韓國水資源學會論文集
第39卷 第8號 · 2006年 8月
pp. 659~668

Gumbel분포형을 이용한 위험도에 관한 불확실성 해석

A Study on Uncertainty of Risk of Failure Based on Gumbel Distribution

허 준 행* / 이 동 진** / 신 홍 준*** / 남 우 성****
Heo, Jun-Haeng / Lee, Dong-Jin / Shin, Hongjoon / Nam, Woosung

Abstract

The uncertainty of the risk of failure of hydraulic structures can be determined by estimating the variance of the risk of failure based on the methods of moments, probability weighted moments, and maximum likelihood assuming that the underlying model is the Gumbel distribution. In this paper, the variance of the risk of failure was derived. Monte Carlo simulation was performed to verify the characteristics of the derived formulas for various sample size, design life, nonexceedance probability, and variation coefficient. As the results, PWM showed the smallest relative bias and root mean square error than the others while ML showed the smallest ones for relatively large sample sizes regardless of design life and nonexceedance probability. Also, it was found that variation coefficient does not effect on the relative bias and relative root mean square error.

keywords : uncertainty, risk of failure, variance, Monte Carlo simulation, Gumbel distribution

요 지

수공구조물의 위험도에 관한 불확실성을 검토하기 위하여 본 연구에서는 빈도해석을 통하여 추정되는 설계홍수량의 분산량을 고려한 불확실성 해석을 실시하였다. Gumbel 분포형을 기본 분포형으로 가정하였으며, 모멘트법, 최우도법, 확률가중모멘트법을 이용하여 각 매개변수 추정방법별로 추정된 설계홍수량에 대한 이론적인 분산량을 산정하였다. 이론적으로 유도한 분산량의 특성을 규명하기 위하여 다양한 표본크기와 설계연한, 비초과확률 및 변동계수 조건에 대하여 Monte-Carlo 모의를 실시하고 각 매개변수 추정방법별 비교를 실시하였다. 그 결과 확률가중모멘트법을 사용한 경우 위험도에 대하여 상대적으로 가장 작은 상대편의 및 상대제곱근오차를 발생시키는 것으로 나타났으며, 최우도법의 경우에는 상대적으로 큰 표본자료에 대해서는 설계연한 및 비초과확률에 관계없이 작은 상대편의 및 상대제곱근오차를 발생시키는 것으로 나타났다. 또한 다양한 변동계수 조건은 상대편의 및 상대제곱근오차의 측면에서 고려하여 볼 때 거의 영향을 주지 않는 것으로 나타났다.

핵심용어 : 불확실성, 위험도, 분산, 모의실험, Gumbel 분포형

* 연세대학교 공과대학 사회환경시스템공학부 교수
(e-mail: jhheo@yonsei.ac.kr)
** Colorado State University, Dept. of Civil Engrg., 박사과정
(e-mail: djlee@engr.colostate.edu)
*** 연세대학교 대학원 토목공학과 박사과정
(e-mail: sinong@yonsei.ac.kr)
**** 연세대학교 대학원 토목공학과 박사과정
(e-mail: nws77@yonsei.ac.kr)

1. 서 론

일반적으로 수공구조물의 설계를 위해서는 홍수량에 대한 빈도해석 결과가 이용되고 있으며, 수공구조물의 규모는 설계 재현기간에 기초하여 빈도해석을 통해 최종적으로 계산되는 설계홍수량을 이용하여 결정하게 된다. 또한 수공구조물이 이미 건설되었거나 운영되고 있는 경우, 구조물의 홍수방어능력을 재검토하기 위해서는 설계홍수량에 대한 해당 재현기간의 크기 또는 비초과확률을 재산정 할 필요성이 대두된다.

연최대자료를 이용한 빈도해석시에는 이용 가능한 홍수량자료에 대하여 여러 확률분포형을 적용한 후 분포형의 매개변수를 추정하게 되고 각 확률분포형들이 홍수량자료의 특성을 얼마나 잘 재현하는지 여부에 따라 적정 확률분포형을 결정하며 이를 이용하여 해당 재현기간에 대한 확률홍수량을 결정하게 된다. 하지만 실제로 사용 가능한 수문자료는 그 수가 제한적이기 때문에 추정되는 확률분포형의 매개변수와 확률홍수량은 불확실성을 가지게 되며, 지금까지 이러한 불확실성을 정량적으로 규명하기 위해서 많은 연구가 진행되어 왔다(Kite, 1988; Chowdhury and Stedinger, 1991). 또한, 설계홍수량에 대하여 재현기간을 결정하는 경우에도 산정된 재현기간은 불확실성을 가지게 되며 결과적으로 해당 수공구조물의 위험도 및 신뢰도 역시 불확실성을 가지게 된다. 하지만 재현기간의 불확실성 규명에 관한 연구에 비하여 빈도해석에 기초한 구조물의 위험도의 불확실성 해석에 관한 연구는 지금까지는 거의 없는 실정이다. 따라서 본 연구에서는 빈도해석을 바탕으로 수공구조물의 위험도의 불확실성을 정량적으로 규명하고자 하며 이를 위한 일련의 해석과정을 제안하고자 한다.

2. 재현기간과 위험도

재현기간과 위험도는 상호 밀접한 관계를 가진다. 일반적으로 수공구조물의 설계시 재현기간의 개념은 해당 홍수량을 기준하여 이보다 같거나 큰 양이 다시 처음으로 발생하는 데 소요되는 평균시간으로 정의된다(Benjamin and Cornell, 1970). 만일 발생한 홍수들이 서로 독립적이며 홍수량을 초과하여 발생할 확률 p 가 연도별로 일정하다면, 연도에 기초한 재현기간 T 는 $1/p$로 정의될 수 있다. 지금까지 재현기간의 개념을 규명하기 위하여 여러 연구가 진행되어 왔는데, 대표적인 연구로서 Vogel (1987)은 여러 통계적 조건하에서 재현기간의 개념을 검토하였으며 Loaiciga and Mariño (1991)는 확률분포형의 정상성(consistency)에 기초하여 재현기간의 개념과 관련된 연구를 수행하였다. 또한, 홍수빈도해석 및 발생 홍수량의 재현기간 결정과 관련된 일반적인 연구와 함께 홍수량의 발생이 서로 독립적이고 통계학적으로 동질하다는 가정에 기초한 관련 연구들도 볼 수 있다(NERC, 1975; Kite, 1988; Stedinger et al., 1993). 극치계열에 대한 재현기간의 불확실성에 대한 연구로서 Davis et al.(1972)은 Bayesian 방법을 이용하여 Poisson 분포 및 지수분포를 가지는 경우 연최대 강우량에 대한 재현기간의 분포에 관한 연구를 수행하였으며, Rasmussen and Rosbjerg(1989)는 부분계열에 대하여 위험도개념을 도입하였고, 또한 Rasmussen and Rosbjerg(1991)는 계절적 요인을 고려한 부분시계열에 대하여 지수분포와 Poisson 분포를 가지는 설계홍수량의 정확도 문제에 대한 연구를 수행한 바 있다. Rosbjerg and Madsen(1992)은 평균제곱근오차에 입각하여 모멘트법과 확률가중모멘트법에 의한 방법별 추정도를 비교, 분석하였으며, Watt and Wilson(1978)는 항만보호공 및 수로개선 문제에 대하여 재현기간과 관련된 불확실성을 고려하였다. Thompson et al.(1997)은 다양한 해석적 방법들과 Monte Carlo 모의기법을 이용하여 댐붕괴 확률 및 피해분포 등을 해석하였다.

암거, 교각 등과 관련된 수공구조물의 홍수위험도는 다음과 같이 재현기간 T 또는 초과확률 p를 이용하여 정의할 수 있으며, 여기서 위험도의 불확실성은 설계연한 n 년 동안에 기준홍수량보다 크거나 같은 홍수가 발생할 확률로 정의된다. 따라서 발생하는 연최대홍수량이 서로 독립적이라고 가정한다면 n 년 동안의 수공구조물에 대한 위험도는 다음 Eq. (1)과 같다(Yen, 1970).

$$\hat{R} = 1-(1-\hat{p})^n = 1-\hat{q}^n = 1-(1-1/T)^n \qquad (1)$$

여기서 R 은 위험도, n 은 설계연한, q 는 비초과확률과 같다. 물론 수공구조물의 경우 수문학적인 요인 이외에 다른 중요한 원인들에 의해서 파괴가 발생할 수 있으나(Kite, 1988; Chow et al., 1988), 본 연구에서는 수문학적인 원인들에 의해서 발생하는 구조물의 파괴위험도만을 고려하기로 한다.

3. 위험도의 불확실성

일반적으로 빈도해석을 통하여 계산되는 재현기간 $T=1/p$에 대한 설계홍수량은 가정된 누가분포함수 $F(x_q)=q=1-p$의 역함수을 이용한 해를 산정함으로

써 추정할 수 있다. 하지만 실제로 주어진 재현기간에 대한 홍수량은 Fig. 1(a)에서 나타낸 바와 같이 불확실성을 가지게 되며, 이러한 설계홍수량을 기준한 재현기간 역시 수공구조물의 시공완료 후 그 값이 변화할 불확실성을 내포하고 있다. 또한 적정분포형으로 채택한 확률분포형 역시 수문자료가 보완, 확충됨으로써 다른 분포형으로 바뀔 가능성을 가지고 있으며, 설계 당시와 동일한 분포형이 적정분포형으로 계속 사용된다 하더라도 분포형의 매개변수는 추가된 수문자료에 따라 바뀌게 될 것이다. 이러한 경우 기추정된 설계홍수량에 대한 재현기간을 새로이 산정할 필요가 있으며, 추정된 설계홍수량에 대한 재현기간은 Fig. 1(b)에서 도시한 바와 같이 불확실성을 가지게 된다.

본 연구에서는 수공구조물의 파괴위험도 추정치에 대한 불확실성을 규명하기 위하여 Gumbel 분포형을 적정분포형으로 가정하여 위험도의 불확실성을 정량적으로 해석하고자 한다. Gumbel 분포형의 누가분포함수(cumulative distribution function, CDF)는 다음과 같이 정의할 수 있다.

$$F(x) = \exp\left\{-\exp\left[-\frac{x-x_0}{\alpha}\right]\right\} \tag{2}$$

여기서, x_0와 α는 Gumbel 분포형의 위치매개변수(location parameter)와 규모매개변수 (scale parameter)이며, 표본자료 $X_1,\cdots,X_N$ (N은 자료개수)로부터 추정된 매개변수의 추정치를 각각 $\hat{x_0}$와 $\hat{\alpha}$로 정의한다. Eq. (2)로부터 계산되는 임의의 비초과확률 q에 대한 설계홍수량 x_q는 다음 Eq. (3)과 같다.

$$\hat{q} = \exp\left\{-\exp\left[-\frac{x_q-\hat{x_0}}{\hat{\alpha}}\right]\right\} \tag{3}$$

이 경우 설계연한 n 년에 대한 수공구조물의 위험도는 다음과 같이 정의할 수 있다.

$$\hat{R} = 1-(1-\hat{p})^n = 1-\hat{q}^n \tag{4}$$

Eq. (4)로부터 추정되는 위험도 $\hat{R}$의 기대값 $E(\hat{R})$을 Taylor 급수를 이용하여 산정하면 Eq. (5)와 같으며, 비초과확률에 대한 기대값 $E(\hat{q})$는 다음 Eq. (6)과 같다.

$$E(\hat{R}) \approx 1-[E(\hat{q})]^n \tag{5}$$

$$E(\hat{q}) \approx \exp\left\{-\exp\left[-\frac{x_q-E(\hat{x_0})}{E(\hat{\alpha})}\right]\right\} \tag{6}$$

만약 $\hat{x_0}$와 $\hat{\alpha}$가 비편의량(unbiased estimator)이라고 가정하면 $E(\hat{q}) \approx q$와 같으며, 따라서 $E(\hat{R})$를 Eqs. (4)-(6)을 사용하여 다시 표현하면 다음과 같다.

$$E(\hat{R}) \approx 1-q^n = 1-\left[\exp\left\{-\exp\left(-\frac{x_q-x_0}{\alpha}\right)\right\}\right]^n \tag{7}$$

한편, 위험도 $\hat{R}$ 의 분산량에 대한 1차 근사식은 다음과 같이 정의할 수 있다.

$$\begin{aligned} Var(\hat{R}) \approx & \left(\frac{\partial \hat{R}}{\partial \hat{x_0}}\right)_\mu^2 Var(\hat{x_0}) \\ & + 2\left(\frac{\partial \hat{R}}{\partial \hat{x_0}}\right)_\mu \left(\frac{\partial \hat{R}}{\partial \hat{\alpha}}\right)_\mu Cov(\hat{x_0},\hat{\alpha}) \\ & + \left(\frac{\partial \hat{R}}{\partial \hat{\alpha}}\right)_\mu^2 Var(\hat{\alpha}) \end{aligned} \tag{8}$$

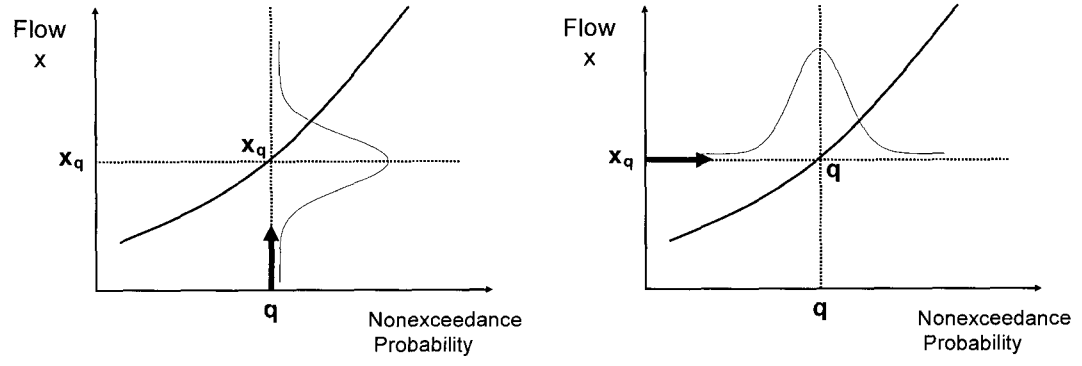

Fig. 1. Schematic Illustrating (a) Flood Quantile Estimation for a Given Nonexceedance Probability (q) and Inverse Problem (b) Estimation of q for a Given Flood

여기서, $(\ )_{\mu}$는 편미분항에 있어 각각 기대값에서의 값을 의미한다.

Gumbel 분포형의 누가분포함수를 이용하여 위치매개변수에 대한 위험도 $\hat{R}$ 의 변화량을 계산하면 다음과 같다.

$$\frac{\partial \hat{R}}{\partial \hat{x}_0} = \frac{n}{\hat{\alpha}} \exp[-\hat{y} - ne^{-\hat{y}}] = \frac{n}{\hat{\alpha}}(-\ln \hat{q})\hat{q}^n \quad (9)$$

여기서, $\hat{y} = \dfrac{x - \hat{x}_0}{\hat{\alpha}} = -\ln[-\ln \hat{q}]$는 Gumbel reduced variate를 의미한다.

동일한 방법으로 규모매개변수에 대한 위험도 $\hat{R}$ 의 변화량을 계산하면 다음과 같다.

$$\frac{\partial \hat{R}}{\partial \hat{\alpha}} = \frac{n}{\hat{\alpha}}(-\ln \hat{q})\ln(-\ln \hat{q})\hat{q}^n \quad (10)$$

Eq. (8)에서의 분산량과 공분산량은 매개변수의 함수이기에 추정방법에 따라 각기 다른 값을 가진다. 본 연구에서는 매개변수 추정방법으로 모멘트법(method of moments, MOM), 최우도법(maximum likelihood method, ML), 확률가중모멘트법(probability weighted moments method, PWM)를 고려하였다.

3.1 모멘트법(Method of Moments, MOM)

MOM을 이용하여 Gumbel 분포형의 매개변수들에 대한 분산량과 공분산량을 산정하면 다음과 같다 (NERC, 1975).

$$Var(\hat{x}_0) = 1.168\frac{\alpha^2}{N} \quad (11)$$

$$Var(\hat{\alpha}) = 1.10\frac{\alpha^2}{N} \quad (12)$$

$$Cov(\hat{x}_0, \hat{\alpha}) = 0.096\frac{\alpha^2}{N} \quad (13)$$

따라서, Eq. (8)에서 정의한 위험도 $\hat{R}$ 의 분산량을 유도하면 다음과 같다.

$$Var(\hat{R})_{MOM} \approx \frac{n^2}{N}\exp[-2y - 2ne^{-y}] \times [1.168 + 0.192y + 1.10y^2] \quad (14)$$

여기서, 위험도 $\hat{R}$ 의 분산량은 설계연한 n과 비초과확률 q(재현기간 T) 및 표본자료 개수 N의 함수로 나타남을 알 수 있다. 다음 Table 1에서는 MOM을 바탕으로 다양한 자료개수N = 10, 50, 100과 설계연한 n = 10, 50, 100, 비초과확률 q = 0.9, 0.98, 0.99에 대한 Eqs. (7) and (14)로부터 산정된 위험도 $\hat{R}$ 의 기대값과 표준편차를 나타내고 있다.

3.2 확률가중모멘트법(Method of Probability Weighted Moments, PWM)

PWM을 통하여 산정된 Gumbel 분포형 매개변수의 분산량 및 공분산량은 다음과 같이 계산된다(Phien, 1987)

$$Var(\hat{x}_0) = \frac{\alpha^2}{N(N-1)}(1.128N - 0.9066) \quad (15)$$

$$Var(\hat{\alpha}) = \frac{\alpha^2}{N(N-1)}(0.8046N - 0.1855) \quad (16)$$

$$Cov(\hat{x}_0, \hat{\alpha}) = -\frac{\alpha^2}{N(N-1)}(0.2287N - 0.5861) \quad (17)$$

Eqs. (9)-(10) 및 Eqs. (15)-(17)를 Eq. (8)에 대입하여 위험도 $\hat{R}$ 의 분산량을 유도하면 다음 Eq. (18)과 같다.

여기서, PWM을 이용하여 유도된 위험도 $\hat{R}$ 의 분산량은 모멘트법과 비교하였을 때 그 식이 약간 복잡하기는 하나, 두 방법 모두 설계연한 n과 비초과확률 q (또는 재현기간 T) 및 표본자료 개수 N 의 함수로 나타남을 알 수 있다.

3.3 최우도법(Method of Maximum Likelihood, ML)

ML을 이용하여 추정된 위치매개변수 x_0 와 규모매개변수 $\hat{\alpha}$ 의 분산량과 공분산량을 산정하면 다음과 같다(NERC, 1975; Phien, 1987).

$$Var(\hat{R})_{PWM} \approx \frac{n^2}{N(N-1)}\exp[-2y - 2ne^{-y}] \times [1.128N - 0.9066 - (0.4574N - 1.1722)y + (0.8046N - 0.1855)y^2] \quad (18)$$

Table 1. Expected Value and Standard Deviation of Risk of Failure

N	q	Method	n=10		n=50		n=100	
			$E(\hat{R})$	$\sigma(\hat{R})$	$E(\hat{R})$	$\sigma(\hat{R})$	$E(\hat{R})$	$\sigma(\hat{R})$
10	0.90	MOM	0.651	0.3111	0.995	0.0230	0.99997	0.00024
		PWM		0.2525		0.0187		0.00019
		ML		0.2686		0.0198		0.00020
	0.98	MOM	0.183	0.2255	0.636	0.5026	0.86738	0.36604
		PWM		0.1880		0.4190		0.30519
		ML		0.1836		0.4091		0.29799
	0.99	MOM	0.096	0.1447	0.395	0.4839	0.63397	0.58547
		PWM		0.1216		0.4067		0.49213
		ML		0.1162		0.3886		0.47021
50	0.90	MOM	0.651	0.1391	0.995	0.0103	0.99997	0.00011
		PWM		0.1074		0.0079		0.00008
		ML		0.1201		0.0089		0.00009
	0.98	MOM	0.183	0.1009	0.636	0.2248	0.86738	0.16370
		PWM		0.0804		0.1791		0.13042
		ML		0.0821		0.1830		0.13326
	0.99	MOM	0.096	0.0647	0.395	0.2164	0.63397	0.26183
		PWM		0.0520		0.1741		0.21062
		ML		0.0519		0.1739		0.21028
100	0.90	MOM	0.651	0.0984	0.995	0.0073	0.99997	0.00007
		PWM		0.0755		0.0056		0.00006
		ML		0.0849		0.0063		0.00006
	0.98	MOM	0.183	0.0713	0.636	0.1589	0.86738	0.11575
		PWM		0.0565		0.1259		0.09172
		ML		0.0581		0.1294		0.09423
	0.99	MOM	0.096	0.0457	0.395	0.1530	0.63397	0.18514
		PWM		0.0366		0.1224		0.14815
		ML		0.0367		0.1229		0.14869

$$Var(\hat{x}_0) = 1.1086\frac{\alpha^2}{N} \tag{19}$$

$$Var(\hat{\alpha}) = 0.6079\frac{\alpha^2}{N} \tag{20}$$

$$Cov(\hat{x}_0, \hat{\alpha}) = 0.2570\frac{\alpha^2}{N} \tag{21}$$

따라서, $\hat{R}$ 의 근사적 분산량은 Eqs. (8)-(10) 및 Eqs. (19)-(21)를 이용하여 유도할 수 있으며 다음과 같다.

$$Var(\hat{R})_{ML} \approx \frac{n^2}{N}\exp[-2y-2ne^{-y}] \times [1.1086+0.5140y+0.6079y^2] \tag{22}$$

4. 비교분석

Eq. (7)에서 나타난 바와 같이 위험도의 기대값 $E(\hat{R})$은 표본자료의 수 N의 함수로는 표현되지 않고 비초과확률 q와 설계연한 n의 함수로 표현되며, 각 매개변수 추정방법에 관계없이 동일한 값을 가짐을 알 수 있다. Table 1에서 나타난 바와 같이 $E(\hat{R})$은 q가 증가할수록 감소하며 n이 증가할수록 그 값이 증가하는 것을 알 수 있다. 또한, 식으로 유도된 위험도의 분산량 $Var(\hat{R})$ 또는 표준편차 $\sigma(\hat{R})$는 세 가지 매개변수 추정방법 모두 n, q, N 의 함수로 나타났으며 이들은 각각 유사한 형태를 가짐을 알 수 있다.

유도된 분산량으로부터 주어진 재현기간에 대하여 최대의 분산량을 발생시키는 설계연한 n을 산정할 수 있는데, Eqs. (14), (18), (22)로부터 $Var(\hat{R})$은 Eq. (23)과 같은 일반적인 형태로 표시할 수 있다.

$$Var(\hat{R}) \propto n^2 (q)^{2n} (-\ln q)^2 C \tag{23}$$

이 식을 이용하여 n에 대한 도함수를 산정한 다음 그 값이 0이 되도록 하는 해를 계산함으로써 최대 분산량을 발생시키는 n을 산정할 수 있다.

$$2nq^{2n}(1+n\ln q)(-\ln q)^2 C = 0 \tag{24}$$

따라서, Eq. (24)를 만족하는 n은 다음 Eq. (25)와 같으며, Table 1에서 제시한 각 비초과확률에 대하여 최대분산량을 발생시키는 설계연한을 산정한 결과 매개변수 추정방법들과 표본자료 수에는 무관하게 q= 0.90인 경우 $n \approx 10$, q=0.98인 경우 $n \approx 50$, q=0.99인 경우 $n \approx 100$로 다음의 관계식을 만족함을 알 수 있다.

$$n = -1/\ln q \tag{25}$$

Table 1에서 나타난 바와 같이 $\sigma(\hat{R})$은 T = 10년인 경우 n이 증가함에 따라 감소하나 T = 100년인 경우에는 n의 증가에 따라 그 값이 증가하게 된다. T= 50년인 경우, n이 50년 이하에서 50년까지 그 값이 증가하게 될 때에는 $\sigma(\hat{R})$은 값이 증가하게 되며 50년보다 큰 범위에서 n값이 증가하게 될 때에는 $\sigma(\hat{R})$이 감소함을 알 수 있다.

유도된 식을 이용하여 $\sigma(\hat{R})_{MOM} / \sigma(\hat{R})_{ML}$의 비를 고려하여 보면 단순히 q 의 함수로만 이루어져 있음을 알 수 있는데, 이 값은 q 값이 증가함에 따라 증가하며 q 값이 일정한 경우에는 Table 2에서 나타난 바와 같이 항상 같은 값을 가짐을 알 수 있다. 반면, $\sigma(\hat{R})_{PWM} / \sigma(\hat{R})_{ML}$의 비는 q와 N의 함수로 구성됨을 알 수 있는데 Table 2에 나타난 바와 같이 q 가 일정한 경우 N이 증가함에 따라 이 값은 감소하게 되며 N 이 일정한 경우에는 q 가 증가함에 따라 그 값이 증가하는 것을 알 수 있다.

위험도의 분산량의 크기에 대하여 각 매개변수 추정방법별로 그 크기를 비교하면 다음과 같은데, 먼저 $Var(\hat{R})_{MOM}$ 과 $Var(\hat{R})_{PWM}$의 차는 다음 Eq. (26)과 같이 표현할 수 있다.

$$DIFE^{\hat{R}}_{MOM-PWM} = \frac{n^2}{N}(q)^{2n}(-\ln q)^2(a_1 y^2 + b_1 y + c_1) \tag{26}$$

여기서,

$$a_1 = 1.10 - \frac{1}{N-1}(0.8046N - 0.1855) \tag{27}$$

Table 2. Ratio of $\sigma(\hat{R})_{MOM} / \sigma(\hat{R})_{ML}$ 및 $\sigma(\hat{R})_{PWM} / \sigma(\hat{R})_{ML}$ as a Function of Sample Size and Nonexceedance Probability

N	q	Ratios	
10	0.9	$\sigma(\hat{R})_{MOM} / \sigma(\hat{R})_{ML}$	1.1584
		$\sigma(\hat{R})_{PWM} / \sigma(\hat{R})_{ML}$	0.9403
	0.98	$\sigma(\hat{R})_{MOM} / \sigma(\hat{R})_{ML}$	1.2283
		$\sigma(\hat{R})_{PWM} / \sigma(\hat{R})_{ML}$	1.0242
	0.99	$\sigma(\hat{R})_{MOM} / \sigma(\hat{R})_{ML}$	1.2451
		$\sigma(\hat{R})_{PWM} / \sigma(\hat{R})_{ML}$	1.0466
50	0.9	$\sigma(\hat{R})_{MOM} / \sigma(\hat{R})_{ML}$	1.1584
		$\sigma(\hat{R})_{PWM} / \sigma(\hat{R})_{ML}$	0.8944
	0.98	$\sigma(\hat{R})_{MOM} / \sigma(\hat{R})_{ML}$	1.2283
		$\sigma(\hat{R})_{PWM} / \sigma(\hat{R})_{ML}$	0.9787
	0.99	$\sigma(\hat{R})_{MOM} / \sigma(\hat{R})_{ML}$	1.2451
		$\sigma(\hat{R})_{PWM} / \sigma(\hat{R})_{ML}$	1.0016
100	0.9	$\sigma(\hat{R})_{MOM} / \sigma(\hat{R})_{ML}$	1.1584
		$\sigma(\hat{R})_{PWM} / \sigma(\hat{R})_{ML}$	0.8890
	0.98	$\sigma(\hat{R})_{MOM} / \sigma(\hat{R})_{ML}$	1.2283
		$\sigma(\hat{R})_{PWM} / \sigma(\hat{R})_{ML}$	0.9733
	0.99	$\sigma(\hat{R})_{MOM} / \sigma(\hat{R})_{ML}$	1.2451
		$\sigma(\hat{R})_{PWM} / \sigma(\hat{R})_{ML}$	0.9964

Table 3. Nonexceedance Probability for Comparison of Each Estimation Method

Sample Size	Solutions	MOM vs. PWM		ML vs. PWM	
		y	q	y	q
10	1	-0.027	0.3580	0.050	0.386
	2	-2.488	0.00001	3.310	0.964
25	1	-0.051	0.3490	0.031	0.379
	2	-2.247	0.00008	4.201	0.985
50	1	-0.057	0.3470	0.025	0.377
	2	-2.188	0.00013	4.546	0.989
100	1	-0.060	0.3460	0.023	0.376
	2	-2.160	0.00017	4.728	0.991
200	1	-0.062	0.3450	0.021	0.376
	2	-2.148	0.00019	4.822	0.992

$$b_1 = 0.192 + \frac{1}{N-1}(0.4574N - 1.1722) \quad (28)$$

$$c_1 = 1.168 - \frac{1}{N-1}(1.128N - 0.9066) \quad (29)$$

따라서 a_1, b_1, c_1는 N의 함수로 표현됨을 알 수 있고, 4개의 항으로 이루어진 Eq. (26)에서 세 번째까지의 항은 항상 양의 값을 가지게 되므로 $Var(\hat{R})_{MOM}$ 과 $Var(\hat{R})_{PWM}$ 은 Eq. (26)의 마지막 항에 따라 크고 작음이 결정되며, 다시 말해 표본자료 수에 따라 그 크고 작음이 결정된다고 할 수 있다. 그러므로 계수 a_1이 양의 값을 가지고($N > 3$), y값이 2차식의 두 해 사이에 존재한다면 $Var(\hat{R})_{MOM}$ 은 항상 $Var(\hat{R})_{PWM}$ 보다 작은 값을 가지게 된다.

Table 3에서는 여러 표본자료 수에 대하여 Eq. (26)에서의 y의 해 및 이에 해당하는 비초과확률을 나타내고 있는데, 여기서 $Var(\hat{R})_{MOM} < Var(\hat{R})_{PWM}$을 만족시키는 재현기간은 2년보다 작음을 알 수 있으며, 일반적으로 홍수빈도해석에서는 2년 이상의 재현기간을 고려하는 점에서 $Var(\hat{R})_{PWM}$이 $Var(\hat{R})_{MOM}$보다 항상 작은 값을 가짐을 알 수 있다. 이와 유사한 방법으로 $Var(\hat{R})_{ML}$와 $Var(\hat{R})_{PWM}$의 차를 식으로 나타내면 다음과 같다.

$$DIFF^{\hat{R}}_{ML-PWM} = \frac{n^2}{N}(q)^{2n}(-\ln q)^2(a_2 y^2 + b_2 y + c_2) \quad (30)$$

여기서,

$$a_2 = 0.6079 - \frac{1}{N-1}(0.8046N - 0.1855) \quad (31)$$

$$b_2 = 0.5140 + \frac{1}{N-1}(0.4574N - 1.1722) \quad (32)$$

$$c_2 = 1.1086 - \frac{1}{N-1}(1.128N - 0.9066) \quad (33)$$

여기서, 계수 a_2는 항상 음의 값을 가지기 때문에 y가 Eq. (30)의 마지막 항으로 표현된 2차방정식의 2개의 해 사이에 존재하는 경우에는 $Var(\hat{R})_{PWM}$는 항상 $Var(\hat{R})_{ML}$보다 작은 값으로 가지게 된다. 이 경우 Table 3의 5, 6열에서 볼 수 있는 바와 같이 표본자료 수에 따라 T=2.0~125년의 범위에서 $Var(\hat{R})_{PWM}$은 $Var(\hat{R})_{ML}$보다 작은 값을 가지게 된다. 또한 표본자료 수가 증가함에 따라 PWM이 ML에 비해 적은 분산량을 가지게 되는 재현기간의 범위는 점점 증가하게 됨을 알 수 있다.

$Var(\hat{R})_{ML}$과 $Var(\hat{R})_{MOM}$의 차를 식으로 나타내면 다음과 같다.

$$DIFF^{\hat{R}}_{ML-MOM} = \frac{n^2}{N}(q)^{2n}(-\ln q)^2(a_3 y^2 + b_3 y + c_3) \quad (34)$$

여기서, $a_3 = -0.4921$, $b_3 = 0.3220$, $c_3 = -0.0594$로서 일정한 값을 가진다. 여기서 계수 a_3 는 항상 음의 값을 가지며 Eq. (34)의 마지막 항으로 표현된 2차식의 해를 산정할 수 없기 때문에 마지막 항은 항상 음의 값

을 가지게 된다. 따라서 $Var(\hat{R})_{MOM}$은 재현기간에 무관하게 항상 $Var(\hat{R})_{ML}$보다 큰 값을 가지게 된다. 한편, 추정된 초과확률 및 재현기간에 대한 매개변수 추정방법별 차를 식으로 유도하여 나타내면 다음과 같다.

$$DIFF^{\hat{P}}_{MOM-PWM} = \frac{1}{N}(q)^2(-\ln q)^2(a_1 y^2 + b_1 y + c_1) \quad (35)$$

$$DIFF^{\hat{P}}_{ML-PWM} = \frac{1}{N}(q)^2(-\ln q)^2(a_2 y^2 + b_2 y + c_2) \quad (36)$$

$$DIFF^{\hat{P}}_{ML-MOM} = \frac{1}{N}(q)^2(-\ln q)^2(a_3 y^2 + b_3 y + c_3) \quad (37)$$

$$DIFF^{\hat{T}}_{MOM-PWM} = \frac{1}{N}(q)^2(1-q)^{-4}(-\ln q)^2(a_1 y^2 + b_1 y + c_1) \quad (38)$$

$$DIFF^{\hat{T}}_{ML-PWM} = \frac{1}{N}(q)^2(1-q)^{-4}(-\ln q)^2(a_2 y^2 + b_2 y + c_2) \quad (39)$$

$$DIFF^{\hat{T}}_{ML-MOM} = \frac{1}{N}(q)^2(1-q)^{-4}(-\ln q)^2(a_3 y^2 + b_3 y + c_3) \quad (40)$$

여기서, 계수 $a_1, \cdots, c_3$는 Eqs. (26)-(34)에서와 동일한 값을 가지며, Eqs. (26)와 (35), Eqs. (30)과 (36), Eqs. (34)과 (37)의 차이점은 2차식으로 구성된 마지막 항 이외의 부분이므로 $\hat{R}$의 분산량에 대해 수행한 방법별 비교에서와 동일한 결과를 기대할 수 있는데 즉, 재현기간이 2년 이상인 경우 $Var(\hat{p})_{PWM}$은 항상 $Var(\hat{p})_{MOM}$보다 작은 값을 가지며, $Var(\hat{p})_{PWM}$은 Table 3에서 제시한 구간에 대해 $Var(\hat{p})_{ML}$보다 작은 값을 가지고, $Var(\hat{p})_{ML}$은 $Var(\hat{p})_{MOM}$보다 항상 작은 값을 가지게 된다. 재현기간의 경우에도 역시 동일한 결과를 기대할 수 있다.

5. Monte-Carlo 모의

제한된 표본자료 수에 대해 각 매개변수 추정방법별로 이론적으로 유도한 위험도 분산량의 적용가능성을 검토하기 위하여 Monte-Carlo 모의실험을 수행하였다. Gumbel 분포형에 대한 모집단 위치매개변수 x_0는 1000으로 가정하였으며 모집단의 변동계수 C_v= 0.5, 1.0, 2.0인 각각의 경우를 구분하여 이를 만족하도록 규모매개변수 $\hat{\alpha}$= 503, 1006, 2012로 가정하였고, 표본집단 수 N=10, 25, 50, 100, 200, 설계연한 n= 10, 25, 50, 100, 200, 500, 비초과확률 q= 0.90, 0.96, 0.98, 0.99, 0.995, 0.998인 경우에 대하여 각각 모의를 실시하였다. 가정된 Gumbel 분포형의 매개변수를 바탕으로 10,000개의 자료를 발생시킨 다음 MOM, PWM, ML을 이용하여 각각 유도된 위험도 분산량의 적용가능성을 검토하기 위하여 상대편의(relative bias, RBIAS) 및 상대제곱근오차(relative root mean square error, RRMSE)를 산정하여 비교, 분석하였다.

Simulation을 통한 각 매개변수 추정방법별 비교분석을 실시한 결과는 다음과 같다. 먼저 표본자료 수 N이 작은 경우를 고려하여 보면 PWM은 비교적 작은 q에서 가장 작은 RBIAS를 보이는 것으로 나타났으며 q가 큰 경우에는 ML이 가장 작은 RBIAS 값을 보이는 것으로 나타났다. 하지만, 설계연한 n을 증가시키는 경우에는 고려된 모든 q 범위에서 PWM은 가장 작은 RBIAS를 제공하며, 표본자료 수 N을 증가시키는 경우에도 작은 설계연한 n에 대해서 PWM이 가장 작은 RBIAS를 제공하는 것을 알 수 있다. 반면 설계연한 n을 증가시키는 경우 큰 값의 q에 대하여 PWM은 가장 작은 RBIAS를 가지는 반면 상대적으로 작은 q에 대해서는 ML이 가장 작은 RBIAS 값을 가짐을 알 수 있다.

Fig. 2에서는 n=50, q=0.99인 경우 모의실험을 통하여 계산된 표본자료 수 N에 따른 각 방법별 위험도 $\hat{R}$의 RBIAS를 비교, 도시하고 있는데, 이 경우 PWM이 가장 작은 RBIAS를 제공하는 것을 알 수 있으며, 특히

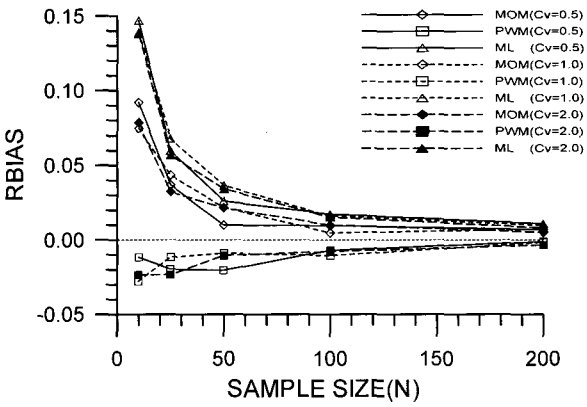

Fig. 2. RBIAS vs. Sample Size of Risk of Failure for MOM, PWM, and ML (n=50, q=0.99, C_v=0.5, 1.0, 2.0)

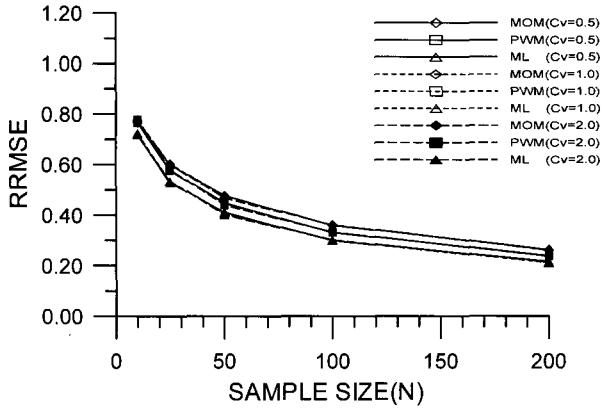

Fig. 3. RRMSE vs. Sample Size of Risk of Failure for MOM, PWM, and ML (n=50, q=0.99, C_v=0.5, 1.0, 2.0)

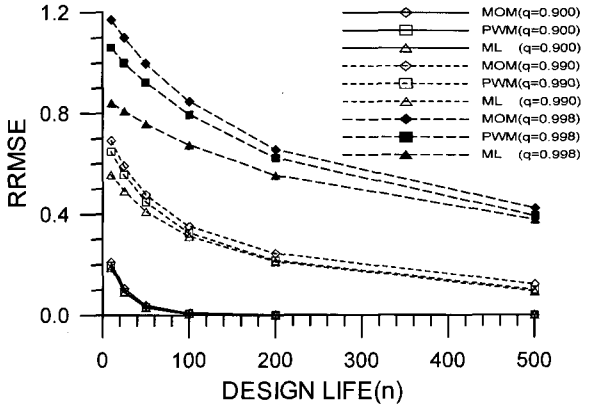

Fig. 4. RRMSE vs. Design Life of Risk of Failure for MOM, PWM, and ML (N=50, q=0.90, 0.99, 0.998, C_v=0.5)

표본자료 수 N이 작을수록 다른 방법에 비해 더 큰 차이를 보이며 우수한 결과를 제공하는 것으로 나타났다. 일반적으로 PWM은 표본자료 수가 적은 경우 가장 우수하며 자료수가 큰 경우에는 ML이 우수한 것으로 알려져 있으며, RBIAS는 자료수가 증가함에 따라 감소한다. 한편, 변동계수의 변화가 RBIAS에 미치는 영향을 검토한 결과, Fig. 2에서 도시된 바와 같이 RBIAS값에 큰 변화를 주지 않는 것으로 나타났으며 변동계수 조건은 RBIAS값에 별다른 영향을 주지 않는 것으로 나타났다.

크기가 작은 표본자료(N =10)와 설계연한 n의 경우에 대해 RRMSE를 고려한 결과, PWM은 상대적으로 비교적 작은 재현기간에서 가장 작은 값을 보이며, 재현기간이 증가할수록 ML이 가장 작은 RRMSE을 가지는 것으로 나타났다. 또한, 설계연한 n값을 증가시킴에 따라 PWM은 점차 넓은 q의 범위에서 가장 작은 RRMSE를 보이는 것으로 나타났으며, n=500인 경우에는 고려한 모든 q에서 가장 작은 RRMSE을 보임을 알 수 있다. 반면, $N \geq 50$인 경우에는 고려한 모든 q와 n에서 ML이 가장 작은 RRMSE을 가지는 것으로 나타났다.

Fig. 3에서는 표본자료 수 N 이 RRMSE에 미치는 영향에 대하여 n=50, q=0.99인 경우를 도시하고 있다. 그림에서와 같이 ML을 이용한 경우가 전체적으로 가장 작은 값을 보이며, 세 가지 방법에서 계산된 RRMSE 모두 N 이 증가할수록 그 값이 감소하여 점차 동일한 값으로 수렴하는 것을 알 수 있다. 또한 변동계수의 변화가 RRMSE에 미치는 영향은 무시할 수 있음을 알 수 있다.

설계연한이 RRMSE에 미치는 영향을 도시하면 Fig. 4와 같다(N=50, C_v=0.5). q= 0.90 인 경우 각 방법별로 계산된 RRMSE 값은 거의 동일하며, n이 증가할수록 ($n \geq 100$) 그 값은 점차 0으로 수렴함을 알 수 있다. 또한 재현기간이 작을수록 계산된 RRMSE이 작음을 알 수 있으며, 모든 q 에서 n이 증가함에 따라 RRMSE는 감소함을 알 수 있는데 이는 표본자료 수가 다른 경우에도 동일한 경향을 보인다. RRMSE는 N =10인 경우를 제외하고는 q 가 증가할수록 그 값이 증가하는 것으로 나타났으며, N 과 RRMSE의 관계에서 분석한 바와 같이 변동계수가 RRMSE에 미치는 영향을 무시할 수 있는 것으로 나타났다.

전체적으로 모의를 통하여 계산된 RBIAS에 대하여 매개변수 추정방법 별로 비교분석한 결과, PWM은 표본자료 수가 작은 경우 가장 작은 값을 보이는 것으로 나타났으며, 반면 ML은 설계연한과는 관계없이 비교적 큰 표본자료 수에서 가장 작은 RBIAS를 발생시키는 것으로 나타났다. 또한 RRMSE의 측면에서 PWM은 작은 표본자료 수와 큰 설계연한에서 가장 우수한 것으로 나타났으며, ML은 큰 표본자료 수에서 우수한 것으로 나타났다.

6. 결 론

본 연구에서는 수공구조물의 위험도의 불확실성을 정량적으로 규명하기 위하여 Gumbel 분포형을 기본 분포형으로 가정한 다음 MOM, PWM, ML의 매개변수 추정방법별로 수공구조물의 위험도의 기대값 및 분산량에 대한 이론적인 식을 유도하였다. 그 결과 위험도의 분산량은 수공구조물의 설계연한, 설계홍수량의 재현기간(비초과확률) 및 사용된 수문자료의 크기의 식으로 표현됨을 알 수 있었다. 유도된 분산량에 대한 이론적 비교 결과 $Var(\hat{R})_{PWM}$은 재현기간이 2년 보다 큰 경우 $Var(\hat{R})_{MOM}$보다 작은 값을, 그리고 표본자료 수에 따라

재현기간 2-125년 범위에서는 $Var(R)_{ML}$보다 작은 것으로 나타났으며, $Var(\hat{R})_{ML}$은 재현기간에 상관없이 $Var(\hat{R})_{MOM}$보다 항상 작은 것으로 나타났다. 유도된 분산량의 적용가능성을 검토하기 위하여 다양한 크기의 표본자료 및 조건들에 대하여 모의실험을 실시한 결과, 자료수가 작은 경우 PWM을 이용하였을 때 가장 작은 상대편의를 보이는 것으로 나타났으며, 비교적 자료수가 많은 경우($N \geq 25$)에는 ML이 가장 작은 상대편의를 나타내었다. 표본이 작은 경우(N=10)에는 설계연한이 증가함에 따라 PWM의 경우 가장 작은 상대제곱근오차를 발생시키는 것으로 나타났으나, 상대적으로 큰 자료에 대해서는 설계연한 및 비초과확률과는 관계없이 ML이 가장 작은 상대제곱근오차를 발생시키는 것으로 나타났다. 또한 변동계수의 변화는 위험도의 분산량에 대하여 별다른 영향을 미치지 않는 것으로 나타났다.

감사의 글

본 연구는 연세대학교 교내연구과제의 학술지원에 의해 수행되었으며, 이에 감사를 드리는 바입니다.

참 고 문 헌

Benjamin, J.R. and Cornell, C.A. (1970). *Probability, statistics and decision for civil engineers*, McGraw Hill Book Co., N.Y.

Chawdhurym J.U. and Stedinger, J.R. (1991). "Confidence intervals for design floods with estimated skew coefficient." *Journal of Hydraulic Engineering*, ASCE, Vol. 117, No. HY1, pp. 811-831.

Chow, V.T., Maidment, D.R., and Mays, L.W. (1988). *Applied hydrology*, McGraw Hill Book, Co., N.Y.

Davis, D., Duckstein, L., Kisiel, C., and Fogel, M. (1972). "Uncertainty in the return period of maximum events: A Bayesian approach." *Proc. Inter. Symp. on Uncertainties in Hydrologic and Water Resources Systems*, Tucson, Arizona, pp. 853-862.

Kite, G.W. (1988). *Frequency and risk analysis in water resources*, Water Resources Publications, Littleton, Colorado.

Loaiciga, H.A. and Mariño, M.A. (1991). "Recurrences interval of geophysical events." *Journal of Water Resources Planning and Management*, ASCE, Vol. 117, No. 3, pp. 367-382.

NERC (National Environment Research Council) (1975). *Flood studies report, Vol. I, Hydrological studies*, Whiterfriars Press Ltd., London.

Phien, H.N. (1987). "A review of methods for parameter estimation for the extreme value type-1 distribution." *Journal of Hydrology*, Vol. 90, pp. 251-268.

Rasmussen, P.F. and Rosbjerg, D. (1989). "Risk estimation in partial duration series." *Water Resources Research*, Vol. 25, No. 11, pp. 2319-2330.

Rasmussen, P.F. and Rosbjerg, D. (1991). "Prediction uncertainty in seasonal partial duration series." *Water Resources Research*, Vol. 27, No. 11, pp. 2875-2883.

Rosbjerg, D. and Madsen, H. (1992). "Prediction in partial duration series with generalized Pareto-distributed exceedances." *Water Resources Research*, Vol. 28, No. 11, pp. 3001-3010.

Stedinger, J. R., Vogel, R.M., and Foufoula-Georgiou, E. (1993). "Frequency analysis of extreme events." *Handbook of hydrology*, Edited by Maidment, D.R., McGraw Hill Book Co., New York.

Thompson, K.D., Stedinger, J.R., and Heath, D.C. (1997). "Evaluation and presentation of dam failure and flood risks." *Journal of Water Resources Planning and Management*, ASCE, Vol. 123, No. 4, pp. 216-227.

Vogel, R.M. (1987). "Reliability indices for water supply systems." *Journal of Water Resources Planning and Management*, ASCE, Vol. 113, No. 4, pp. 563-579.

Watt, W.E. and Wilson, K.C. (1978). "An approach to optimal design of hydraulic structures." *Reliability in water resources management*, Edited by McBean, E.A. et al., Water Resources Publications, Littleton, Colorado, pp. 75-90.

Yen, B.C. (1970). "Risks in hydrologic design of engineering projects." *Journal of Hydraulic Engineering*, ASCE, Vol. 96, No. HY4, pp. 959-966.

(논문번호:06-34/접수:2006.02.27/심사완료:2006.07.14)