

부록 2. 분석방법 및 사후검정

IBM Corp. SPSS Statistics V.21: x86, Active Memory Exp. 1,028MB
MS Office 365 Excel: x64 (≈1028MB)

□ (데이터 설계/수집/분석/검정) N=545개 직위자의 프로필을 검색 및 수집하여 아래 항목과 같이 데이터 수집하여 분석/검정함.

항목	수집/관측 데이터 값	측도	분석/검정 방법
순번	숫자(ID)	명목	• 실명확인
사진	이미지	-	※ 분석/검정 시 성명 및 사진 삭제, 순번으로만 구분
성명	문자	명목	
출생년도 <도표4>	숫자	등간척도	• 실명확인 • 빈도분석 • 교차분석(독립/동질성검사)
중복/공석 (필터)	동일인(겸직자, 고등고시 동시 합격자) 공석	명목/실질	• 실명확인 • 측도기준 변경 통계표작성
기관형태 (소분류) <도표11(좌), 도표16(좌)>	행정부(정부기관), 공공기관(위탁집행형 준정부기관, 기금관리형 준정부기관, 시장형 공기업, 준시장형 공기업, 기타공공기관)		낙하산 인사 검정 <도표15> 회전문 인사 검정 <도표25>
기관업무 (소분류) <도표11(우), 도표16(우)>	경제(과학기술, 국토교통, 금융, 노동, 농업, 조달, 재정세제, 중소기업, 산업통상, 경제일반), 사회(보건, 복지, 교육, 문화, 법무, 안전, 환경, 사회일반), 외교안보(외교, 안보, 국방, 통일), 총괄(행정, 인사)	명목적도	• 기술통계 • 빈도분석 • 교차분석(독립/동질성검사)
자기 전문분야 <도표17>			
현재 소속기관 <도표12, 13, 14>	현재 소속기관(명) 내 출신자(수, 비중) 직전 소속기관(명) 내 출신자(수, 비중) 현재-직전 소속기관 내 출신자 증/감(수, 성장률) 공직출신별(명, 비중) 민간출신별(명, 비중)		분산도, 집중도, 이합집산도 측정 <도표22> • 적합도 검증 비중편차 확률밀도함수 가우스 적분값으로 적합도 단축 검증 - 편미분(베르누이, 블랙숄츠) - 등비수열(야코비 행렬) - 베타함수(가우스-페르마적분법)
직전 소속기관	출신기관 내/외 일원분산 표준편차 <도표22 이하 분산도> 전/후 비중편차값 <도표22 이하 집중도>	명목적도 등간척도 파이제곱측도	• 신경망분석 신뢰도 검증 F-분포 공변량 표준화 분석 및 방식형 누적분포함수 정규화 검정 <도표23> • 군집밀도 분석 및 품질 검증 - 다변량 공분산행렬 동질성 검정 균형위험도(Risk Parity) 측정 <도표23> 기관별균형위험도= $\sqrt{[(\omega_{\text{공직}} \cdot \sigma_{\text{전}})^2 + (\omega_{\text{민간}} \cdot \sigma_{\text{공직-민간}})^2 + 2(\omega_{\text{공직}} \cdot \sigma_{\text{전}} \cdot \omega_{\text{민간}} \cdot \sigma_{\text{공직-민간}})]} / \sum \omega \cdot \sigma_{\text{전체}}$
출신기관 <도표2, 18>	비중편차= $\sqrt{[(\omega_{\text{전}} \cdot \sigma_{\text{전}})^2 + (\omega_{\text{후}} \cdot \sigma_{\text{후}})^2 + 2(\omega_{\text{전}} \cdot \omega_{\text{후}} \cdot \sigma_{\text{전}} \cdot \sigma_{\text{후}})]}$ ω =관피아 비중, σ =일원분산 표준편차 출신기관별 누적분포함수 미분값(확률밀도함수)		



현재/직전 계급 <도표5, 6>	원수, 총리, 부총리, 장관, 차관, 준차관(위탁형, 시장형, 기타), 차관보, 1급, 2급, 3급	순서척도	<ul style="list-style-type: none"> •교차분석(독립/동질성검사) •신뢰도 검증
	전체 데이터수(545개) ÷ 계급 위 순서척도 (1~12)	등간척도	
공직/민간 출신 <도표2> 등	행정고시, 기술고시, 사법고시 및 사법연수원, 외부고시, 군인기수(육사, 해사), 4급, 5급, 7급, 9급, 비고시, 기타(미확인), 경찰, 소방, 공공기관, 학계(공공기관)	등간척도	<ul style="list-style-type: none"> •출신별 평균분석 •출신간 교차분석 •출신내 상관분석
	학계(사립대), 개안단체, 정치, 기업, 언론, 종교	명목척도	
경력 <도표26>	정부경력(유/무, 개월수)	명목척도	낙하산 인사 검정 <도표15>
	정치경력(유/무, 개월수)	명목척도	회전문 인사 검정 <도표25>
	민간경력(유/무, 개월수)	등간척도	<ul style="list-style-type: none"> •출신별 다중응답빈도/교차분석(독립/동질성검사) •출신간 상관분석
예산 <도표20>	행정부처별 예산(총지출기준), 공공기관별 사업수입 및 기금수입(정부지원)	등간척도	<ul style="list-style-type: none"> •출신별/간 평균분석 •출신간 교차분석(독립/동질성검사)
보수 <도표21>	「국가공무원법」(계급별 고정급 / 호봉별 기본급 + 직무급+ 12% 상여급) 연봉(상임기관장, 비/상임 이사 및 감사)	등간척도	<ul style="list-style-type: none"> •경력기간별 상관분석
권력순서	계급, 고등고시기수, 출생년도, 민간출신 경력기간(개월수)	순서척도 등간척도	<ul style="list-style-type: none"> •서열순 통계표작성 (부록1) •교차분석(독립/동질성검사) •<도표> 전체 신뢰도 검증

□ 검정기준

○ 정규분포: 표본 집단 신뢰도 95%수준

- (95%신뢰구간 가정) F-분포 양측 검정, 정규화 단측 검정, 3/4분위 일괄 폐기
- (표기) $P^* < 0.05$, $P^{**} < 0.01$, $P^{***} < 0.001$, $P \geq 0.05$: “신뢰할 수 없음 (기각)”

○ 교차분석: 출신 집단간/내 종속변수의 변량차이를 분석하여 독립변수를 조사함

- (원칙) 중립성 추정에 따라 모든 항목에 대해 출신 집단간/내 교차검증한 수치가 0.4 이상일 경우 종속관계 추정, 0.5 이상일 경우 상관관계 추정, 0.7 이상일 경우 영가설 가정 및 동질성을 검정함. 1.0의 경우 관측수 N=21개 미만일 경우 일괄 기각함 <상관계수 해석>.



- (검정) 명목 대 구간 : 람다¹²⁾

명목 대 명목척도: 파이 및 크레머V (Cramer V, 표기: “CV”),¹³⁾

명목 대 순서척도: 켄달의 타우-b, 타우-c¹⁴⁾

명목 대 등간척도: 에타¹⁵⁾

☞ 2개이상 독립변수 동질성 검정: U-검정 및 T-검정¹⁶⁾

< 상관계수 R 해석 >

(기각) $0.0 \leq \text{약} < \pm 0.4 \leq \text{중} < \pm 0.7 \leq \text{강} < \pm 1.0$ (기각역)

※ CV=±1.0 기각역: 1/4분위 내에서 관측수 N=21개(빈도수) 미만일 경우 기각¹⁷⁾

< 표기 >

0.0: “종속성 없음”

0.1~0.3수준 : “유의미한 차이 없음(0.1),” “종속성 추정 어려움(0.2),”
“약한 종속성(0.3)”

0.4~0.6수준 : “종속성 추정(0.4),” “유의미한 차이를 보임(0.5),”
“종속성을 갖음(0.6),”

0.7~0.9수준 : “강한 종속성을 갖음(0.7),” “결정 기준(0.8),” “명백한 차이(0.9)”

1.0: “신뢰할 수 없음 (표본 신뢰도: N/21%)”

12) 독립변수의 값이 종속변수의 값을 예측하는 데 사용될 때 오차 내 비례 축소를 반영하는 연관 척도입니다. 값 1은 독립변수가 종속변수를 완전하게 예측한다는 의미이고, 값 0은 독립변수가 종속변수를 예측하는 데 도움이 되지 않는다는 의미입니다 (IBM, 2021) URL: <https://www.ibm.com/docs/ko/spss-statistics/beta?topic=crosstabs-statistics>

13) 파이는 카이제곱 통계량을 표본 결과로 나누고 그 결과의 제곱근을 구하는 연관 척도에 기반한 카이제곱입니다. Cramer의 V는 카이제곱을 기반으로 한 연관 척도입니다 (ibid).

14) Kendall의 타우-b: 동률을 고려하는 순서변수나 순위변수에 대한 비모수 상관 척도입니다. 계수의 부호는 관계의 방향을 나타내고 이에 대한 절대값은 강도를 나타냅니다. 이때 절대값이 클수록 관계가 더 강력합니다. 가능한 값의 범위는 -1부터 1까지이지만 -1 또는 +1 값은 제곱표에서만 구할 수 있습니다; Kendall의 타우-c: 동률을 무시하는 순서변수에 대한 비모수 연관 척도입니다. 계수의 부호는 관계의 방향을 나타내고 이에 대한 절대값은 강도를 나타냅니다. 이때 절대값이 클수록 관계가 더 강력합니다. 가능한 값의 범위는 -1부터 1까지이지만 -1 또는 +1 값은 제곱표에서만 구할 수 있습니다 (ibid).

15) 0과 1 사이에 존재하는 연관 척도로서, 0은 행 변수와 열 변수 간에 연관이 없음을 나타내고 1에 가까울수록 연관 정도가 높음을 의미합니다. 에타는 구간 척도(예: 수입)에 대해 측정된 종속변수와 범주(예: 성별)의 수가 제한된 독립변수에 적합합니다. 두 에타 값이 계산되는데, 하나는 행 변수를 구간변수로 처리하고 다른 하나는 열 변수를 구간변수로 처리합니다 (ibid). 에타의 경우 0.14보다 클 때 ‘강한 종속’을 갖음

16) 독립성 검정에서 제외된 외부적인 다른 요인에 의해 영향을 받을 수 있으므로, 검정하려는 요인과 다른 요인 간의 평균적인 차이가 있는지 여부를 판별하여 독립변수의 신뢰도를 검정하는 것을 말함. URL: <https://www.ibm.com/docs/ko/spss-statistics/beta?topic=tests-independent-samples-t-test>

17) 본 통계분석에서는 기획재정부(N=10, CV=1.0) 등 집단 내 관측값 21개 미만일 경우 신뢰하지 않고 일괄 기각함. 그러나 관점에 따라서는 ‘통계적 사실’로서 추정할 수도 있는 실측값임을 알림. 참고로, SPSS가 제시하는 CV 해석 기준은, ‘약 $\leq 0.2 < \text{중} \leq 0.6 < \text{강}$ ’으로서, 본 통계조사는 보다 엄격한 기준에서 독립성 검정을 실시함.



□ 모델 검정 및 적합도 검증

○ (계급) 아래 <분류기준>에 따라 1~12순위로 순서/등간척도로 분류하여 교차검정 및 그 신뢰도를 검증함.

- (순서척도) 1=원수, 2=총리, 3=부총리, 4=장관, 5=차관, 준차관(6=위탁형, 7=시장형, 8=기타), 9=차관보, 10=1급, 11=2급, 12=3급

- (등간척도) 순서척도의 신뢰도를 검증하기 위해 '전체 데이터 수(N=545개) ÷ 계급 순서척도(1~12)'로서, 원수=545, 총리=545/2, 부총리=545/3, 장관=545/4, 차관=545/5, 준차관(위탁형공공기관장=545/6, 시장형공공기관장=545/7, 기타공공기관장=545/8), 차관보=545/9, 1급=545/10, 2급=545/11, 3급=545/12 순으로 계급간 권력값을 역수로 측정하여 영점을 제거하고 등간척도(모델값)를 사용하여 위 순서척도의 적합도 (영가설: "출신집단별 계급간 상명하복의 상관관계는 동일하다.")를 검증함.

∴ (검정결과) 순서척도 및 등간척도 모두 동일한 관측값이 각각 관측되었으므로 영가설 기각은 타당하고 분석결과를 신뢰할 수 있음 <도표5, 6>.

< 직책에 따른 공직자 계급 분류기준 >

Table with columns for rank (대수, 급수) and job titles (일반공무원, 외교공무원, 조속등교원, 고등교원, 차안공무원, 국정직공무원, 소방공무원, 군인, 검찰, 연구직, 지도직, 전문경력관, 일반부, 사법부, 지방자치행정, 지방관련행정, 공기업, 기타). Rows list various ranks from 원수 to 1급 and corresponding job titles.



○ (이합집산도) ‘이합집산(離合集散)’은 흩어졌다가 모였다가 한다는 뜻으로, 출신집단별 분산 및 집중의 변동성(비중간 표준편차)에 따라 확률공간 내 분포가 중첩된 방사형 누적분포함수의 정규값<도표22>로서 명명하고 이를 아래와 같이 측정한 뒤 모델값으로 적합도를 검증함 .

- (정의) 누적분포함수는 관측된 확률공간(Ω -표본공간, F-분포, Pr-확률) 내 변수(X)가 특정 실수값(x_n)보다 작거나 같은 확률변수(R)를 등비수열 (1)의 합 공식으로 나타낸 베타함수(2)로 정의됨.¹⁸⁾

$$(1) - F_X(x_1, \dots, x_n) = \Pr(X \leq x_n \in -\infty, x) \quad \forall x \in R$$

$$(2) - X = (X_1, \dots, X_n) : \Omega \rightarrow (R, \beta(R^n))$$

$$\beta(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}, \quad \Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$$

$$\beta(x, y) = 2 \int_0^{\pi/2} \sin^{2x-1} \theta \cos^{2y-1} \theta d\theta, \quad R(x) > 0, R(y) > 0$$

$$\beta(x, y) = \int_0^\infty \frac{t^{x-1}}{(1+t)^{x+y}} dt, \quad R(x) > 0, R(y) > 0$$

18) 베타함수(Euler, 1730)는 ①오일러적분(1종 베타), ②가우스적분(Gauss-Wantzel, 1801-1836), 외에도 변수(t)로 치환하여 ③감마적분(2종 베타), ④야코비행렬(Jacobi, 1804-1851)로 계산할 수 있음. 본 통계분석에서는, 관피아 비중에 따른 분산도 및 집중도에 가중치가 부여된 표준화된 비교척도인 ‘비중편차[조화평균과 달리, 비중편차는 포트폴리오 선택이론(Markowitz, 1952)의 분산 위험에 따른 비중간 표준편차(피타고라스 평균)]’를 측정하여 F-분포 집단 내 일원분산 및 집단간 분산-공분산행렬을 표준화(standardization)하여 확률밀도를 (1)과 같이 추정하였고, 이에 따라 편미분방정식을 적용한 관피아 비중의 변화량(확률밀도)을 (3)과 같이 델타(d_1, d_2)계수로 하는 정규화(normalization, R=95%)된 불완전 베타함수(4)를 적용함. 이에 불완전 베타함수를 구하기 위해 이합집산도 검증 단계에서 누적분포 내 확률밀도의 다변량 공변량을 측정하기 위해 ⑤베르누이 미분방정식(Bernoulli, 1738)으로 비선형 베타함수 모수를 추정하여, 야코비행렬 누적분포도함수 및 가우스적분 누적분포함수를 구하여 적합도를 검증하고 이합집산도를 제시함. 사후검증단계에서는 해당 모형(불완전 베타함수)의 불확실성(극좌표계 벡터 및 표준 오차 허수값의 브라운 운동에 따른 엔트로피)를 확인하여 신뢰도를 검증하기 위해 ⑥블랙-숄츠 미분방정식(Black-Scholes, 1997) 모델을 인용하여 파이제공급척도로서 이합집산도를 재검증함. 즉, 정규화된 이합집산도 모델 (5)은 중립성 추정에 따라 출신지(n=545개 중 389개)를 요인으로, 출신집단(직전소속기관)내 이항계수 $F(d_1, d_2)$ 를 분할변수(3)로, 출신집단(현재소속기관)간 다항계수 $F(x_1, \dots, x_n)$ 를 종속변수(1)로 알고리즘을 설계하여 ‘방사형 공변량분석(신경망분석)’을 통해 앞서 추정된 모수, 배타계수, 상수, 독립변수의 적합도를 재검증하고, 또한 관 피아간 비중편차 가중치 역시 종속변수(H)간 군집밀도에 따른 신경망 가중치(synaptic weight)로 재검정하여 엔트로피를 측정하여 신뢰도를 검증했다. 마지막으로, 표준화된 F-분포 내 출신집단별 요인들의 집중도와의 동 질성/상이성의 차이를 살펴보기 위해 K-평균군집분석 및 정규분포 내 요인들간 군집밀도(DBSCAN, 소위 ‘가장 가까운 이웃’)분석을 통해 데이터의 품질을 검증했고, F-분포의 벡터에 따라 조정된 정규분포인 <도표22>와 더불어 누적분포 모수의 총량을 1로 조정된 균형위험도인 <도표20>로 환산하여 함께 제시함. 즉, ①②③⑤로 측정 분석, ④⑥으로 재검증했다.



- (분산도) 출신기관별 누적분포를 조사하기 위해 표본공간 n=545개 직위의 개별 인사이동에서 일어날 수 있는 직전 소속기관과 현재 소속기관 간의 1:1 인사이동에 따른 두 기관 내 직위 수(변수)의 변화량을 각각 d_1, d_2 의 변화율(델타계수, dx, dy 편미분값)로 하는 일원분산을 측정(3)하여 출신기관(중립가정 Z =종속변수)별 분산도를 표준편차(σ)로서 제시함. 19)
- ☞ 해석: 절대값이 클수록 분산도가 높음

$$(3) - d_2 = \frac{R_{\text{후}}/R_{\text{후}}^2}{R_{\text{전}}/R_{\text{전}}^2} = d_1 - \sigma \sqrt{t}, t = \sin^2 \theta, \theta = \frac{n_z}{n}$$

$$\sigma = \sqrt{(w_{\text{전}} * \sigma_{\text{전}})^2 + (w_{\text{후}} * \sigma_{\text{후}})^2 + 2(w_{\text{전}} * \sigma_{\text{후}})}$$

< (3)일원분산 분석 결과 >

변동요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
집단-내	13769.96	51	269.9992	5.688801	1.95E-09	1.587389
집단-간 (조합)	5.46E-12	1	5.46E-12	3.43E-14	1	3.934253

- ☞ 해석: $n_z=51$ 개 출신기관, $n=545$ 개 직위 중 공석 12개를 제외한 케이스처리 유효값 $N=533$ 개 직위의 비중편차(분산도)에 영향을 주는 독립변수=51개, “직전/현재 소속기관간 1:1 인사이동에 따른 분산도(비중편차)가 동일하다.” 는 영가설은 거짓임을 확인.

∴ 출신여부에 따라 일원분산에는 차이가 있고 출신기관을 독립변수로 가정하는 것은 타당함. 다만, 기관간 동분산에도 영향을 미치는지 일원분산만으로 추정할 수 없음. 다변량 공분산분석이 요구됨.

19) 분산도는 ①분산, ②표준편차, ③확률밀도함수를 일반적으로 사용함. 집중도/분산도 모두 일원분산 값에 관피아 비중간 표준편차를 적용하여 통일된 척도로 제시함.

한편, 분산도 외에도 ‘군집도 분석’을 통해 관피아 51개 출신기관 내 산하기관을 2계층으로 분할하고 주부 부처 내 공공기관을 포함한 직위 점유율 대비 타 부처 내 점유율 간의 ‘분산-공분산 행렬’을 추정하여 집단내/간 K-군집평균(조합평균)을 분석함으로써 동분산성의 차이를 통해 독립변수를 추정하는 방법도 있으나, 본 통계분석에서는 군집도 분석에 사용되는 조합평균을 이용치 아니하고 관피아 비중을 가중시키기 위해 피타고라스 평균을 채택함. 이에 따라 야코비행렬 누적분포함수(도)로 우선 재검정하였고, 신경망분석을 통해 집단내/간 다변량 공분산을 추정하여 정규화된 방사형 누적분포함수도를 제시함. 또한 사후 검증단계에서 F-분포 내 군집밀도(소위 ‘가장 가까운 이웃’)분석을 통해 집중도(군집도)의 품질검사를 했고 모형의 신뢰도를 살펴봄. 데이터 설계 단계에서 계층분할 모형을 좀 더 개량한다면 군집분석을 유의미한 측정도구로서 사용할 수도 있음을 알림.



- (집중도) 출신기관별 분산도가 중첩된 확률, 즉 관피아 세력별 비중편차 (3)의 결과값들을 각각 (2)에 대입하여 누적분포함수[F(x)]의 미분값인 확률밀도함수[f(x)]를 구한 그 결과값들(4)이 관피아별 비중의 변화량 [F(d₁, d₂)]이 (1)의 정의에 따라 그 조건을 만족하지 않는 경우(X > x_n, 허수의 확률)를 제외하고, 나머지 유의미한 출신기관별 비중편차(3)를 집중도로서 제시함. 해석: 절대값이 0에 가까울수록 집중도가 높음.

$$(4) - F(x) = \int_{-\infty}^x f(t)dt, f(x) = e^{-x^2} \frac{1}{NMI} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}y^2} dy = 1 - \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{1}{2}y^2} dy$$

$$F(d_1, d_2) = \frac{1}{\beta(d_1/2, d_2/2)} \left(\frac{d_1x}{d_1x+d_2}\right)^{d_1/2} \left(1 - \frac{d_1x}{d_1x+d_2}\right)^{d_2/2} x^{-1}$$

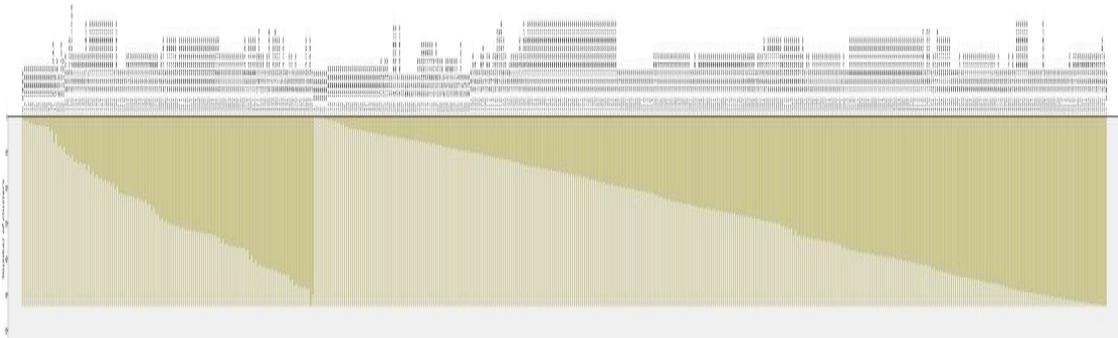
$$F(z_1 \dots z_n) \approx \Pr(X \leq z_n \in x, \infty) \quad \forall_x \in R_z$$

<(4)출신기관별 F(d₁, d₂), Rn=545개 케이스처리 결과 >

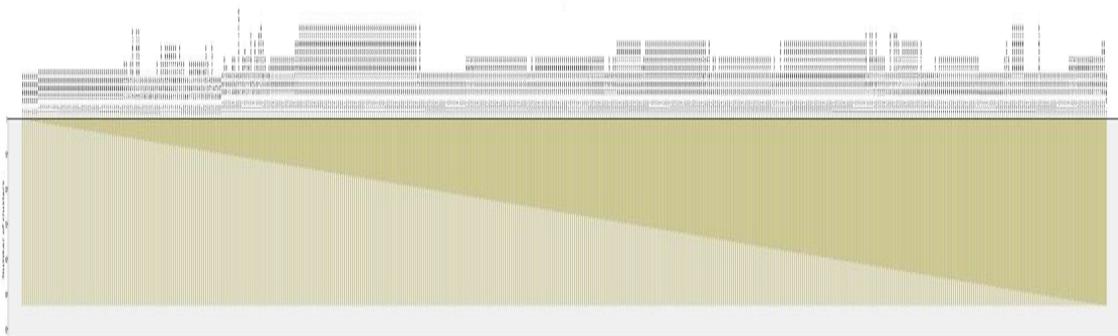
- 야코비행렬 누적분포도함수 -

*척도: 표본공간 / 확률공간 / (기울기) 확률밀도

*F(x,y)편미분 구간: $-\infty < d_1 < x, y < d_2 < \infty$



- (1)의 F(X1...Xn) 등비수열의 벡터합 $\forall_{545} = R$ -



〈 모형 요약 〉		〈 분산분석 결과 〉					
다중 R	.997	제곱합	자유도	평균 제곱	F	유의확률	
R 제곱	.995	회귀분석	6358873.194	2	3179436.597	2775.747	.000
수정된 R 제곱	.994	잔차	34363.040	30	1145.435		
추정값의 표준오차	33.844	합계	6393236.233	32			
로그-우도 함수 값	193.069						

	비표준화 계수		표준화 계수		t	유의확률
	B	표준 오차	베타	표준 오차		
상수(α)	.001	.000			13.634	.000
집단내(d_1)	.001	.000	.475	.027	17.851	.000
집단간(d_2)	.001	.000	.558	.027	20.968	.000

해석: 51개 출신기관별 관피아 비중 변화량(전)을 독립변수, 출신기관간 비중 변화량(후)를 종속변수로 가정하였을 때, 기관 내 일원분산 값으로부터 독립변수로 가정된 51개 중 기관간 비중 변화량에 유의미한 차이가 있는(동질성이 없는) 독립변수는 최소 2개(정규분포 오차범위를 감안 하면 최대 3개: 기획재정부, 대통령실, 교육부 순)임.

위 모형(누적분포도함수)의 설명력은 99.4%, 누적분포함수의 총량(모수)은 43.117이며, 확률밀도는 정규분포 내 비표준화 베타적분 계수는 0.001로서, 즉 관피아가 1명 증가하면 그 비중은 +0.1%수준으로 증함.

∴ 위 누적분포함수 모형은 표준분포(오차율 2.7% 수준의 F-분포)로서, 이를 95%신뢰구간 내 정규분포로 불완전 베타적분치환(즉, 비표준화 계수를 사용²⁰⁾) 하는 것은 적합함. 물론, 야코비행렬에서 음수의 표준편차는 발견되지 아니하였으나, 정규화 과정에서 발생하는 표준오차 구간의 극좌표계 내 허수확률(치환과정에서 발생될 수 있는 음의 확률로 인한 누적분포함수의 불확실성)까지 재검증하고자 아래와 같이 신경망분석(다변량 공변량분석)하고 이를 통해 도출된 방사형 누적분포함수도를 이합집산도(<도표22>)로서 제시하고자 함.

20) 유체역학에서 사용하는 확률밀도함수를 구하기 위한 베르누이 편미분방정식에 의한 가우스-페르마의 비표준화(비선형 다항계수) 베타적분법을 말함. 물론, 야코비행렬 누적분포도함수를 살펴봤을 때 선형계수가 눈에 보이므로 위 사례에서는 뉴턴-베르누이 베타적분법이나 고속적분법을 그냥 사용해도 무관할 것으로 추정됨. 그러나 본 통계분석에서는 비선형 다항계수를 가정했고, 사후검증을 위해 역산과 치환적분이 가능한 가우스적분법을 사용함.

- (신경망분석) 모델 검증을 위해 다른 편미분방정식(블랙-숄츠, 1997)을 사용하여 n=545개(요인, 독립변수), 출신여부(분할변수), 인사이동에 따른 출신기관간 전/후 비중 변화량(종속변수) 간의 ‘다변량 공변량분석(신경망분석)’ 알고리즘(5)을 설계하면 다음과 같음.

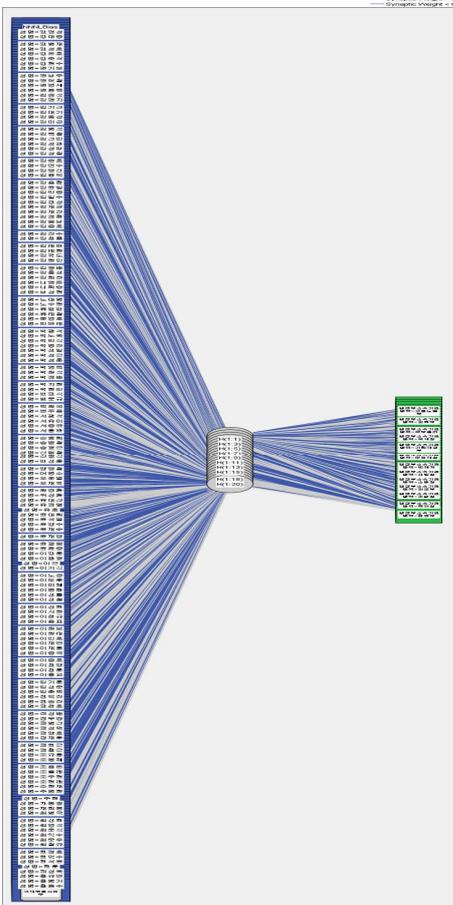
$$(5) - C(n_{\text{전}}, z) = n_{\text{후}} \Phi(d_1) - n_{\text{전}} e^{-\alpha z} \Phi(d_2)$$

$$d_1 = \frac{\ln\left(\frac{n_{\text{후}}}{n_{\text{전}}}\right) + \left(\alpha + \frac{\sigma^2}{2}\right)z}{\sigma\sqrt{z}}, \quad d_2 = \frac{\ln\left(\frac{n_{\text{후}}}{n_{\text{전}}}\right) + \left(\alpha - \frac{\sigma^2}{2}\right)z}{\sigma\sqrt{z}} = d_1 - \sqrt{z}$$

$$n_{\text{후}} \phi(d_1) = n_{\text{전}} \phi(d_2), \quad \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{1}{2}z^2} dz = 1 - \frac{1}{\sqrt{2\pi}} \int_z^{\infty} e^{-\frac{1}{2}z^2} dz$$

<도표23> 신경망분석 및 엔트로피 측정



< 시넵스 가중치 케이스처리 요약 >

		N	퍼센트
표본	훈련	389	100.0%
	유효	389	100.0%
	제외됨	162	
	합계	551	

< 네트워크 정보 >

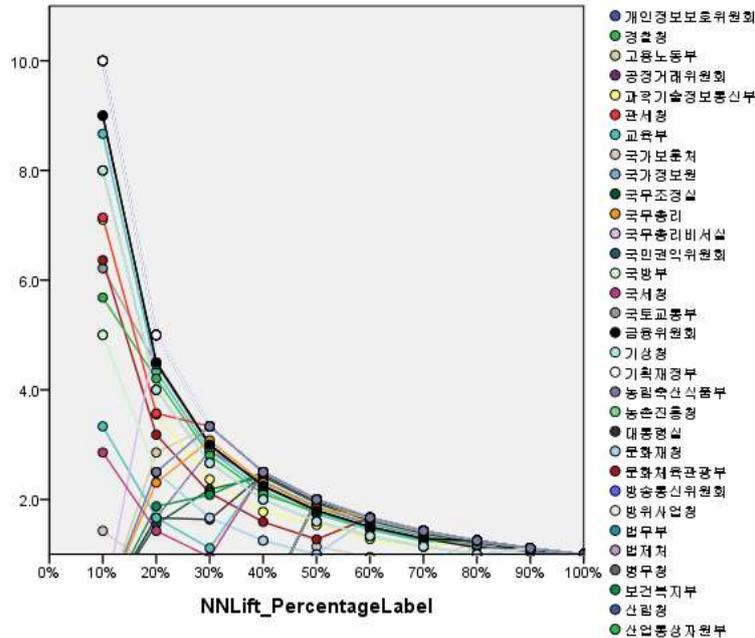
입력 레이어	요인	1	성명
	공변량	1	공직출신비중
	단위 수a		357
숨겨진 레이어	공변량 조정 방법		정규화
	숨겨진 레이어 수		1
	숨겨진 레이어 1에서 단위 수a		20
출력 레이어	활성화 함수		쌍곡 탄젠트
	종속변수	1	현재행정부소속기관
	단위 수		46
	활성화 함수		Softmax
	오차 함수		교차-엔트로피

a. bias 단위 제외

< 모형 엔트로피 검정 측정 요약 >

	교차 엔트로피 오차	20.847
	퍼센트 잘못된 예측	2.3%
훈련	사용된 중지 규칙	기간의 최대 수(100)가 초과되었습니다.
	훈련 시간	0:00:01.83
	종속변수:	현재행정부소속기관

<도표23> 방사형 누적분포도함수 다변량 공변량분석 결과



해석: <도표23>의 관피아 n=389개 직위 중 출신여부에 따라 관피아 비중에 영향을 주는 출신기관(분할노드)의 변수는 20개로 추정되며, H=20개 중 출신여부에 따라 전체 공변량에 영향을 주는 설명변수로서 1위 기획재정부, 2위 대통령실, 3위 교육부 순으로서 독립변수의 중요도가 나타나 (1)의 결과와도 일치하는 경향을 보였고 (위 방사형 누적분포도함수의 상단부터 하단 순, 신뢰구간 95% 내 설명변수 총 20개 기관들로 한정된 <도표22>의 시계열 순위), 위 (3)의 집단간 동분산 가정에 대한 유의미한 차이가 있는 독립변수는 기획재정부 및 대통령실로 확인됨. 위 방사형 누적분포도함수의 10~20%구간에서 표준오차 내 벡터값에 확률밀도가 하락하는 기관들(검찰청, 국토교통부, 국무조정실)의 순위는 좀 달라질 수 있으나 95% 신뢰구간 범위 내에 있어서 유의미한 차이가 없으며 “동률”로 추정할 수 있음.

- (신뢰도) 한편, <도표22>에 대한 총 10회 기계학습 결과 정규화과정에서 불확실성은 2.3%(엔트로피: 20.7~20.8사이)로 측정되어 위 모형의 예측력은 97.7%로 확인됨. 독립/종속 변수간에 미치는 영향에 대한 신뢰도를 검정하기 위해 더미변수 총 n=10개(5개×2회)를 무작위로 랜덤기계학습(훈련수: 364, 370회)한 검정 오차율은 다음과 같음.

<도표22> 모형 검정결과 요약

	오차제곱합	55.838		오차제곱합	48.154
훈련 (370회)	퍼센트 잘못된 예측	22.2%	훈련 (364회)	퍼센트 잘못된 예측	15.9%
	훈련 시간	0:00:38.53		훈련 시간	0:00:35.38
검정 (더미변수 5개)	오차제곱합	.021a	검정 (더미변수 5개)	오차제곱합	.156a
	퍼센트 잘못된 예측	0.0%		퍼센트 잘못된 예측	0.0%

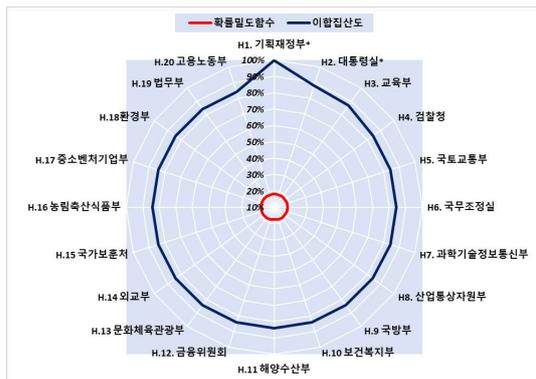
종속변수: 현재행정부소속기관

a. 숨겨진 단위 수는 검정 데이터 기준에 따라 결정됩니다. 숨겨진 단위의 "최상의" 수는 검정 데이터에서 가장 작은 오차가 발생하는 1입니다.

< (4) 및 (5)에 따른 정규분포 모형적합도 검정 결과>

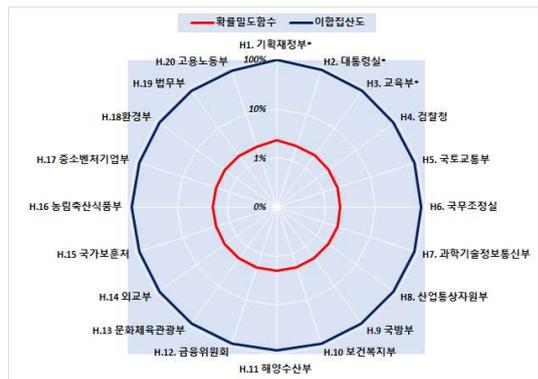
- (4)가우스 분포 <도표22> -

산식: 베르누이(1738), 척도: 0~1, 설명력=99.4%**, 자유도=2



- (5)로그 분포 <도표23> -

산식: 블랙-숄츠(1997), 척도: 파이제곱, 설명력=97.7%, 자유도=3



∴ 정규분포 신뢰구간 95% 범위 내 이합집산도 <표22>의 기관들과 일치하며, 순위는 1~3순위까지 강력한 신뢰가 추정됨. 위 검증결과와 같이 정규분포 95%신뢰구간 내 (4) 비선형 누적분포도함수의 설명력은 99.4%, (5) 방사형 누적분포도함수의 설명력은 97.7%, 모형의 예측력은 100%이므로 <표22>의 이합집산도는 타당함.

다만 군집도 분석 결과, 검찰청, 국토교통부, 국무조정실의 경우 ‘잡음(노이즈, 허수확률에 의한 누적분포값)’으로 추정됨. 그 원인은, ①원데이터 오류가 아니라면, ②표준오차의 정규화 과정에서 오류, ③위 모형의 연역적 한계로 추정할 수 있음. 이는, 추정된 독립변수 2~3개[즉, (4)에서 2개, (5)에서 3개]의 경우를 몇 개로 가정하냐에 따라서, 군집밀도분석에서 추정된 잡음을 제외하고 이합집산도 내 순위는 다소 조정될 수도 있음. 이에 따라, F-분포에서



확인된 H=20개 기관의 각 벡터값에 따라 정규분포 내 시계열로 조정하고, 조정된 정규화 방사형 누적분포도함수에서 소수점 둘째자리 미만 오차를 반올림하여 누적분포도함수의 결과값이 같은 기관들은 “동순위”로 추정함.

- (균형위험도) 위 (3)비중편차에 따라 모수로 추정된 (4)의 누적분포함수의 총량(43.177) 대비 관피아 비중값인 기관별 확률밀도에 민간출신자와의 비중편차를 가중시켜 더한 값(6)을 내림차순으로 정리하면 <도표24>과 같음.

$$(6) - 1 = \frac{\sum PDF_n \sqrt{(\omega_{\text{공직}\sigma_{\text{전후}}})^2 + (\omega_{\text{민간}\sigma_{\text{공직-민간}}})^2 + 2(\omega_{\text{공직}\sigma_{\text{전후}}})(\omega_{\text{민간}\sigma_{\text{공직-민간}})}}{CDF \sum \omega_n \sigma_n}$$

$$\sum PDF_n + SE = CDF \rightarrow \frac{\sum \sqrt{(\omega_{\text{공직}\sigma_{\text{전후}}})^2 + (\omega_{\text{민간}\sigma_{\text{공직-민간}}})^2 + 2(\omega_{\text{공직}\sigma_{\text{전후}}})(\omega_{\text{민간}\sigma_{\text{공직-민간}})}}{\sum \omega_n \sigma_n}$$

$$Risk Parity = \frac{\sqrt{(\omega_{\text{공직}\sigma_{\text{전후}}})^2 + (\omega_{\text{민간}\sigma_{\text{공직-민간}}})^2 + 2(\omega_{\text{공직}\sigma_{\text{전후}}})(\omega_{\text{민간}\sigma_{\text{공직-민간}})}}{\sum \omega_n \sigma_n}$$

또한 포트폴리오 선택이론(Markowitz, 1952)에 따라 균형위험도를 이용하여 위험비용이나 위험수익(소위 “프리미엄”)을 예측(7)할 수 있음

$$(7) - \text{OUTPUT} = \text{INPUT} \times Risk Parity$$

나아가, 여기서 향후 정부가 관피아 비중을 위험분산 대책(5의 델타헤지)에 따른 예측변수로 설정할 경우 산식은 (8)과 같이 표현할 수 있음.

$$(8) - \omega_{\text{미래}} = \omega_{\text{현재}} + \frac{1}{2} \Delta Risk Parity^2, \Delta = \Phi(d_1)$$

□ 검증 의견

출신여부와 동분산성을 갖는 다른 독립변수가 발견되지 않는 이상, 적어도 모피아 출신여부와 나머지 변수들을 종속관계로 일반화하는 것은 신뢰도 95%수준에서 설명력을 가지므로 일견 타당함.



□ 검정 결과값

*불완전 베타계수=0.001, 신뢰구간 95%수준

중요변수 (IV 독립, H 설명)	비중편차 $w\sigma$	회귀분석 R^2	잔차 χ^2	유의확률 p	상수 α	확률밀도함수 PDF	누적분포함수 CDF	균위위험도 Risk Parity
IV-H1. 기획재정부	2.736	1401695	45331	0.032	0.785	0.023105009	0.996241798	0.3729
IV-H2. 대통령실	0.767	1239577	2580	0.002	0.728	0.020442632	0.881445393	0.1567
IV-H3. 교육부	0.627	1222771	1155	0.001	0.719	0.020164709	0.869461915	0.1131
H4. 검찰청	0.192	1186478	5	0.000	0.703	0.019568687	0.843762653	0.0261
H5. 국토교통부	0.165	1185631	10	0.000	0.703	0.019556784	0.843249421	0.0419
H6. 국무조정실	0.172	1185631	10	0.000	0.703	0.019556592	0.843241145	0.0334
H7. 과학기술정보통신부	0.072	1184078	22	0.000	0.704	0.019529853	0.842088192	0.1124
H8. 산업통상자원부	0.000	1182948	34	0.000	0.705	0.019510890	0.841270541	0.0307
H.9 국방부	0.050	1182665	37	0.000	0.704	0.019505952	0.841057627	0.0068
H.10 보건복지부	0.026	1182524	39	0.000	0.704	0.019504725	0.841004747	0.0198
H.11 해양수산부	0.025	1182524	39	0.000	0.704	0.019504355	0.840988796	0.0093
H.12. 금융위원회	0.037	1182524	39	0.000	0.704	0.019503886	0.840968540	0.0061
H.13 문화체육관광부	0.017	1182383	41	0.000	0.704	0.019502544	0.840910695	0.0345
H.14 외교부	0.022	1182383	41	0.000	0.704	0.019502102	0.840891654	0.0042
H.15 국가보훈처	0.023	1182383	41	0.000	0.705	0.019502096	0.840891374	0.0036
H.16 농림축산식품부	0.000	1182383	41	0.000	0.704	0.019501671	0.840873066	0.0045
H.17 중소벤처기업부	0.010	1182383	41	0.000	0.704	0.019501570	0.840868690	0.0030
H.18 환경부	0.010	1182383	41	0.000	0.704	0.019501570	0.840868690	0.0020
H.19 법무부	0.017	1182383	41	0.000	0.704	0.019501570	0.840868690	0.0023
H.20 고용노동부	0.007	1182383	41	0.000	0.704	0.019501317	0.840857767	0.0046
식품의약품안전처	0.006	1182383	41	0.000	0.704	0.019501219	0.840853565	0.0014
행정안전부	0.006	1182383	41	0.000	0.704	0.019501219	0.840853565	0.0014
특허청	0.000	1182383	41	0.000	0.705	0.019501184	0.840852060	0.0011
산림청	0.000	1182242	43	0.000	0.704	0.019501126	0.840849539	0.0005
방위사업청	0.005	1182242	43	0.000	0.704	0.019501100	0.840848418	0.0007
통일부	0.003	1182242	43	0.000	0.704	0.019501083	0.840847683	0.0006
경찰청	0.003	1182242	43	0.000	0.704	0.019501083	0.840847683	0.0005
인사혁신처	0.003	1182242	43	0.000	0.704	0.019501083	0.840847683	0.0005
여성가족부	0.000	1182242	43	0.000	0.704	0.019501048	0.840846177	0.0027
기상청	0.000	1182242	43	0.000	0.704	0.019501048	0.840846177	0.0000
원자력안전위원회	0.000	1182242	43	0.000	0.704	0.019501048	0.840846177	0.0000
공정거래위원회	0.000	1182242	43	0.000	0.704	0.019501048	0.840846177	0.0000
농촌진흥청	0.000	1182242	43	0.000	0.704	0.019501048	0.840846177	0.0000
소방청	0.000	1182242	43	0.000	0.704	0.019501048	0.840846177	0.0000
관세청	0.002	1182242	43	0.000	0.704	0.019501044	0.840846002	0.0003
국가정보원	0.002	1182242	43	0.000	0.704	0.019501044	0.840846002	0.0003
문화재청	0.000	1182242	43	0.000	0.704	0.019501028	0.840845337	0.0000
국무총리비서실	0.000	1182242	43	0.000	0.704	0.019501028	0.840845337	0.0000
국세청	0.000	1182242	43	0.000	0.704	0.019501028	0.840845337	0.0000
해양경찰청	0.000	1182242	43	0.000	0.704	0.019501028	0.840845337	0.0000
개인정보보호위원회	0.001	1182242	43	0.000	0.704	0.019501024	0.840845162	0.0001
국무총리	0.001	1182242	43	0.000	0.704	0.019501024	0.840845162	0.0001
법제처	0.001	1182242	43	0.000	0.704	0.019501024	0.840845162	0.0001
병무청	0.001	1182242	43	0.000	0.704	0.019501024	0.840845162	0.0001
새만금개발청	0.001	1182242	43	0.000	0.704	0.019501024	0.840845162	0.0001
조달청	0.001	1182242	43	0.000	0.704	0.019501024	0.840845162	0.0001
통계청	0.001	1182242	43	0.000	0.704	0.019501024	0.840845162	0.0001
방송통신위원회	0.000	1182242	43	0.000	0.704	0.019501022	0.840845057	0.0007
국민권익위원회	0.000	1182242	43	0.000	0.704	0.019501022	0.840845057	0.0002
질병관리청	0.000	1182242	43	0.000	0.704	0.019501022	0.840845057	0.0002
행정중심복합도시건설청	0.000	1182242	43	0.000	0.704	0.019501022	0.840845057	0.0002
합계	5.010	60628030	50952	0.037	36.039	1.000	43.117	1.000
평균	0.098	1188785	999	0.001	0.707	0.020	0.845	0.020