

다층 신경망을 이용한 비중심F분포 확률계산

구 선 희†

요 약

ANOVA 검정에서 검정통계량은 단일 또는 이중 비중심F분포를 따르며 비중심F분포는 일반적인 선형 가설 검정에서 검정함수 계산에 적용되고 있다. 기존 비중심F분포의 함수 계산에 대한 연구로 여러 접근 방법이 제시되었지만, 하나의 정확한 함수값을 구하는데도 많은 시간이 소요되는 문제점이 발생되었다. 본 논문에서는 기존 함수 계산의 문제점을 해결하기 위하여 다층 퍼셉트론 네트워크로부터 역전파 학습 알고리즘을 적용하여 비중심F분포의 함수값을 구하는 방법을 제안하였다. 제안된 신경망에 의한 함수값과 기존 Patnaik이 제시한 분포식에 의한 함수값의 차이를 표와 그림을 통하여 비교하였으며, 정확성과 계산속도를 고려할 때 Patnaik의 함수식에 의한 방법보다 신경망을 이용한 방법이 효율적임을 알 수가 있다.

Computation of Noncentral F Probabilities using multilayer neural network

Sun Hee Gu†

ABSTRACT

The test statistic in ANOVA tests has a single or doubly noncentral F distribution and the noncentral F distribution is applied to the calculation of the power functions of tests of general linear hypotheses. Although various approximations of noncentral F distribution are suggested, they are troublesome to compute. In this paper, the calculation of noncentral F distribution is applied to the neural network theory, to solve the computation problem. The neural network consists of the multi-layer perceptron structure and learning process has the algorithm of the backpropagation. Using Tables and figs, comparisons are made between the results obtained by neural network theory and the Patnaik's values. Regarding of accuracy and calculation, the results by neural network are efficient than the Patnaik's values.

키워드 : 신경망(Neural Network), 역전파, 비중심F분포(Noncentral F Distribution), Patnaik 함수값(Patnaik's Values)

1. 서 론

비중심F분포는 1938년 Tang에 의하여 처음으로 제시되었으며 교호작용이 있는 분산분석에서 검정력 계산 및 선형 가설 검정에 적용되고 있다[10, 18]. 비중심F분포에서 분포함수(distribution function)의 값을 계산하는데에는 많은 시간이 소요될 뿐만 아니라 정확성이 문제가 되고 있다. 이로 인하여 이 분포함수에 대한 연구가 Chattamvelli, Johnson, Price, Seber, Sibuya, Tiku 등에 의하여 이루어졌다[3, 9, 12, 14, 15, 19].

지금까지 비중심F분포에 대한 연구가 주로 분포식을 유도하여 함수값을 구하는 방법으로 이루어졌으며, 이 방법이 함수의 정확한 값을 구할 수는 있으나, 단 하나의 함수값을 구하는데도 많은 시간이 걸리는 문제점이 발생되었다. 이로 인하여 본 논문에서는 비중심F분포의 함수값을 인공지능

분야에서 많은 연구가 이루어지고 있는 다층 퍼셉트론 네트워크로부터 역전파 학습 알고리즘[6-8, 17, 21]을 적용하여 이러한 문제점을 해결하고자 하였다. 비록 학습과정에서 적절한 은닉노드 및 여러 모수를 찾는 과정이 다소 시간이 걸리는 단점이 있지만 계속되는 하드웨어의 발전으로 인하여 앞으로 이러한 문제는 어느 정도 해결될 수 있으리라 생각한다.

먼저 자유도와 비중심모수에 따라 모형을 분류한 후, 학습과정에 필요한 자료는 Mathematica Program을 이용하여 Patnaik이 제시한 분포함수식[11]을 통하여 하나 하나 직접 구했으며, 학습과정에서 나온 모수를 가지고 구한 검정자료 값과 Patnaik의 함수값을 표와 그림을 통하여 비교하였다.

2. 회귀검정에서의 비중심F분포 확률

비중심F분포는 1928년 Fisher에 의하여 비중심베타분포에 대한 연구와 관련하여 처음 언급된 이후로, 1938년 Tang

† 정 회 원 : 전주대학교 교양학부 컴퓨터강의전담 전임강사
논문접수 : 2001년 1월 16일, 심사완료 : 2002년 4월 8일

에 의하여 처음으로 분포식이 유도되었다[5, 18]. 그후 1949년 Patnaik에 의하여 이 분포에 대한 연구가 활발히 이루어져 지금은 Tang보다는 Patnaik이 유도한 분포식이 더 잘 알려져 있다[11].

비중심F분포의 정의는 다음과 같다.

정의

자유도가 n_1 과 n_2 인 두 독립인 χ^2 통계량에 대하여 다음과 같은 비(ratio)를 F분포라 한다.

$$F_{n_1, n_2} = \left\{ \frac{\chi^2_{n_1}}{n_1} \right\} \left\{ \frac{\chi^2_{n_2}}{n_2} \right\}^{-1} \quad (1)$$

만일 중심 χ^2 이 비중심 χ^2 이라면 자유도가 n_1, n_2 이고 비중심모수 λ_1, λ_2 를 갖는 이중비중심F분포(doubly noncentral F distribution)는 다음과 같다.

$$F_{n_1, n_2}^{\lambda_1, \lambda_2} = \left\{ \frac{\chi^2_{n_1}(\lambda_1)}{n_1} \right\} \left\{ \frac{\chi^2_{n_2}(\lambda_2)}{n_2} \right\}^{-1} \quad (2)$$

이중비중심F분포에서 $\lambda_2=0$ (또는 $\lambda_1=0$)인 경우 단일비중심F분포(single noncentral F distribution)라 하며 일반적으로 단일비중심F분포를 간단히 비중심F분포라하며 F_{n_1, n_2}^λ 로 나타낸다[2, 10].

다음은 1949년 Patnaik가 제시한 비중심F분포의 누적합수식이다.

$$F_{n_1, n_2}^\lambda(x) = \sum_{j=0}^{\infty} \left(\frac{\left[\frac{\lambda}{2} \right]^j}{j!} e^{-\frac{\lambda}{2}} \right) \times I_w \left(\frac{n_1}{2} + j, \frac{n_2}{2} \right) \quad (3)$$

여기서 $w = \frac{n_1 x}{n_2 + n_1 x}$ 이며, $I_w(u, v)$ 는

$$I_w(u, v) = \frac{1}{B(u, v)} \int_0^w t^{u-1} (1-t)^{v-1} dt$$

로서 불완전 베타 함수(incomplete beta function)이다[11].

즉, 포아송과 불완전 베타함수 곱의 무한 급수 합으로 분포식이 이루어짐을 알 수 있다.

위 함수식의 정확한 값을 구하기 위하여 여러 접근 방법들이 연구 되었는데, 이 중에서 분포의 자유도 n_2 가 짝수인 경우에 Sibuya는 무한 급수 합을 유한 급수 합으로 유도하였으며[15], Seber는 불완전 베타함수를 다항식의 형태로 유도하여 분포 함수값을 구하였다[14]. 비록 위의 접근 방법으로도 함수값을 구할 수 있으나, 단 하나의 함수값을 구하는데도 다소 시간이 걸리는 문제가 발생함으로 본 논문에서는 함수값의 정확성과 계산시간을 고려하여 효율적인 방법을 찾게 되었다. 즉, 여기에서는 함수식을 유도하여 값을 구하기보다는 다층 퍼셉트론 네트워크로부터 역전파 학

습 알고리즘을 적용하여 함수값을 구한다.

역전파 학습 알고리즘에서 학습과정은 변환함수를 통과한 출력값과 목표값 사이에 가장 잘 적합되는 가중값을 찾는 과정이라 말할 수 있으며, 그 모수들에 대하여 살펴보면 다음과 같다.

은닉노드에 대한 가중값 a_{ij} 와 출력노드에 대한 가중값 b_j 를 가중값 ω 라고 하면, m 번째 학습 후의 변화된 가중값은 다음과 같다.

$$\omega_m = \omega_{m-1} + c_m \quad (4)$$

c_m 은 m 번째 학습 후의 가중값 ω 의 변화량이며 다음 식과 같이 적응적 학습률을 이용하여 구한다.

$$c_m = \mu c_{m-1} - (1 - \mu) e_m d_m, \quad 0 \leq \mu \leq 1 \quad (5)$$

여기서 e 는 가중값들이 변화됨에 따라 최소값으로 수렴되어 가는 비율을 조절하는 모수이며 다음과 같다.

$$e_m = \begin{cases} e_{m-1} + \phi, & d_m f_m > 0 \\ e_{m-1} \times \phi, & d_m f_m \leq 0 \end{cases}, \quad 0 < \phi < 1 \quad (6)$$

d_m 은 모든 예제들에 대하여 MSE를 각 학습량에서의 가중값으로 미분한 값들의 누적합이다. 따라서 m 번째 학습 후의 누적 가중값들은 다음과 같다.

$$d_m = \sum_{n=1}^N \left(\frac{\partial E}{\partial \omega_m} \right)_n \quad (7)$$

f_m 은 누적합이 최근 변화에 더 큰 영향을 받도록 하기 위한 향으로서 다음과 같다.

$$f_{m+1} = \theta f_m + (1 - \theta) d_m, \quad 0 \leq \theta < 1 \quad (8)$$

여기서 ϕ, θ 그리고 μ 는 학습률을 조정하기 위한 모수들로서 m 번째의 학습이 끝난 후에 주어진 허용오차를 만족할 때까지 역전파 학습 알고리즘은 계속 반복된다[1, 4, 13, 16, 20].

3. 시뮬레이션 및 결과 고찰

본 논문에서는 자유도와 비중심모수를 고려하여 여러 모형으로 분류한 후 다층 퍼셉트론 네트워크로부터 역전파 학습 알고리즘을 적용하여 비중심F분포의 함수값을 구하고자 한다.

네트워크는 입력층의 입력노드가 4개(분자와 분모의 자유도, 비중심모수, 백분위수)이며, 출력층의 출력노드가 1개(확률값) 그리고 1개의 은닉층으로 구성된 다층 퍼셉트론이다. 먼저 학습 알고리즘을 적용하기 전에 학습자료와 검증

자료에 대하여 확률은 0.95에서 0.99까지의 범위내에서 자유도(대·중·소)와 비중심모수에 따라 다음과 같이 모형을 분류한다. 임의로 자유도와 비중심모수의 크기를 고려하지 않고 추출하여 학습을 하면 학습오차가 커질뿐 아니라, 검정자료 또한 Patnaik이 제시한 함수값과 차이를 보이게 된다. 즉, 함수값의 정확성과 학습 계산시간을 고려한다면 분자·분모 자유도의 차이가 작을수록 그리고 모형을 세분화할수록 효율적임을 알 수가 있었다. 검정자료는 분자와 분모 자유도는 학습자료와 동일한 값으로, 그러나 비중심모수는 학습자료와 다른 값을 취하여 추출하였다.

〈표 1〉 모형 1

자료	확률(P)	분자자유도 (n_1)	분모자유도 (n_2)	비중심모수 (λ)
학습자료	P = 0.95 - 0.99	$n_1 = 1$	$n_2 = 2$	$\lambda = 1$ $\lambda = 5$
			$n_2 = 5$	$\lambda = 3$ $\lambda = 9$
		$n_2 = 2$	$n_2 = 2$	$\lambda = 2$ $\lambda = 6$
			$n_2 = 5$	$\lambda = 4$ $\lambda = 12$
검정자료	P = 0.95 - 0.99	$n_1 = 1$	$n_2 = 2$	$\lambda = 3$ $\lambda = 7$
			$n_2 = 5$	$\lambda = 6$ $\lambda = 12$
		$n_2 = 2$	$n_2 = 2$	$\lambda = 4$ $\lambda = 8$
			$n_2 = 5$	$\lambda = 8$ $\lambda = 16$

〈표 2〉 모형 2

자료	확률(P)	분자자유도 (n_1)	분모자유도 (n_2)	비중심모수 (λ)
학습자료	P = 0.95 - 0.99	$n_1 = 5$	$n_2 = 5$	$\lambda = 5$ $\lambda = 15$
			$n_1 = 10$	$n_2 = 10$
		$n_2 = 15$		$\lambda = 13$ $\lambda = 39$
		$n_2 = 20$		$\lambda = 15$ $\lambda = 45$
				$\lambda = 20$ $\lambda = 60$
		검정자료	P = 0.95 - 0.99	$n_1 = 5$
$n_1 = 10$	$n_2 = 10$			
	$n_2 = 15$			$\lambda = 26$ $\lambda = 52$
	$n_2 = 20$			$\lambda = 30$ $\lambda = 60$

〈표 3〉 모형 3

자료	확률(P)	분자자유도 (n_1)	분모자유도 (n_2)	비중심모수 (λ)
학습자료	P = 0.95 - 0.99	$n_1 = 15$	$n_2 = 15$	$\lambda = 15$ $\lambda = 45$
			$n_2 = 20$	$\lambda = 17$ $\lambda = 51$
		$n_1 = 20$	$n_2 = 20$	$\lambda = 20$ $\lambda = 60$
			검정자료	$n_1 = 15$
$n_2 = 20$	$\lambda = 34$ $\lambda = 68$			
$n_1 = 20$	$n_2 = 20$	$\lambda = 40$ $\lambda = 80$		

다음은 각 모형에 대한 학습자료이다. 동일한 자유도와 비중심모수에 대하여 확률 0.95부터 0.99 사이의 임의의 5개

〈표 4〉 모형 1의 학습자료

분자의 자유도 (n_1)	분모의 자유도 (n_2)	비중심모수 (λ)	백분위수 (x)	확률 (P)
1	2	1	37.5	0.949997
1	2	1	47.5	0.959998
1	2	1	64.1	0.969969
1	2	1	97.5	0.979999
1	2	1	197	0.989974
1	2	5	115	0.949940
1	2	5	145	0.959960
1	2	5	195	0.969976
1	2	5	295	0.979989
1	2	5	595	0.989997
1	5	3	22	0.946416
1	5	3	25	0.957506
1	5	3	30	0.969900
1	5	3	36	0.978999
1	5	3	51	0.989816
1	5	9	50	0.949205
1	5	9	56	0.959330
1	5	9	65	0.969897
1	5	9	79	0.979959
1	5	9	108	0.989830
2	2	2	38	0.949692
2	2	2	48	0.959802
2	2	2	64	0.969583
2	2	2	98	0.979950
2	2	2	198	0.989987
2	2	6	77	0.949931
2	2	6	97	0.959955
2	2	6	130	0.969898
2	2	6	197	0.979988
2	2	6	397	0.989996
2	5	4	15	0.944367
2	5	4	17	0.956045
2	5	4	20	0.967986
2	5	4	25	0.979657
2	5	4	34	0.989426
2	5	12	33	0.947046
2	5	12	37	0.957838
2	5	12	43	0.969007
2	5	12	52	0.979251
2	5	12	72	0.989833

의 값을 추출하였으며, 여기에서 제시한 확률값은 Patnaik가 제시한 비중심F분포의 누적 함수식을 가지고 Mathematica Program을 이용하여 계산한 함수값이다. 일반적으로 가설 검정에서 확률값은 0.95에서 0.99까지의 값만이 의미가 있으므로 가급적 그 범위내에서 백분위수를 정하였다.

〈표 5〉 모형 2의 학습자료

분자의 자유도 (n_1)	분모의 자유도 (n_2)	비중심모수 (λ)	백분위수 (x)	확률 (P)
5	5	5	9	0.940882
5	5	5	10	0.951846
5	5	5	12	0.966627
5	5	5	15	0.979068
5	5	5	21	0.989952
5	5	15	18	0.945523
5	5	15	20	0.955969
5	5	15	24	0.969857
5	5	15	29	0.979911
5	5	15	39	0.989579
10	10	10	5.7	0.948388
10	10	10	6.1	0.958201
10	10	10	6.7	0.969111
10	10	10	7.6	0.979806
10	10	10	9.2	0.989789
10	10	30	10.9	0.948724
10	10	30	11.7	0.959261
10	10	30	12.8	0.969881
10	10	30	14.3	0.979551
10	10	30	17.3	0.989848
10	15	13	5.5	0.948625
10	15	13	5.8	0.958053
10	15	13	6.3	0.969832
10	15	13	6.9	0.979404
10	15	13	8.1	0.989967
10	15	39	11	0.948399
10	15	39	11.7	0.959918
10	15	39	12.5	0.969757
10	15	39	13.7	0.979895
10	15	39	15.8	0.989756
10	20	15	5.4	0.947068
10	20	15	5.7	0.958100
10	20	15	6.1	0.969192
10	20	15	6.6	0.978868
10	20	15	7.6	0.989784
10	20	45	11.2	0.949991
10	20	45	11.7	0.959374
10	20	45	12.4	0.969525
10	20	45	13.4	0.979631
10	20	45	15.2	0.989883

자료를 앞에서 제시한 각 모형에 따라 분류하여 학습한 결과 다음과 같은 모수에서 학습오차가 최저가 됨을 알 수 있었다. 각 모형 모두 입력노드가 4개 (분자·분모의 자유도, 비중심모수, 백분위수), 은닉노드가 12개 그리고 출력노드가 1개 (확률값)로 이루어진 다층 퍼셉트론구조로서, 확률 0.95부터 0.99값에 대하여 식 (4)부터 식 (8)까지의 모수가 $x=0.009$, $\phi=0.3$, $\mu=0.4$, $\theta=0.3$ 에서 최적의 값을 가짐을 알 수가 있었다. 따라서, 이 모수로부터 구한 최종 학습결과와 연결강도를 가지고 비중심F분포의 확률계산을 하였다.

〈표 6〉 모형 3의 학습자료

분자의 자유도 (n_1)	분모의 자유도 (n_2)	비중심모수 (λ)	백분위수 (x)	확률 (P)
15	15	15	4.6	0.947679
15	15	15	4.9	0.959252
15	15	15	5.2	0.968086
15	15	15	5.7	0.978490
15	15	15	6.7	0.989762
15	15	45	8.9	0.949977
15	15	45	9.3	0.958349
15	15	45	10	0.969551
15	15	45	10.9	0.979365
15	15	45	12.6	0.989687
15	20	17	4.5	0.949671
15	20	17	4.7	0.958670
15	20	17	5	0.969135
15	20	17	5.4	0.978934
15	20	17	6.2	0.989912
15	20	51	8.8	0.949604
15	20	51	9.2	0.959382
15	20	51	9.7	0.968886
15	20	51	10.5	0.979525
15	20	51	11.9	0.989895
20	20	20	4.1	0.948839
20	20	20	4.3	0.959034
20	20	20	4.5	0.967122
20	20	20	4.9	0.978654
20	20	20	5.6	0.989694
20	20	60	7.9	0.949908
20	20	60	8.2	0.958318
20	20	60	8.7	0.969209
20	20	60	9.4	0.979681
20	20	60	10.6	0.989784

다음에서 제시하고 있는 표는 모형1, 2, 3의 전체 학습자료와 검정자료중에서 일부만을 임의로 추출한 것으로서 Patnaik의 함수값과 신경망이론에 의한 함수값의 차이인 오차를 비교하고 있다. 그림은 각 모형의 전체자료(110개 = 모형1의 40개 + 모형 2의 40개 + 모형 3의 30개)에 대한 학습자료와 검정자료로서 가로축은 모형1, 2, 3 순서로 모형 1의 첫 번째 자료 ($n_1=1, n_2=2, \lambda=1, x=37.5$)에서 모형 3의 마지막 자료 ($n_1=20, n_2=20, \lambda=60, x=10.6$)까지를 일련번호 1에서 110까지로 나타내고 있으며 세로축은 각 자료에 대한 Patnaik의 함수값과 신경망이론에 의한 함수값의 차이인 오차를 나타내고 있다. 그림에서 보면 학습자료는 전체자료중 ($n_1=10, n_2=20, \lambda=45, x=15.2$)인 80번째 자료가, 검정자료는 ($n_1=1, n_2=2, \lambda=7, x=194$)인 7번째 자료에서 가장 큰 오차를 보이고 있음을 알 수 있다. 참고로, 여기에서 사용한 검정자료도 학습자료처럼 주어진 모형하에서 확률이 0.95에서 0.99 사이에 있는 백분위수를 추출하여 사용하였으며 Patnaik의 확률값 또한 Mathematica Program을 이용하여 계산하였다. 즉, 표는 각 모형에서 일부 자료만을 추출하여 학습과 검정결과를 비교하고 있으며 그림은 모형에서 설정된 전체자료를 가지고 학습과 검정결과를 비교하고 있다.

〈표 7〉 Patnaik의 함수값과 신경망이론에 의한 함수값 비교 (학습자료)

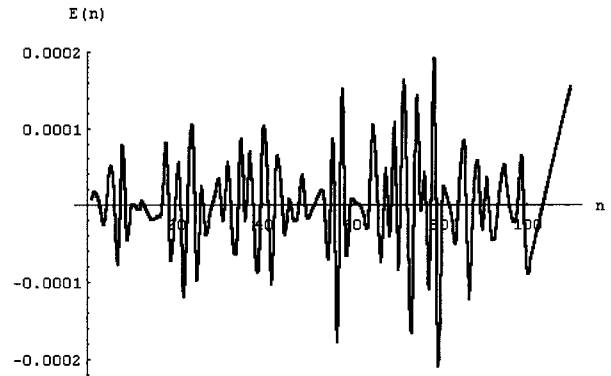
분자의 자유도	분모의 자유도	비중심 모수	백분 위수	Patnaik의 확률	신경망의 확률	오 차
1	2	1	37.5	0.949997	0.950004	0.000007
1	2	5	145	0.959960	0.960039	0.000079
1	5	3	30	0.969900	0.969875	0.000025
1	5	9	79	0.979959	0.980032	0.000073
2	2	2	198	0.989987	0.990087	0.000100
2	5	4	20	0.967986	0.968018	0.000032
2	5	12	52	0.979251	0.979337	0.000086
5	5	5	21	0.989952	0.989991	0.000039
5	5	15	24	0.969857	0.969870	0.000013
10	10	10	7.6	0.979806	0.979795	0.000011
10	10	30	17.3	0.989848	0.989844	0.000004
10	15	39	11	0.948399	0.948376	0.000023
10	20	15	5.7	0.958100	0.957930	0.000170
10	20	45	12.4	0.969525	0.969639	0.000114
15	15	15	5.7	0.978490	0.978542	0.000052
15	15	45	12.6	0.989687	0.989721	0.000034
15	20	17	6.2	0.989912	0.989860	0.000052
15	20	51	10.5	0.979525	0.979459	0.000066
20	20	20	4.5	0.967122	0.967095	0.000027
20	20	60	8.2	0.958318	0.958185	0.000133

〈표 8〉 Patnaik의 함수값과 신경망이론에 의한 함수값 비교 (검정자료)

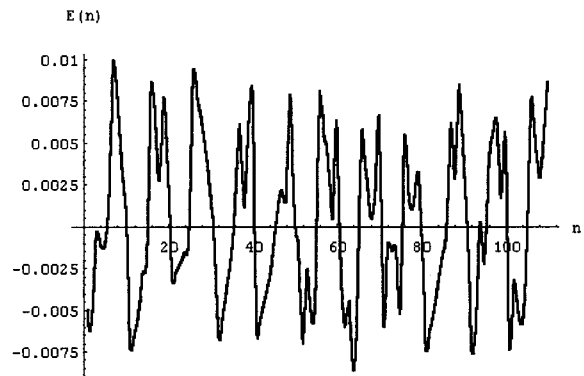
분자의 자유도	분모의 자유도	비중심 모수	백분 위수	Patnaik의 확률	신경망의 확률	오 차
1	2	3	196	0.979975	0.982733	0.002758
1	2	7	794	0.989998	0.988550	0.001448
1	5	6	80	0.989814	0.991718	0.001904
1	5	12	99	0.979789	0.972091	0.007698
2	2	4	297	0.989977	0.990701	0.000724
2	2	8	121	0.959812	0.952130	0.007682
2	5	8	32	0.969388	0.974584	0.005196
2	5	16	90	0.989775	0.981642	0.008133
5	5	10	30	0.989640	0.991179	0.001539
5	5	20	23	0.949092	0.946322	0.002770
10	10	20	8.9	0.958530	0.955746	0.002784
10	10	40	15.7	0.969402	0.964809	0.004593
10	15	26	10.3	0.979399	0.989041	0.009642
10	15	52	19.6	0.989802	0.983200	0.006602
10	20	30	8.3	0.948188	0.944346	0.003842
10	20	60	14.6	0.959086	0.957297	0.001789
15	15	30	7.6	0.968684	0.973720	0.005036
15	15	60	13.5	0.979869	0.971408	0.008461
15	20	34	9	0.989366	0.991707	0.002341
20	20	68	10.9	0.949250	0.945679	0.003571
20	20	40	6.3	0.959703	0.952970	0.006733
20	20	80	10.7	0.968614	0.965680	0.002934

각 모형에서 추출한 자료를 가지고 학습을 시킨 후 나온 모수를 가지고 다시 동일한 모형에서 비중심모수만 다르게 하여 추출한 새로운 자료를 가지고 검정한 후에 그 결과값을 Patnaik의 함수값과 비교하여 다음과 같은 결론을 얻을 수 있었다.

그림에서 확인할 수 있듯이 학습자료의 함수값은 대략 소수점 셋째 자리까지 Patnaik의 함수값과 동일하게 나오며 학습자료중 Patnaik의 함수값과 가장 큰 차이는 ($n_1=10$,



(그림 1) 학습결과의 오차 : 오차 $E(n)$ 는 Patnaik의 함수값과 신경망 방법을 이용한 함수값과의 차이이며, $E(80) = 2.14 \times 10^{-4}$ 의 최대오차를 나타내고 있다.



(그림 2) 검정결과의 오차 : 오차 $E(n)$ 는 Patnaik의 함수값과 신경망 방법을 이용한 함수값과의 차이이며, $E(7) = 9.799 \times 10^{-3}$ 의 최대오차를 나타내고 있다.

$n_2=20, \lambda=45, x=15.2$)에서 오차 $|0.989883(Patnaik) - 0.989669(신경망)| = 0.000214$ 값으로 거의 차이를 보이지 않고 있다고 볼 수 있다. 또한 검정자료의 함수값은 대략 소수점 둘째자리까지 Patnaik의 함수값과 동일하며 Patnaik의 함수값과의 최대오차는 ($n_1=1, n_2=2, \lambda=7, x=194$)에서 오차 $|0.959980(Patnaik) - 0.969779(신경망)| = 0.009799$ 값으로 검정 결과 또한 Patnaik의 함수값과 거의 차이가 없음을 알 수 있다. 즉, 검정자료가 Patnaik의 함수값과 신경망에 의한 함수값과의 차이인 오차로부터 좋은 결과임을 알 수 있으며 또한, 최적의 모수로부터 구한 최종 학습결과의 연결강도를 가지고 비중심F분포의 누적함수값 여러 개를 동시에 소요되는 시간 없이 구할 수 있었다. 이로부터 하나의 함수값을 구하는데도 시간이 걸리는 Patnaik 누적함수식에 의한 방법보다 많은 함수값을 빠르게 구할 수 있는 신경망을 이용한 접근 방법이 효율적임을 알 수가 있다.

4. 결 론

본 논문은 신경망 이론을 적용하여 비중심F분포의 확률

계산을 하였다. 이 신경망 이론은 다층의 구조를 갖는 복잡한 신경망 역전파 학습 알고리즘으로 연결강도 공간에서 오차표면의 기울기 감소를 기본으로 하였다.

비중심F분포의 확률 계산을 Patnaik의 함수값과 신경망을 이용한 함수값과 비교하여 학습결과와 검증결과 그리고 그에 따른 오차들을 그림을 이용하여 제시하였다. 신경망을 이용한 함수값과 Patnaik의 함수값을 비교한 결과 다음과 같은 결론을 얻을 수 있었다.

첫째, 학습자료의 함수값은 대략 소수점 셋째 자리까지 Patnaik의 함수값과 정확하며 검증자료의 함수값은 소수점 둘째자리까지 Patnaik의 함수값과 정확하게 나타남을 알 수 있다.

둘째, 학습하는데 소요되는 시간, 즉 신경망의 최적의 모수를 찾는 데 있어 많은 시간과 노력이 필요하지만 한 번 학습이 끝난 후에는 최종의 연결강도를 가지고 소요되는 시간 없이 비중심F분포의 확률 계산을 한다.

셋째, 정확성과 계산속도를 고려하여 볼 때 비중심F분포의 확률 계산을 Patnaik의 누적함수식에 의한 방법보다 신경망을 이용한 방법이 효율적임을 알 수 있다.

참 고 문 헌

[1] 김대수, "신경망 이론과 응용 (I), (II)", 하이테크정보, 1994.
 [2] P. J. Bickel and K. A. Doksum, "Mathematical Statistics : Basic ideas and selected topics," SanFrancisco : Holden-Day, 1977.
 [3] R. Chattamvelli, "On the doubly non-central F distribution," Computational Statistics and Data Analysis, Vol.20, pp. 481-489, 1995.
 [4] B. Cheng and D. M. Titterington, "Neural Networks : A review from a statistical perspective," Statistical Science, Vol.9, pp.2-54, 1994.
 [5] R. A. Fisher, "The general sampling distribution of the multiple correlation coefficient," Proceedings of the Royal Society of London, Series A, Vol.1, pp.654-673, 1928.
 [6] J. N. Hwang and H. Li. M. Maechler and D. Martin and J. Schimert, "Regression modeling in backpropagation and projection pursuit learning," IEEE transactions on neural networks, Vol.5, pp.342-353, 1994.
 [7] L. K. Jones, "A simple lemma on greedy approximation in Hilbert space and convergence rates for projection pursuit regression and neural network training," The Annals of Statistics, Vol.20, pp.608-613, 1992.
 [8] H. John and K. Anders and G. P. Richard, "Introduction to the Theory Neural Computation," Addison Wesley Publishing Co. Inc, 1991.

[9] N. L. Johnson, "On an extension of the connexion between the Poisson and χ^2 distributions," Biometrika, Vol.46, pp. 352-363, 1993.
 [10] N. L. Johnson and S. Kotz, "Continuous Univariate Distributions," Boston : Houghton Mifflin, 1970.
 [11] P. B. Patnaik, "The noncentral χ^2 and F-distributions and their applications," Biometrika, Vol.36, pp.202-232, 1949.
 [12] R. Price, "Some noncentral F-distributions expressed in closed form," Biometrika, Vol.51, pp.107-122, 1964.
 [13] D. E. Rumelhart and E. H. Geoffrey and J. W. Fonald, "Learning representations by backpropagating errors," Nature, Vol.323, pp.533-536, 1986.
 [14] G. A. F. Seber, "The noncentral chi-squared and beta distributions," Biometrika, Vol.50, pp.542-544, 1963.
 [15] M. Sibuya, "On the noncentral beta distribution function," Unpublished manuscript, 1967.
 [16] M. Simth, "Neural Networks for Statistical Modeling," New York : Van Nostrand Reinhold, 1993.
 [17] D. F. Specht, "A general regression neural network," IEEE transactions on neural networks, Vol.2, pp.568-576, 1991.
 [18] P. C. Tang, "The power function of the analysis of variance tests with tables and illustrations of their use," Statistical Research Memoirs, Vol.2, pp.126-150, 1938.
 [19] M. L. Tiku, "A note on approximating to the noncentral F-distribution," Biometrika, Vol.53, pp.606-610, 1975.
 [20] H. White, "Learning in artificial net-works : A Statistical perspective," Neural Computation, Vol.1(4), pp.425-469, 1989.
 [21] H. White, "Connectionist nonpara-metric regression : multilayer feed-forward networks can learn arbitrary mappings," Neural Networks, Vol.3, pp.535-549, 1990.



구 선 희

e-mail : guseonhee@hanmail.net

1986년 숙명여자대학교 수학과 졸업 (이학사)

1989년 숙명여자대학교 일반대학원 수학과 졸업(이학석사)

1996년 성균관대학교 일반대학원 통계학과 졸업(통계학박사)

1996년~2001년 전주대학교 정보기술컴퓨터공학부 컴퓨터 객원 교수

2001년~현재 전주대학교 교양학부 컴퓨터강의전담 전임강사
 관심분야 : 신경망이론, 전산수학