac{n}{\alpha\beta}(Q^{1-\beta}-Q)\,q^n \tag{9}$$

$$\frac{\partial\hat{R}}{\partial\hat{\beta}} = \frac{n}{\beta}\left\{-\ln Q\times Q-\frac{1}{\beta}(Q^{1-\beta}-Q)\right\}q^n \tag{10}$$

여기서, $Q=[-\ln q]$와 같다.

4. 매개변수의 분산량 산정

식 (8)에서 제시한 파괴위험도를 산정하기 위해서는 규모매개변수와 형상매개변수의 분산량 및 공분산량의 산정이 우선되어야 한다. 본 연구에서는 대표적인 매개변수 추정방법인 모멘트법(Methods of Moments; MOM), 최우도법(Maximum Likelihood Methods; ML), 확률가중모멘트법(Prabability Weighed Moments; PWM)을 이용하여 2변수 GEV 분포형의 매개변수 값의 분산량 및 공분산량을 각각 산정하였다.

4.1 모멘트법

모멘트법을 이용하여 GEV 분포형의 매개변수 분산량 및 공분산량을 산정하기 위해서는 우선적으로 중심극한정리를 이용하여 표본의 모멘트와 모집단 모멘트의 차가 Bivariate

Normal 분포형으로 수렴한다는 가정하에 표본모멘트들과 분포형의 매개변수의 관계를 이용하여 변환(transformation)시킴으로써 최종적으로 각 매개변수들의 분산량 및 공분산량을 산정할 수 있다. 모멘트법을 이용하여 산정된 각 매개변수의 분산량과 공분산량은 다음 식 (11) - 식 (13)과 같다.

$$Var(\hat{\alpha})_{MOM} = \frac{A_1^2 C_{11} + 2A_1 A\beta C_{13} H + A_\beta^2 C H^2}{N} \tag{11}$$

$$Cov(\hat{\alpha}, \hat{\beta})_{MOM} = \frac{A_1 C_{13} H + A_\beta C H^2}{N} \tag{12}$$

$$Var(\hat{\beta})_{MOM} = \frac{CH^2}{N} \tag{13}$$

여기서, A_1, A_β는 변환시 사용되는 Jacobain Matrix의 구성요소들로서 다음과 같다.

$$A_1 = \beta / [1 - \Gamma(1 + \beta)] \tag{14}$$

$$A_\beta = \frac{\alpha}{\beta[1-\Gamma(1+\beta)]}[1-\Gamma(1+\beta) + \beta\Gamma'(1+\beta)] \tag{15}$$

또한, C_{11}, C_{13}, C, H 변수들 역시 변환과정에서 발생하는 Jacobain Matrix의 구성요소들로서 식으로 정의하면 다음과 같다.

$$H = \frac{[1-\Gamma(1+\beta)]^3}{2\{\Gamma'(1+2\beta)[1-\Gamma(1+\beta)]+\Gamma'(1+\beta)[\Gamma(1+2\beta) - \Gamma(1+\beta)]\}} \tag{16}$$

$$C = 4C_{11}\frac{M_2^2}{M_1^6} - 4C_{12}\frac{M_2}{M_1^5} + \frac{C_{22}}{M_1^4} \tag{17}$$

$$C_{11} = \left(\frac{\alpha}{\beta}\right)^2 [\Gamma(1+2\beta) - \Gamma^2(1+\beta)] \tag{18}$$

$$C_{13} = -2C_{11}\frac{M_2^2}{M_1^3} + C_{12}\frac{1}{M_1^2} \tag{19}$$

$$C_{12} = \left(\frac{\alpha}{\beta}\right)^3 [-2\Gamma^2(1+\beta) + 2\Gamma(1+2\beta) + \Gamma(1+\beta)\Gamma(1+2\beta) - \Gamma(1+3\beta)] \tag{20}$$

$$\begin{aligned} C_{22} = \left(\frac{\alpha}{\beta}\right)^4 [\ &\Gamma(1+4\beta) - 4\Gamma(1+3\beta) - \Gamma^2(1+2\beta) + 4\Gamma(1+2\beta) \\ &+ 4\Gamma(1+\beta)\Gamma(1+2\beta) - 4\Gamma^2(1+\beta)] \end{aligned} \tag{21}$$

4.2 최우도법

식 (1)로부터 유도된 GEV 분포형에 대한 최우도함수는 다음 식 (22)와 같다.

$$LL = -N\ln(\alpha) - (1-\beta)\sum_{i=1}^{N} y_i - \sum_{i=1}^{N} e^{-y_i} \tag{22}$$

여기서, $y_i = -\frac{1}{\beta}\ln\left[\frac{1-\beta}{\alpha}x_i\right]$이며, 식 (22)를 각 매개변수들을 이용하여 2차 미분한 값을 기대값으로 취하는 Information Matrix에 대하여 그 역행렬을 구성하는 요소는 각 매개변수들의 분산량 및 공분산량과 같다. 따라서, 최우도법을 이용한 매개변수들의 분산량 및 공분산량은 다음 식 (23) - 식 (25)와 같이 유도할 수 있다.

$$Var(\hat{\alpha})_{ML} = \frac{\alpha^2\beta^2}{N}\frac{C_B}{C_A C_B - C_{AB}^2} \tag{23}$$

$$Cov(\hat{\alpha}, \hat{\beta})_{ML} = \frac{\alpha\beta^2}{N}\frac{C_{AB}}{C_A C_B - C_{AB}^2} \tag{24}$$

$$Var(\hat{\beta})_{ML} = \frac{\beta^2}{N}\frac{C_A}{C_A C_B - C_{AB}^2} \tag{25}$$

여기서, $C_A = 1 - 2\Gamma(2-\beta) + (1-\beta)^2\Gamma(1-2\beta)$ (26)

$$C_B = 1 - \gamma_E - \frac{(1-\beta)^2}{\beta}\Gamma(1-2\beta) - \frac{1}{\beta}[1-\Gamma(2-\beta)] - \Gamma(2-\beta)\left[\psi(1-\beta) - \frac{1-\beta}{\beta}\right] \tag{27}$$

$$C_{AB} = 1 - \gamma_E - \frac{(1-\beta)^2}{\beta}\Gamma(1-2\beta) - \frac{1}{\beta}[1-\Gamma(2-\beta)] - \Gamma(2-\beta)\left[\psi(1-\beta) - \frac{1-\beta}{\beta}\right] \tag{28}$$

4.3 확률가중모멘트법

확률가중모멘트(Probability Weighted Moments)를 이용하여 규모매개변수와 위치매개변수를 추정하는 경우 양해적으로 산정할 수 있는 방법이 없으므로, PWM과 매개변수의 관계를 이용하여 다음과 같은 여러 단계의 변환과정을 거쳐 산정한다(Heo 등, 1990).

$$\begin{matrix} \hat{B}_0 \\ \hat{B}_1 \end{matrix} \rightarrow \begin{matrix} \hat{B}_0 \\ \hat{B}_1 \\ R \end{matrix} \rightarrow \begin{matrix} \hat{B}_0 \\ \hat{B}_1 \\ \beta \end{matrix} \rightarrow \begin{matrix} \hat{\alpha} \\ \hat{\beta} \end{matrix}$$

여기서, B_r은 GEV 분포형의 모집단 PWM으로 다음 식 (29)와 같이 표현할 수 있다(Hosking 등, 1985).

$$B_r = \left(\frac{\alpha}{\beta}\right)\frac{1}{r+1}\left[1 - \frac{\Gamma(1+\beta)}{(r+1)^\beta}\right] \tag{29}$$

위의 변환단계중 첫번째 변환단계에서의 제시한 $R = \hat{B}_0/2\hat{B}_1$과 같이 β의 함수로 표현할 수는 있지만 직접적으로 $\hat{\beta}$를 산정할 수 없으며 두번째 변환단계에서 음해적으로 $\hat{\beta}$를 산정한다. 이상과 같은 방법으로 PWM을 이용하여 규모매개변수 α와 형상매개변수 β의 분산량 및 공분산량을 산정하면 다음과 같다.

$$Var(\hat{\alpha})_{PWM} = \frac{A_0^2 D_{00} + 2A_0 A_\beta C_1^* K + A_\beta^2 C^* K^2}{N} \tag{30}$$

$$Cov(\hat{\alpha}, \hat{\beta})_{PWM} = \frac{A_0 C_1^* K + A_\beta C^* K^2}{N} \tag{31}$$

$$Var(\hat{\beta})_{PWM} = \frac{C^* K^2}{N} \tag{32}$$

여기서,

$$C^* = \frac{2\beta}{\alpha\{1-2^{-\beta}\Gamma(1+\beta)\}}\left[\{1-2^{-\beta}\Gamma(1+\beta)\}\frac{C_1^*}{2} - \{1-\Gamma(1+\beta)\}C_2^*\right] \tag{33}$$

$$C_1^* = \frac{2\beta}{\alpha\{1-2^{-\beta}\Gamma(1+\beta)\}}\left[\{1-2^{-\beta}\Gamma(1+\beta)\}\frac{D_{00}}{2} - \{1-\Gamma(1+\beta)\}D_{01}\right] \tag{34}$$

$$C_2^* = \frac{2\beta}{\alpha\{1-2^{-\beta}\Gamma(1+\beta)\}}\left[\{1-2^{-\beta}\Gamma(1+\beta)\}\frac{D_{01}}{2} - \{1-\Gamma(1+\beta)\}D_{11}\right] \tag{35}$$

$$K = \frac{\Gamma(1+\beta)}{[1-2^{-\beta}\Gamma(1+\beta)]^2}[\{-1-2^{-\beta}\Gamma(1+\beta)\}\psi(1+\beta) - \{1-\Gamma(1+\beta)\}\{2^{-\beta}\ln 2 - 2^{-\beta}\psi(1+\beta)\}] \tag{36}$$

또한, D_{00}, D_{01}, D_{11}은 표본 PWM $\hat{B}_0$, $\hat{B}_1$ 의 점근적 분산 및 공분산을 나타내며 다음 식 (36) - 식 (38)과 같다(Hosking 등, 1985). 여기서, 식 (36)은 모멘트법에서 제시하였던 식

(18)과 동일함을 알 수 있다.

$$D_{00} = \left(\frac{\alpha}{\beta}\right)^2 [\Gamma(1+2\beta) - \Gamma^2(1+\beta)] \tag{36}$$

$$D_{01} = \frac{1}{2}\left(\frac{\alpha}{\beta}\right)^2 [2^{-2\beta}\Gamma(1+2\beta) + (1-2^{-\beta})\Gamma^2(1+\beta)] \tag{37}$$

$$D_{11} = \left(\frac{\alpha}{\beta}\right)^2 2^{-2\beta} [H(1/2)\Gamma(1+2\beta) - \Gamma^2(1+\beta)] \tag{38}$$

여기서, $H(z)$는 hypergeometric function을 의미한다.

5. 비교분석

홍수빈도해석을 이용한 수공구조물의 위험도를 분석하기 위한 위험도의 불확실성은 식 (8)의 위험도에 대한 분산량으로부터 정량적으로 산정할 수 있다. 본 연구에서는 서로 다른 재현기간 및 내구연한, 그리고 매개변수 추정방법별로 추정된 규모매개변수와 형상매개변수의 분산량 및 공분산량에 따른 위험도의 불확실성의 비교를 실시하였다. 재현기간의 경우 10, 20, 50, 100, 200, 1000년을 적용하였으며, 수공구조물의 내구연한은 2, 10, 50, 100, 200, 500, 1000년을 고려하였다. 표본의 자료수는 10, 25, 50, 100개로 변화를 시켜 그 결과를 분석하였다. 한편 GEV 분포형의 규모매개변수 α는 1로 가정하였으며, 형상매개변수 β는 -0.2, -0.1085, 0.0053, 0.1, 0.2776, 0.4로 변화시키면서 각각의 경우에 대한 위험도의 분산량을 산정하였다. 그 결과, 내구연한의 변화에 대한 각 매개변수 추정방법별 분산량의 크기는 전반적으로 감소하였으며, 동일한 내구연한에서의 각 추정방법별 상대적인 차는 내구연한의 변화에 관계없이 동일한 양상을 보이는 것으로 나타났다. 자료수 및 재현기간, 형상매개변수의 변화에 따른 각 방법별 분산량의 비교분석 결과를 정리하면 다음 표 1과 같다. 표 1에서와 같이 형상매개변수의 크기가 작은 경우 재현기간이 증가할수록 확률가중모멘트법을 이용하여 추정된 분산량의 값이 적어지는 것을 볼 수 있으며, 형상매개변수 값이 증가할수록 최우도법을 이용한 분산량 값이 적어지는 것을 볼 수 있다. 또한 재현기간이 작아지고 형상매개변수의 크기가 증가할수록 추정된 위험도의 분산량의 크기는 확률가중모멘트법, 모멘트법, 최우도법의 순으로 상대적인 크기가 작아지는 것을 볼 수 있다.

6. 결론

본 연구는 GEV분포형을 이용하여 홍수량을 빈도해석하고 이를 수공구조물에 적용할 때 발생되는 위험도의 불확실성을 다양한 조건하에서 각 매개변수 추정방법별로 비교한 것으로 얻어진 결론은 다음과 같다.

1) 재현기간이 증가할수록, 형상매개변수의 크기가 감소할수록 매개변수 추정방법별 상대적인 위험도의 분산량 크기는 모멘트법, 최우도법, 확률가중모멘트법의 순으로 작아지는 것으로 나타났다.
2) 형상매개변수의 크기는 매개변수 추정방법별 상대적인 크기를 비교시 큰 영향을 미치는 것으로 나타났으며, 형상매개변수의 크기가 증가할수록 확률가중모멘트법을 이용한 추정량에서 가장 큰 분산량을 보이며 최우도법에 의한 분산량이 가장 작은 값을 보이는 것으로 나타났다.
3) 표본자료의 수를 증가시키면서 위험도의 분산량 크기를 비교한 결과 표본결과에는 크게 영향을 받지 않고 비슷한 양상을 나타나는 것으로 나타났다.
4) 설계내구연한을 증가시켜 위험도의 분산량을 비교한 결과 분산량의 크기는 점점 감소하는 것으로 나타났으나, 매개변수 추정방법별 분산량의 상대적인 크기는 그대로 유지되

는 것으로 나타났다.

표 1. 매개변수 추정방법별 분산량 비교

자료수 (N)	형상매개변수 (β) 및 해당 왜곡도계수(γ)	재현기간 ($T=\frac{1}{1-q}$)					
		10	20	50	100	200	1000
10	-0.2000 (γ=3.535)	②	②	②	②	③	③
	-0.1085 (γ=1.999)	②	②	②	②	②	③
	-0.0053 (γ=1.108)	①	①	②	②	②	②
	0.1000 (γ=0.638)	①	①	②	②	②	②
	0.2776 (γ=0.000)	①	①	①	①	②	②
	0.4000 (γ=0.359)	①	①	①	①	①	②
25	-0.2000 (γ=3.535)	②	②	②	②	③	③
	-0.1085 (γ=1.999)	②	②	②	②	②	③
	-0.0053 (γ=1.108)	①	①	②	②	②	②
	0.1000 (γ=0.638)	①	①			②	②
	0.2776 (γ=0.000)	①	①	①	①	②	②
	0.4000 (γ=0.359)	①	①	①	①	①	
50	-0.2000 (γ=3.535)	②	②	②	②	③	③
	-0.1085 (γ=1.999)	②	②	②	②	②	③
	-0.0053 (γ=1.108)	①	①	②	②	②	②
	0.1000 (γ=0.638)	①	①	②	②	②	②
	0.2776 (γ=0.000)	①	①	①	①	②	②
	0.4000 (γ=0.359)	①	①	①	①	①	②
100	-0.2000 (γ=3.535)	②	②	②	②	③	③
	-0.1085 (γ=1.999)	②	②	②	②	②	③
	-0.0053 (γ=1.108)	①	①	②	②	②	②
	0.1000 4(γ=0.638)	①	①	①	②	②	②
	0.2776 (γ=0.000)	①	①	①	①	②	②
	0.4000 (γ=0.359)	①	①	①	①	①	②

주) ① : $Var(\hat{\beta})_{ML} < Var(\hat{\beta})_{MOM} < Var(\hat{\beta})_{PWM}$, ②: $Var(\hat{\beta})_{ML} < Var(\hat{\beta})_{PWM} < Var(\hat{\beta})_{MOM}$

③ : $Var(\hat{\beta})_{PWM} < Var(\hat{\beta})_{ML} < Var(\hat{\beta})_{MOM}$

7. 참고문헌

NERC(Natural Environmental Research Council) (1975). *Flood Study Report*, Vol. 1, Hydrological Studies, Whitefriars Press Ltd., London.

Heo, J.H., Boes, D.C., and Salas, J.D. (1990) Regional Flood Frequency Modeling and Estimation, *Water Resources Papers*, No. 101, Colorado State University, Fort Collins, Colorado.

Hosking, J.R.M., Wallis, J.R., and Wood, E.F. (1985) Estimation of the Generalized extreme-value distribution by the method of probability weighted moments, *Technometrics*, 27, pp. 183-257.

Yen, B. C. (1970). "Risks in hydrologic design of engineering projects." *J. Hydr.* Engrg., ASCE, 96(HY4), 959-966.