〈講 座〉

수문학에서의 Robust Estimators [I]

조 원 철*

1. 서 론

수문학은 크게 두개의 별개 분야 즉 과학분야 및 공학분야에서 공헌하고 있다. 이러한 과학분야 및 공학분야의 구분은 모형기법에 의해서 더욱 분명하게 된다. 과학분야는 주로 인과관계(causality)에 관심을 가지며 자현현상에 대한 복잡하고도 정교한 모형을 제안하고 있는 반면에 공학분야는 물관리에 관심을 가지며 관리분석및 결정이론을 포함하는 모형에 밀접한 관계가 있다. 이 두 그룹이 보다 전문화됨에 따라 상호작용 및 공동연구는 점차 사라지고 있다. robust 통계학은 이 두 분야사이의 많은 경계선을 연결하는데 도움을 주기 때문에 그 중요성이 날로 인식되어 가고 있다.

robustness를 정의하기는 매우 어려워서 Fiering[1]은 12개의 정의를 제공한 바 있다. 자료값 사이에서 발생하는 적정수준의 교란에 대한 매개변수의 안정도 개념은 새로운 조건 또는 관리지침하에서의 체제변화(system change)를 예측하는데 필요한 조건이다.

robustness는 매개변수들의 급격한 변화나 system의 실행정밀도 보다는 측정되어지는 조절된 동적응답(controlled dynamic response)의 특징에 더 많은 관련을 가진다. 통계학 및 경제학이 의사결정의 공식적 및 비공식적인 수단인 반면 수문과정모형은 여러가지 선택방안들 중에서 하나의 방안을 선택하는데 도움을 제공해 준다. 이러한 의사결정에 필요한 기초는 모형으로부터 구해지는 예측에 기초하고 있다. 우리는 모형을 관측값에 적용하여 보고 그 모형이 충분히 잘 들어 맞는 지를 검토한다. 그리고 또한 적용유역의 수문현상을 잘 설명하고 있는 지도 검토해야 한다. 그 후에라야 우리는 새로운 적용유역에 대해서 예측을 할 수 있는 것이다.

종래의 수자원 해석은 유역의 가용수의 시간적 추적, 저류시설, 조절설비, 날로 증가하는 물 수요의 사회적 요구, 그리고 제반시설의 설계 및 관리에 필요한 계산등에 촛점을 맞추어 왔다. 물의 균형(water balance)이 모든 분석을 지배하고 있다. 즉 수요와 공급을 균형있게 하기 위한 시도들은 자급자족하는 경향을 보이며 모든 가용수자원을 이용한다. 물의 수요와 공급을 다루는 장기계획은 경제성장과도 절대적인 관계를 가지고 있다. 그러한 장기적인 계획은 중요한 변화를 예측하여야 하며 중요한 지역적 추세를 규명하여야 하고 현재의 혹은 잠재적인 관련지역을 예측하여야만 한다. 그렇게 하기 위하여 수문과정 모형을 이용하는 데는 다음의 4가지의 기본적인 사항을 필요로 한다.

(1) 수문과정을 모형화시키며 이 모형의 매개변수들을 결정할 수 있는 실질적인 능력의 향상.

(2) 개선된 수학적 기법, 계산기법 및 통계적 기법.

(3) 자료수집 및 보관처리 체계의 개선.

(4) 본고에서 제안하고 있는 탄력성(resilience)

* [필자주 : 본 강좌는 수문학 및 수자원공학에서의 Risk와 불확실성문제해결을 위한 새로운 해석기법으로 관심을 모으고 있는 Robust 통계학을 소개하기 위한 것으로 개념의 난해함때문에 표현이 무척이나 힘들었음을 알려드리며 이 분야의 연구자들인 Harvard대학의 Myron B. Fiering, George Kuczera, 그리고 MIT의 John C. Houghton의 글을 중심으로 엮은 것이다.]

* 연세대학교 공과대학 토목공학과

및 robustness개념에 근거한 의사결정 방법이다.

2. Robustness 및 탄력성

설계의 robustness개념은 다른 과학적, 사회적 분야에서와 같은 유사점을 갖는다[2]. 확률현상에 근거한 종래의 설계방법은 해로운 결과의 발생확률을 감소시키고자 하였다. 역사적으로 기술자들은 홍수조절을 위하여 댐 및 제방을 건설하였으나 최근에 들어서야 비로소 기상조절, 조기경보체계, 체계적 홍수방지대책, 지역화 및 홍수보험등에 의하여 자유체(free-body) 즉 system의 구조적 성질과 경제성을 변화시키고 있다. 이러한 대책들은 받아들일 수 없는 사상의 확률을 조절하기보다는 받아들일 수 있는 범위를 확장하는 것이다.

댐 및 제방들은 수문학적 위협을 감소시켜 준다. 그러면서도 현대사회는 대단위의 재해가 가능한 방향으로 발달하고 있으며 이러한 점은 분명히 부적절한 것이라 하겠다. 이런 형태의 행위는 위험에 노출되어 있는 공통적인 상황이다. 예를들자면 우리가 보호벽이 설치된 도로위를 빠른 속도로 달리고 있다고 하자. 이 보호벽은 "system"을 안전하게 해준다. 즉 우리는 보호벽 덕분으로 절벽으로 추락한 위험성을 덜게 된다. 그러나 이러한 안전성이 맞이하게 되는 위험하에서 생존할 수 있는 system의 확실한 능력을 말하는 것은 아니다. 따라서 보호벽이 설치된 도로에서의 사고는 보호벽이 설치 안된 도로에서의 사고보다 더 위험하다. 왜냐하면 보호벽이 있을 때 보다 빠른 속력으로 달리기 때문이다. 안전성이 반드시 시스템 수행의 바람직한 특성인 것만은 아니다. System은 반드시 "safe-fail"이 되도록 설계될 수는 없다. 적어도 합리적인 시간과 합리적인 비용을 넘지 않아야 한다. 즉 System설계는 보수 또는 복구에 드는 시간과 비용을 고려해야만 한다. 홍수조절용 시설물은 system을 보다 실패하기 쉽게 만들며 재난에 상처 입기 쉽게 하는 경향이 있다. 생태학적인 표현을 사용하면 이것은 용납할 수 있는 행동의 범위를 더 줄이는 것과 같은 것이다. 즉 기술자들이 수많은 영향들을 충분히 고려하지 않고 생태계에 대한 "엄한 제한"을 마련하여 놓는 것과 비슷하다.

기상조절로 발생할 홍수위험의 빈도 및 그 성격을 바꾸려는 시도를 할 수 있다. 이러한 시도에는 매우 중요한 불확실성이 있기 때문에 우리는 불가피한 실패 또는 홍수로부터 보호하기 위해 "safe-fail"의 상반된 대안들을 발생하여 우리의 위치에 대한 안전성을 검토하는 것이다. 생태학적인 표현을 빌리자면 "실패"란 생태계가 받아들일 수 없는 상태로 되어 있는 것을 의미한다. 이와같이 받아들일 수 있는 생태계의 한계를 피함으로써 실패를 막으려 노력하고 한편으로는 system의 동성(dynamics)을 변경하므로써 이러한 경계를 탄력이 있으면서도 유연하게 유지하려고 한다. 따라서 받아들일 수 없는 위치로부터 받아들일 수 있는 것을 분리하는 경계를 피함으로써 실패를 피하려는 시도를 하며 그와 동시에 현재의 그리고 예측된 위치로부터 경계를 제거하여 실패를 없애려 한다.

기술자들은 오랫동안 "safe-fail"한 환경을 만듦으로써 공해를 없애려고 노력해 왔다. 대규모 하수처리 시설 및 이중벽으로 건조한 유조선등이 그 한 예라할 수 있다. 불가피한 실패를 회복하려는 것이 탄력성 있는 계획설계의 가장 중요한 요소중의 하나임을 명백히 인식해야 한다. 전반적인 최적조건은 쉽게 구해질 수 있지만 그것의 유용성은 몇가지 매개변수 추정값은 정확도에 달려 있다. 최적해가 그 방법을 제공해줄 수 있는지 혹은 최적값에 근사한 해가 보다 나은 기회를 제게해 줄 수 있는 지를 규명해야만 한다. 최적성(optimality)을 너무 강조하지 않아야 하며 그래야만 근시안적인 최적화를 극복할 수 있는 것이다.

탄력성은 시간의 차원을 갖는다. 결정변수 x_i에 대한 system반응의 민감도는 도함수 $\partial f/dx_i$로 주어지는데 이것은 동적거동을 반영하지 않고 있다[3]. 전미분 df/dx_i가 dx_j/dx_i의 항을 포함하도록 함으로서 만일 x_i에 작은 변화량 dx_i가 발생하면 system반응을 최적화하기 위하여 또 다른 가능한 결정변수 x_j를 다룰 수 있다. 이처럼 결정변수 몇몇에 대한 조정은 다른 것에

비하여 보다 신속히 이루어질 수 있다. 저수지에서의 수문의 개폐 또는 발전소에서의 연료스위치 작동은 빠르게 이루어질 수 있다. 기준 및 규정을 바꾸는 것과 같은 정치적 및 제도적 조정은 수개월이 요구될지도 모른다. 구조물의 구조변화는 수년이 걸릴지도 모르며 토지이용의 변경, 인구분포형태의 변경등과 같은 사회적 변화는 수십년이 걸릴수도 있다. 따라서 이미 구해진 해(solution)의 vecter$\{x\}$로부터 예정한 대로 약간 옮겨간 위치에 대한 새로운 전미분 df/dx_i를 최소로 한다거나 혹은 dx_i의 변화에 따른 제약조건에서 df를 최소로 한다는 것은(즉 구하고자 하는 새운로 해는 원래의 해 x를 중심으로 하는 적절한 범위내에 있다.) 주어진 상태 x로부터 축소된 반응수준(response level)에 대한 탄력성있는 system의 궤적을 나타내는 것이다.

수많은 시스템분석 기법의 응용에서 정해진 범위(환경기준)안에서 system 수행의 실패를 나타내는 벌칙함수(penalty function)개념을 이용하는 것이 관례로 되어 왔으며 이렇게 정해진 범위를 표준조건으로 여겨 왔다. 벌칙함수는 반드시 대칭이 아니라고 하더라도 표준편차에 대하여 증가하는 것이 보통이다.

system 실행의 정확도는 초기위치 및 최종위치사이, 혹은 하단과 상한사이에서의 궤적(trajectory)을 따라서 발생하는 기울기와 관련있다. 만일 그 궤적이 완만하다면 앞으로의 정책변화를 조정할 충분한 시간이 있어서 피해를 줄일 수 있다.

3. 극치확률

수문학자들이 크게 관심을 보이고 있는 확률밀도함수(pdf)는 강우 및 유출이 밀도함수의 중심 가까이에 있을 때에 매우 유용하게 인다. 그러나 가뭄 및 홍수와 같은 극한 현상이 발생했을 때는 문제가 발생한다. 불행하게도 우리의 예측능력은 이러한 현상에 대해서는 정확하지 못하다. 장기간에 걸친 기록들이 그리 많지 않기 때문에 극한상황을 나타내는 분포의 끝단 즉 tail을 정의하는데 필요한 고차 moment의 추정값은 극히 불안정하다. 예를들면 N년의 홍수기록 가운데 가장 작은 홍수값을 key punch하는 과정에서 발생하는 transcription error는 장기재현기간 T년에 대한 홍수추정값에 큰 영향을 줄 수도 있으며 moment를 이용해서 밀도함수를 정의한 경우에는 이러한 현상은 쉽사리 발생한다.

누가함수(cdf)에 보다 큰 관심을 갖고 있을 때는 적합도 판정을 재래의 편차값에 의하여 판정해야 하는가 하는 의문점들이 설계계산에서는 문제시 되지 않아 왔다. 그것들은 평균값 및 분포함수의 중심 가까이의 중앙값(median)에 근거하고 있어서 수문해석에 있어서 조그만 오차에는 민감하지 않았다. 그러나 최근에 이르러 보다 여유있는 수자원 계획들과 수질보호를 중요목적으로 하는 계획들, 그리고 인구의 증가에 직면해서 관개용수를 확보하지 못하는 것과 같은 심각한 문제점들의 해결을 위해서 극한 수문현상에 대한 정확한 평가가 수문사업의 계획, 설계 및 운영에 보다 중요한 문제점으로 인식하게 되었다. 매개변수 하나를 잘 추정한다고 해서 주요 사상에 대한 안정된 주기성 추정을 하기란 어렵다.

우리의 목적은 특정사상 x에 대한 pdf와 cdf를 추정하는데 있어서 가장 신뢰할만한 분포함수를 추정하는데 있다. 여기서 cdf는 다음과 같은 적분으로 구해진다.

$$\int_{-\infty}^{x} f(x)dx = F(x)$$

만일 매개변수의 작은 교란이 pdf 또는 cdf에 커다란 분산을 야기한다면 매개 변수들에 대한 robust 추정기법을 개발한다는 것은 쓸모가 없게 된다. 확률 및 누가 밀도함수의 추정값을 가지고 무엇을 할 것인가에 대하여 생각하는 것이 더욱 중요하다. 그리고 매개변수 추정값의 안정성에 대하여 지나치게 걱정하는 것보다는 재현기간 및 기대손실을 가지고 무엇을 할 것인가를 생각하는 것이 더 중요하다. 매개변수는 고유값을 갖고 있지 않으며 정확한 추정값이라 할지라도 "보다 나은" 모형을 만들기 위한 과학적 목적에도 불구하고 아무런 유용한 실용적 결과를 제시하지 못한다. 많은 종류의 밀도함수는

지수항을 포함하고 있어서 매개변수 추정값은 작은 오차라도 tail에서는 큰 변동을 야기시킨다. 따라서 매개변수 자체의 robustness 혹은 안정성은 유용성에 대한 본질적인 문제가 아니다.

다시 말하면 매개변수값들이 변화하더라도 재현확률이 안정하고도 확실한 추정값들을 주는 밀도함수를 이용하는 것이 중요하다. 극치 수문량을 분석하는 데는 moments 보다는 매개변수들을 사용하는 것이 더 좋다고 제안된 바 있는데 이는 그 매개변수들이 광범위한 변위에서 변동하고 있으므로 다양한 값을 가질 수 있기 때문이며 특히 직접적으로 관련된 명확한 물리적 의미가 없는 경우에도 사용할 수 있기 때문이다.

4. Wakeby 분포함수

다섯개의 매개변수를 가지며 역함수(inverse form)의 형태로 쓰여진 다음과 같은 식이 있다[4].

$$x=e+c(1-F)^{-d}-a(1-F)^{b}$$

여기서 F는 누가밀도함수 $F(x)$이며 상수 a, b, c, d, e는 $f(x)$의 매개변수들이다. 그리고 x는 주어진 값이다. Thomas 및 Houghton[4]은 이 함수를 많은 수의 연간 홍수자료에 적합시켜 보았다. 이 Wakeby 밀도함수의 다섯개의 매개변수를 산정하는 데는 계속적인 연구가 필요하나, 대량의 홍수자료에 적합시킬 때는 직교함수를 이용하는 최소자승법[5]을 이용하는 것이 가장 정밀하고도 계산시간이 적게 드는 것으로 판단된다. 이 함수가 역함수 형태로 되어 있는 자체가 매우 강한 장점으로서 인수(argument)들이 점이나 함수값이 아닌 적분값 또는 누가면적이기 때문이다.

Fiering[6, 7] 역시 지역화한 Wakeby 함수를 여러 지점의 갈수자료에 적합시켜 보았다. 또 Greenwood 등[8]은 Wakeby 밀도함수의 다섯개 매개변수를 산정하기 위해 확률 moments를 이용한 것을 제안하였으며 매개변수가 자료에 따라서는 매우 불안정한 Wakeby 밀도함수로 부터 구해진 확률산정값의 안정성을 Monte Carlo 해석에 의하여 분석한 바 있다.

일반적으로 다섯개의 매개변수는 Wakeby 분포의 두 tail의 위치 및 형태를 정의해준다. 그러나 실제로는 두개의 매개변수가 tail의 각각을 정의하기 위하여 필요하며 다섯번째 매개변수는 위치매개변수이다. 적절히 매개변수를 고정시킴으로써 Wakeby 밀도함수는 정규분포, 대수정규분포, Log-Pearson Ⅲ 및 흔히 사용하는 다른 중요한 밀도함수가 된다. 두 tail에 독립적으로 적합시킬 수 있는 유연성은 중요한 장점을 많이 나타낸다. 즉 다섯개의 매개변수가 자료집단 간에 큰 변화를 보인다 하더라도 누가 확률밀도함수(cdf)는 안정하게 된다.

5. 기존의 해석기법에 대한 새로운 판단기준

적합도 판정을 위한 종래의 수치기법과 두 함수를 비교하는 도해법은 둘 다 수문해석의 필요를 만족시키지 못하고 있다. 미국 수자원 위원회(US-WRC)가 규정하여서 특별한 경우를 제외하고서는 연방단위의 수자원 사업 분석에는 표준 홍수밀도함수로 Log-Pearson Ⅲ를 그리고 관련된 매개변수 추정방법을 사용할 것을 규정하고 있다[9]. 이러한 규정에서 한가지 흥미있는 것은 Log-Pearson Ⅲ에 적합시킬때 발생하는 고차의 moment의 표본변동폭을 고르게 하기 위해서 지역통계량(regional statistics)을 사용하는 것이다. 이러한 점은 수문해석에서 종종 부딪히는 작은 크기의 표본에서 왜곡계수(skew coefficient)가 불안정하기 때문이다. 지역통계량의 이용은 작은 교란에도 극히 민감하지 않은 통계기법 내지는 robust 통계기법의 채택에 중요한 발판을 제공해 준다. moment 들(홍수자료의 평균, 표준편차 및 왜곡도)을 추정하는 James-Stein[10] 기법은 손쉬우면서도 사용이 편하다. 이들의 기법에서는 주요 오차들이 쉽게 검출될 수 있다. 왜냐하면 수치들이 바르게 나타나지 않기 때문이다. 종래의 해석기법에서는 이들 모멘트들의 안정된 추정값을 얻는 것을 목표로 하고 있다. 그러나 작은 추정오차는 밀도함수가 관측값에 근사한 것으로 보일지 몰라도 중요홍수의 재현기간 산정에는 현저한 차이를 일으킬 수도 있다. 왜냐하면 평균값 주위에 밀도함수의 두터운 부분이 놓일때는, 즉 표준편차가 작은 경우에는

작은 오차라도 극한 수문사상의 재현기간 산정에 상당히 큰 변화를 초래하기 때문이다. 지점특성과 전국적인 관측 뿐만아니라 robustness와 관련된 수문사상의 분포를 해석하는 새로운 방법이 필요하다. 그러한 과정은 홍수지역으로 부터의 지역전체의 정보를 포함하게 되어 특정지점에서의 개별사상의 중요성은 강조되지 않게 된다.

지역홍수와 지점홍수의 조합해석은 Kuczera [11]에 의하여 연구되었으며 두개 매개변수의 분포형인 대수정규분포로 요약된다. James-Stein의 과정과 개념적으로 흡사한 Empirical Bayes(EB) 과정은 지점별 홍수특성과 지역적 홍수자료의 해석을 연결시켜 주고 있다. EB과정은 지역 홍수 정보가 홍수의 평균값 및 표준편차의 선행분포형의 매개변수를 추정하기 위하여 사용된다는 점외에는 Bayesian이다. 이러한 선행분포형은 지점홍수자료에 의하여 후속분포로 수정되어진다. 주어진 손실함수(loss function)에 의해서 재현기간 T-년 홍수의 최적 추정자가 유도될 수 있다. 모의발생에 관한 많은 연구들은 이용가능한 지역정보가 존재하면 짧은 자료기간에 대하여서도 언제나 100년 빈도의 홍수에 대한 EB추정자가 최우추정법에 의한 것보다 상당히 우수하다는 사실을 입증해주고 있다.

미국의 몇개 지역에서의 자료분석 결과는 공간 상관(spatial correlation)을 고려할 때 표준편차의 선행분포형에 따라서는 필요한 지역정보가 상당히 감소될 수 있음을 보여주고 있다. 선행분포는 지역화 과정 (regionalization procedure)을 사용하는 지형 자료와 함께 지역 홍수기록 확충에 의해 보다 정확함을 보여준다. New England 유역에 대한 자료해석결과는 표준편차의 선행분포에 의하여 수행되는 정보자료의 약 20년의 유효자료기간과 등가임을 나타내 보이고 있다.

이상치(outlier)문제는 수문학자들에게 있어서 가장 힘든 부분으로 종래의 해석 방법으로는 그것들을 어떻게 다루어야 할지를 모른다. moment 추정에 있어서 그것들을 포함함은 평균값과 표준편차 값에는 커다란 변화를 주지 않으나 설계사상과 관련있는 재현기간 산정에서처럼 왜곡계수 산정에는 크게 영향을 준다.

이상치가 극치사상(extreme events)의 추정에 주는 어려움을 규명하기 위하여 Houghton이 설정한 네가지 Wakeby분포를 사용하는 간편한 Monte Carlo실험을 수행한 바 있다. 이러한 Wakely분포는 대상지역에서의 홍수의 범위를 나타내준다. 즉 기록연한이 15년과 30년인 100년 빈도의 홍수에 대한 다음의 3가지 추정자에 대하여 평균제곱오차(mean-squared error)를 산정하였다. 추정자들은 대수정규 최대우도 추정자(log-normal maximum likelihood : ML), 위치추정에 대한 Tukey의 biweight 추정자[12]와 Lax의 축척추정자[13]를 이용하는 robust 대수정규 모멘트법(log-normal method of moment)추정자, 그리고 모멘트 추정기법인 log-Pearson Ⅲ을 이용하여 추정하였다. ML추정자에 대한 다른 추정자의 상대적 효율은 추정자의 평균 제곱오차에 대한 ML평균 제곱오차의 비로서 정의된다.

이러한 Monte Carlo추정결과는 표 1에 요약되어 있다. 표의 유형 Ⅳ에서 보는 바와 같이 robust 대수정규 추정자는 ML추정자보다 훨씬 좋은 결과를 보여주고 있다. 그러나 유형 Ⅲ의 경우에 있어서 robust추정자가 현저하게 나쁜 결과를 보여준다. 이것은 주어진 분포가 정규분포보다 더 두터운 꼬리를 가질때 위치 및 축적에 대한 robust추정자들은 moment추정가를 안전하게 하도록 고안됨에 반해 유형 Ⅲ의 경우는 그렇지 않기 때문이다. 따라서 위치 및 축적에 대한 적응 robust추정자가 필요하다. 다음으로, Log-Pearson Ⅲ추정자는 유형 Ⅲ을 제외하고는 추정자에 비교해 볼 때 좋은 결과가 나오지 않는다. 두 ML추정자가 위치 및 축적에 대한 거의 동일한 추정값을 사용하기 때문에 Log-Pearson Ⅲ추정자의 부정확한 결과치가 이상치 때문에 생기는 왜곡계수에 의한 것이라고 보는 것이 옳겠다. 그러나 Log-Pearson의 접근방법이 단지 ML추정자가 높은 편차를 보이기 때문에 버려야 할 것이라는 것은 아니다. 오히려 왜곡계수의 robust추정값이 개발되어야 하며 지점별 왜곡계수 추정자를 안정한 값으로 구하기 위해서는 먼저 지역왜곡계수를 사용해야 할 것이다.

表 1. Efficiency of Robust Log-Normal and Log-Pearson Estimators of the 100-year Flood for Different Distributional Assumptions and Record Length

Wakeby Distribution Parent Type	Mean-Squared-Error Efficiency of Estimator Relative to Log-Normal Maximum Likelihood Estimator(*)			
	Robust Log-Normal		Log-Pearson Ⅲ	
	15 Yr	30 Yr	15 Yr	30 Yr
Ⅰ	1.042 (0.037)	0.966 (0.013)	0.125	0.037 (0.058)
Ⅱ	0.966 (0.020)	0.922 (0.013)	0.397 (0.042)	0.502 (0.038)
Ⅲ	0.591 (0.015)	0.674 (0.006)	0.816 (0.089)	1.178 (0.091)
Ⅳ	1.697 (0.307)	1.235 (0.064)	0.037	0.109 (0.035)

* : where possible, standard errors are reported; plus and minus ranges are parentheses.

참 고 문 헌

1) Fiering, M.B., "*Alternative indices of resilience*", Water Resour. Res., Vol., 18, No. 1, 1982, pp. 33~39.
2) Fiering, M.B., and C.S. Holling, "*Management and standards for perturbed ecosystems*", Agro-Ecosystems, 1, 1974, pp. 301~321.
3) Fiering, M.B., "*A screening model to quantify resilience*", Water Resour. Res., Vol. 18, No. 1, 1982. pp. 27~32.
4) Houghton, J., Robust Estimation of the Frequency of Extreme Events in a Flood Frequency Context, Doctoral Thesis, Harvard University. 1977.
5) 조원철, 이재준. "직교함수를 이용한 최소자승법의 정밀도 향상에 관한 연구", 대한토목학회 논문집, 제 6 권 제 4 호, 1986 년 12 월, pp. 43~52.
6) Fiering, M.B., Consulting report prepared for the State of New Jersey, unpublished. 1978.
7) Fiering, M.B., Consulting report prepared for the Miami Conservancy District, unpublished. 1979.
8) Greenwood, J.A., et al., Probability Weighted Moments: Definition and Application, IBM Research Report RC 7108 (30439), 1978.
9) U.S. Water Resources Council, Guidelines to Determining Flood Flow Frequency, Bulletin 17A of the Hydrology Committee, Washington, DC. 1977.
10) Effron, B., and C. Morris. "*Stein's paradox in statistics*", Sci. Am., May 1977, pp. 119~127.
11) Kuczera, G., Robust Hydrologic Models at the Watershed Scale, Doctoral Thesis, Harvard University, 1980.
12) Launer, R.L., and G.N. Wilkinson, Robustness in Statistics, Academic Press, Inc., 1979, p. 5.
13) Mosteller, F., and J.W. Tukey, Data Analysis and Regression, Addison-Wesley Publ. Co., Inc., Reading, Mass. 1977.