JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 27 No. 6, pp. 167-174, June 2022
https://doi.org/10.9708/jksci.2022.27.06.167

Testing the Equality of Several Correlation Coefficients by Permutation Method

Yonghwan Um*

*Professor, Dept. of Industrial and Management Engineering, Sungkyul University, Anyang, Korea

[Abstract]

In this paper we investigate the permutation test for the equality of correlation coefficients in several independent populations. Permutation test is a non-parametric testing methodology based upon the exchangeability of observations. Exchangeability is a generalization of the concept of independent, identically distributed random variables. Using permutation method, we may construct asymptotically exact test. This method is asymptotically as powerful as standard parametric tests and is a valuable tool when the sample sizes are small and normality assumption cannot be met. We first review existing parametric approaches to test the equality of correlation coefficients and compare them with the permutation test. At the end, all the approaches are illustrated using Iris data example.

▸ **Key words**: permutation test, equality of correlation coefficients, non-parametric testing, exchangeability, normality assumption

[요 약]

본 논문에서는 여러 개의 독립적인 모집단들 사이에서 상관계수들의 등가성에 대한 퍼뮤테이션 검정을 조사한다. 퍼뮤테이션 검정은 관측값들의 상호교환성에 기초하는 비모수적인 검정 방법이며 상호교환성이란 독립적이고 동일한 확률변수들의 개념을 일반화한 개념이다. 퍼뮤테이션 검정을 사용함으로써 근사적으로 정확한 검정에 가까운 검정을 실시할 수 있다. 퍼뮤테이션 검정은 근사적으로 보수적인 검정만큼의 검정력을 지니며, 표본의 크기가 작거나 정규성 가정이 충족되지 않을 때 유용한 방법이다. 본 논문에서는 먼저 상관계수들의 등가성을 검정하는 모수적인 방법들을 소개하고 이들을 퍼뮤테이션 검정과 비교한다. 끝으로 모든 검정들은 Iris 데이터를 예를 들어 비교된다.

▸ **주제어**: 퍼뮤테이션 검정, 상관계수들의 등가성, 비모수 검정, 상호교환성, 정규성 가정

• First Author: Yonghwan Um, Corresponding Author: Yonghwan Um
*Yonghwan Um (uyh@sungkyul.ac.kr). Dept. of Industrial and Management Engineering, Sungkyul University
• Received: 2022. 05. 12, Revised: 2022. 06. 10, Accepted: 2022. 06. 10.

http://www.ksci.re.kr pISSN:1598-849X | eISSN:2383-9945

I. Introduction

상관계수는 두 개의 연속형 확률변수들 사이의 상호 연관성을 나타내는 척도로서 다양한 연구분야의 통계추론에서 중요한 역할을 하는 통계치이다. 특히 피어슨 상관계수(Pearson's correlation Coefficient)와 스피어만 순위 상관계수(Spearman's rank Correlation Coefficient)가 일반적으로 널리 사용되고 있으며, 크기가 n인 이변량 데이터 $(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)$ 에서 피어슨 상관계수(γ)는 다음과 같이 정의된다.

$$\gamma = \frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\overline{y})^2}},$$

여기서 $\overline{x}=\sum_{i=1}^{n}x_i/n$, $\overline{y}=\sum_{i=1}^{n}y_i/n$ 이다. 데이터로부터 계산된 피어슨 상관계수(γ)는 모집단 상관계수(ρ)의 추정치이고 두 변수 X와 Y가 모두 거리 또는 비례척도이며 정규분포를 따른다는 가정하에서 계산된다. 이 때 두 변수간 상관성이 없다는 귀무가설($H_0:\rho=0$)을 검정할 때 사용되는 통계량은 자유도 n-2의 t분포를 따르는 t = $\gamma\sqrt{n-1}/\sqrt{1-\gamma^2}$ 이고, 더 일반적인 귀무가설 $H_0:\rho=\rho_0$에 대해서는 Fisher의 z변환 z = $\frac{1}{2}\ln\frac{1+\gamma}{1-\gamma}$ 에 근거한 검정통계량을 사용한다. 두 변수의 분포가 정규분포가 아닐 때는 데이터의 순위를 이용한 스피어만 순위 상관계수(γ_s)를 사용하며 귀무가설을 검정하기 위해 근사적으로 표준정규분포를 따르는 통계량 z = $\gamma_s\sqrt{n-1}$ 을 사용한다.

$$\gamma_s = 1-\frac{6\sum_{i=1}^{n}d_i^2}{n(n^2-1)},$$

여기서 d_i는 x값들의 순위와 y값들의 순위 사이의 차이이며, 동일한 데이터값이 존재할 때는 해당되는 순위들의 평균치를 각각 순위로 정한다.

본 논문에서는 여러 개의 독립 표본으로부터 얻은 상관계수들이 서로 동일한지의 여부를 검정하고 이 검정방법들을 비교하고자 한다. 실제로 이 문제는 여러 상황에서 많이 발생하는데, 예를 들어 학생들의 어학능력과 수리능력 사이의 상관성이 학년에 따라 차이가 있는지 또는 혈압(diastolic blood pressure)과 체중 사이의 상관성이나 두 개의 혈압(diastolic blood pressure 와 systolic blood pressure)사이의 상관성이 연령대별(20대, 30대, 40대 등)로 동일한지를 검정하는 것이다. 지금까지 두 개 또는 여러 개의 상관계수들 사이의 등가성을 검정하는 여러 방법들이 연구되어 왔다. David는 두 개의 상관계수의 등가성 검정을 위해 Fisher의 z변환(z-transformation)을 이용하였고, Krishnamoorthy와 Xia는 일반화된 변수 연구법(generalized variable approach)과 Olkin과 Finn이 제안한 두 상관계수간의 차이를 근사적으로 구하는 방법을 사용하였다[1-3]. David가 가설($H_0:\rho_1=\rho_2$ vs. $H_1:\rho_1\neq\rho_2$)을 검정하기 위해 제안한 통계량 T는 다음과 같다.

$$T = \frac{z_1-z_2}{\sqrt{1/(n_1-3)+1/(n_2-3)}},$$

여기서 $z_i = \frac{1}{2}\ln\frac{1+\gamma_i}{1-\gamma_i}$ 이고 T는 데이터가 정규분포를 따른다는 가정하에서 근사적으로 표준정규분포를 따른다.

또한 세 개 이상의 상관계수들 사이의 등가성을 검정하기 ($H_0:\rho_1=\rho_2=\cdots\rho_k$ vs. $H_1:$적어도 한개의 상관계수는 다르다) 위한 다양한 연구가 Pearson, David, Kraemer, Donner와 Rosner, Paul, Jafari와 Kazemi, Liu와 Ma 등에 의해 진행되어 왔다[4-10]. David는 Fisher의 z변환에 기초하여 검정을 소개하였고, Kraemer는 t분포를 이용한 근사적인 방법을 제안하면서 David가 사용한 Fisher의 z변환 검정과 비교하였다. Donner와 Rosner는 우도비에 의한 가설검정(likelihood ratio test)을 제안하였고, Paul은 우도비 검정과 다른 두 개의 근사적인 통계량을 제시함과 동시에 이들을 Fisher의 z변환 검정과 비교하였다. 또한 Jafari와 Kazemi는 Olkin과 Finn의 연구법을 상관계수가 여러 개일 때로 확장한 통계량을 제시하였으며, Liu와 Ma는 우도비 검정, Wald-type 검정, Score 검정을 비교하여 Score 검정이 제1종 오류율과 검정력에서 더 우월한 검정임을 보였다.

본 연구에서는 Fisher가 처음 소개한 퍼뮤테이션 검정(permutation test)을 사용하여 여러 개의 독립적인 상관계수들 사이의 등가성을 검정한다. 퍼뮤테이션 검정은

데이터의 상호교환성에 근거한 비모수적인 방법으로서 그 기본원리에 기초하여 다양한 연구(회귀모형들의 등가성, 알파계수 또는 반분검사 신뢰도 등의 신뢰척도들의 등가성, 메타분석의 한 방법인 p값 통합 등)가 진행되어 왔다[11-14]. 특히 표본의 크기가 작거나 정규성 가정을 충족할 수 없는 연구에서 널리 사용할 수 있는 장점을 갖고 있다. 본 논문에서는 퍼뮤테이션에 의한 상관계수들의 등가성 검정을 Fisher의 z변환 검정, Paul의 연구법, Jafari와 Kazemi의 연구법, 우도비 검정(likelihood ratio test)과 비교하고 실제 데이터에 적용해본다.

II. Testing the equality of k correlation coefficients

이변량 데이터 (x_{ij}, y_{ij}), (i=1, , , k 그리고 j=1, , , n_i)를 평균이 $\boldsymbol{\mu}_i = (\mu_{1i}, \mu_{2i})$이고 분산-공분산 행렬이 Σ_i인 k개의 독립적인 이변량 정규분포로부터 수집된 임의표본이라 하자.

$$\Sigma_i = \begin{pmatrix} \sigma_{1i}^2 & \rho_i \sigma_{1i} \sigma_{2i} \\ \rho_i \sigma_{1i} \sigma_{2i} & \sigma_{2i}^2 \end{pmatrix}$$

그리고 k개의 독립적인 상관계수들의 등가성에 대한 귀무가설이 다음과 같을 때,

$$H_0 : \rho_1 = \rho_2 = \cdots \rho_k = \rho \qquad (1)$$

(여기서 ρ는 공통 상관계수)

이를 검정하기 위해 Fisher의 z변환 검정, Paul의 연구법, Jafari와 Kazemi의 연구법, 우도비 검정 등을 사용한다.

1. Fisher's z transformation

Fisher의 z변환(=z_i)은 평균이 $\tanh^{-1}(\rho_i)$이고 분산이 $1/(n_i - 3)$인 정규분포를 따른 것은 잘 알려진 사실이다.

$$z_i = \frac{1}{2} \ln \frac{1+\gamma_i}{1-\gamma_i} = \tanh^{-1}(\gamma_i)$$

따라서 H_0을 검정하기 위한 검정통계량(F_z)은 다음과 같이 주어진다.

$$FZ = \sum_{i=1}^{k} \left(\frac{z_i - \bar{z}}{1/\sqrt{n_i - 3}} \right)^2 = \sum_{i=1}^{k} (n_1 - 3)(z_i - \bar{z})^2 \qquad (2)$$

여기서 $\bar{z} = \sum_{i=1}^{k} (n_i - 3) z_i \,/\, \sum_{i=1}^{k} (n_i - 3)$이다. 검정통계량 FZ는 근사적으로 자유도가 k-1인 카이제곱분포를 따르며 FZ > $\chi^{2(k-1),\alpha}$ 일 때 귀무가설을 기각하며, $\chi^{2(k-1),\alpha}$은 자유도가 k-1인 카이제곱분포의 α번째 상위 분위수(αth upper-quantile)이다.

2. Paul's approach

Paul이 귀무가설 (1)을 검정하기 위해 제안한 통계량(PA)은 다음과 같다.

$$PA = \sum_{i=1}^{k} \frac{n_i (\gamma_i - \gamma_F)^2}{(1 - \gamma_F \gamma_i)^2} \qquad (3)$$

여기서 $\gamma_F = \tanh(\bar{z}) = \dfrac{e^{2\bar{z}} - 1}{e^{2\bar{z}} + 1}$은 공통 상관계수 ρ의 합동 추정량이다(pooled estimate). 통계량 PA는 자유도 k-1의 카이제곱분포를 따르며 PA > $\chi^{2(k-1),\alpha}$ 일 때 귀무가설을 기각한다. 이 통계량은 Neyman에 의해 개발된 일련의 검정들에 기초하고 있는데, 이 일련의 검정들은 가설검정에서 관심대상은 아니지만 분석할 때 고려해야하는 모수(nuisance paramters)가 있을 때 적용될 수 있는 검정들이다[15].

3. Jafari and Kazemi Approach

Jafari and Kazemi는 Olkin과 Finn이 제안한 두 개의 상관계수의 차이에 대한 연구법과 Krishnamoorthy와 Xia의 연구방법을 k개의 상관계수들의 등가성 문제에 확장하였다. Krishnamoorthy와 Xia는 Olkin과 Finn의 연구법을 사용하여 가설검정을 실시하고 $\rho_1 - \rho_2$에 대한 신뢰구간을 구축하였고, 결론적으로 근사적으로 표준정규분포를 따르는 통계량 OF를 제안하였다.

$$OF = \frac{(\gamma_1 - \gamma_2) - (\rho_1 - \rho_2)}{\sqrt{(1-\gamma_1^2)^2/n_1 + (1-\gamma_2^2)^2/n_1}}$$

그리고 Jafari와 Kazemi는 k개의 상관계수들의 등가성을 검정하기 위해 다음의 통계량 OF_k를 제안하였다.

$$JK = \sum_{i=1}^{k} \frac{n_i(\gamma_1 - \bar{\gamma})^2}{(1-\gamma_i^2)^2}, \qquad (4)$$

여기서

$\bar{\gamma} = \sum_{i=1}^{k} n_i \gamma_i (1-\gamma_i^2)^{-2} / \sum_{j=1}^{k} n_j (1-\gamma_j^2)^{-2}$ 이다.

통계량 JK는 귀무가설 하에서 근사적으로 자유도가 k-1인 카이제곱분포를 따르며 JK > $\chi^2_{(k-1),\alpha}$ 일 때 귀무가설을 기각한다.

4. likelihood ratio test

귀무가설 (1)을 검정하기 위해 제안된 우도비 검정의 통계량(LR)은 다음과 같다.

$$LR = \sum_{i=1}^{k} n_i \log\left(\frac{(1-\tilde{\rho}\gamma_i)^2}{(1-\gamma_i)^2(1-\tilde{\rho}^2)} \right), \qquad (5)$$

여기서 $\tilde{\rho}$ 는 공통 상관계수 ρ의 최대우도 추정치(maximum likelihood estimate)이며 다음의 식을 반복적으로 풀어가면서 해를 구한다.

$$\sum_{i=1}^{k} \frac{n_i(\gamma_i - \tilde{\rho})}{(1-\gamma_i \tilde{\rho})} = 0. \qquad (6)$$

통계량 LR은 귀무가설 하에서 근사적으로 자유도가 k-1인 카이제곱분포를 따르며 LR > $\chi^2_{(k-1),\alpha}$ 일 때 귀무가설을 기각한다.

III Permutation

확률분포가 알려져 있지 않을 때 정확한 유의성 검정을 찾는 것이 통계추론에서 중요한 문제이며 Fisher는 이 문제를 풀기 위해 퍼뮤테이션 방법(permutation method)을 제시한 바 있다[16]. 이 방법의 특징은 데이터의 상호교환성에 기초하여 관측된 데이터의 모든 배열(arrangement)의 하나하나를 고려하는 것이며 각각의 배열은 귀무가설 하에서 동일한 가능성을 갖고 발생할 수 있다는 것이다. 달리 말하면 임의표본인 이변량 데이터 (x_j, y_j), (j=1, , , n)에서 벡터 **a**=$(a_1, a_2, \ldots, a_n)$를 (1,2, . . .,n)의 어느 한 퍼뮤테이션이라 할 때 각기 다른 배열이 n!개 존재하게 되는데 만일 두 변수 간에 상관성이 없다는 귀무가설이 사실이라면 y_j값들의 배열순서는 상관계수에 영향을 주지 않게 되며 벡터 **a**는 1/n!의 동일한 확률로 발생한다는 것이다. 따라서 모든 배열의 각각에 대해 통계량을 계산하고, 이 통계량들 중에서 원래 관측된 데이터로부터 얻은 통계량의 값 보다 극한의 값을 갖는 통계량들의 비율을 산출함으로써 p값을 구하고 정확한 검정이 가능해진다. 이 p값들은 관측된 데이터에만 의존하기 때문에 퍼뮤테이션 검정은 모집단의 분포에 대해 어떤 가정도 요구하지 않는 distribution-free 검정이 된다.

표본의 크기가 각각 n_1, n_2, , , , n_k인 k개의 이변량 데이터 (x_{ij}, y_{ij}), (i=1, , , k 그리고 j=1, , , n_i)에서 k개의 독립적인 상관계수들의 등가성을 검정을 실시한다고 하자. 이 때 모든 배열의 수는

M = $(n_1 + n_2 + \cdots + n_k)!$ 이고
p값 = (관측된 통계량 값 보다 극한의 값을 갖는 통계량들의 수) / M

이 된다. 그러나 이 값은 k와 n_i(i=1, , , k)값이 커짐에 따라 매우 큰 무한대에 가까운 수가 되므로 실제적으로 정확한(exact) 퍼뮤테이션 검정을 실시하는 것이 불가능하기 때문에 이 배열들 중에서 일부(L개, 일반적으로 L=1,000,000)를 비복원 임의 추출하여 근사적으로 p값을 계산하는 몬테칼로 재표본(Monte Carlo resampling) 퍼뮤테이션 검정을 사용한다. Mielke와 Berry는 재표본 퍼뮤테이션 검정의 결과와 모든 배열을 전부 사용하는 정확한 퍼뮤테이션 검정결과는 거의 일치한다고 연구된 바 있어 M이 클 때 재표본 퍼뮤테이션 검정은 좋은 대안이 될 수 있다[17]. 추출한 L개의 각 배열에 대해서 k개의 상관계수(γ_i),(i=1, , , k)를 먼저 계산한 후 수식 (2),(3),(4),(5)에서 통계량 FZ, PA, JK, LR를 구하고, 원래의 관측 데이터로부터 계산된 통계량 FZ_0(또는 PA_0, JK_0, LR_0) 보다 극한의 값을 갖는 통계량들의 비율로부터 p값을 각각 계산한다.

p값 = (FZ_0 보다 극한의 값을 갖는 FZ값들의 수) / L

= ($|FZ| \geq |FZ_0|$ 을 만족하는 FZ값들의 수) / L

p값 = (PA_0 보다 극한의 값을 갖는 PA값들의 수) / L

= ($|PA| \geq |PA_0|$ 을 만족하는 PA값들의 수) / L

p값 = (JK_0 보다 극한의 값을 갖는 JK값들의 수) / L

= ($|JK| \geq |JK_0|$ 을 만족하는 JK값들의 수) / L

p값 = (LR_0 보다 극한의 값을 갖는 LR값들의 수) / L

= ($|LR| \geq |LR_0|$ 을 만족하는 LR값들의 수) / L

본 논문에서는 데이터의 배열들을 생성하고, 통계량 계산 및 퍼뮤테이션 p값 등의 계산을 위해 R프로그램을 이용하였으며 특히 통계량 LR을 산출하는 과정에서 $\tilde{\rho}$ (ρ의 최대우도 추정치)를 구하기 위해 R함수 optim()을 사용하였다.

IV Example

여러 개의 독립적인 상관계수들 간의 등가성 검정을 예시하기 위해 Table 1의 붓꽃(iris) 데이터를 사용하였다[18]. 이 데이터는 붓꽃의 종류별 (setosa, versicolor, verginica)로 꽃받침의 길이(sepal length), 꽃받침의 너비(sepal width), 꽃잎의 길이(petal length)와 꽃잎의 너비(petal width)를 측정한 것으로 150개(각 종류별로 50개씩)의 레코드로 구성된다(k=3, n_1= n_2 = n_3= 50). 따라서 3개의 그룹과[(1) setosa, (2) versicolor, (3) verginica] 4개의 변수들(1. sepal length, 2. sepal width, 3. petal length, 4. petal width)에서 각 그룹별로 고려할 수 있는 상관계수의 수는 총 6개가 되고 각각의 상관계수가 세 그룹들 간에 동일한지를 검정하였다. 예를 들어

H_0 : $\rho_{1,2}^{(1)} = \rho_{1,2}^{(2)} = \rho_{1,2}^{(3)}$ 은 변수1(꽃받침의 길이)과 변수2(꽃받침의 너비)간의 상관계수가 세 그룹에서 동일하다는 귀무가설이다. Table2는 각 그룹별로 6개의 변수 조합에서의 상관계수이며 모두 양의 상관성을 보이고 있다. Table 3은 각 조합별로 수식 (2),(3),(4),(5)에 의해 계산된 FZ, PA, JK, LR값과 p값들이다. 예를 들어 변수1과 변수2 사이의 상관계수들은 각각 γ_1=0.743, γ_2=0.523, γ_3=0.457 이므로 Fisher의 z변환에 의해 z_1=0.956, z_2=0.584, z_3=0.494와 $\bar{z}$ =0.678로 계산되어 수식 (2)에서 FZ = 5.6416이 되고, 수식 (3)에서는 γ_F = 0.590이므로 PA = 5.7707이 된다. 또한 수식 (4)에서 $\bar{\gamma}$ = 0.640로 계산되어 JK = 6.5260이 되고, 수식 (6)에서 $\tilde{\rho}$ = 0.589로 추정되어 수식 (5)의 LR = 5.9424이 된다. 귀무가설 $H_0 : \rho_{1,2}^{(1)} = \rho_{1,2}^{(2)} = \rho_{1,2}^{(3)}$에 대한 검정 결과는 JK 경우는 p값=0.03827로 유의하였으나 FZ, PA, LR 검정에서는 p값이 0.05보다 약간 큰 값을 보여 α=0.05에서 유의하지 않았다. 또한 변수1과 변수3, 변수2와 변수4, 변수3과 변수4 에 대한 등가성 검정은 모두 유의한 결과를 나타냈으나, 변수1과 변수4, 변수2와 변수3 에 대한 등가성 검정은 모두 유의하지 않았다. 또한 전반적으로 JK를 이용한 검정의 p값이 FZ, PA, LR에 의한 p값보다 작게 나타났다. Table 4는 퍼뮤테이션 검정에 의해 p값을 산출한 결과이다. 퍼뮤테이션 검정에서 고려해야 할 전체 배열의 수는 M = (50+50+50)!의 무한대에 가까운 큰 수이므로 L=1,000,000개의 재표본 퍼뮤테이션 검정을 실시하였다. 퍼뮤테이션에 의한 검정은 모수 검정(FZ, PA, JK, LR)과 동일한 검정 결과를 보였으며 퍼뮤테이션 검정의 p값은 대응되는 모수검정의 p값과 같거나 약간 큰 값을 나타냈다. (단 변수1과 변수4 간의 상관계수의 등가성 검정에서는 FZ를 이용한 퍼뮤테이션 검정의 p값=0.1881은 FZ 검정의 p값=0.1890보다 약간 작게 나타났음)

Table 1. Iris Data

case no.	sepal length	sepal width	petal length	petal width	species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0,2	setosa
⋮	⋮	⋮	⋮	⋮	⋮
51	7.0	3.2	4.7	1.4	versicolor
52	6,4	3.2	4.5	1.5	versicolor
⋮	⋮	⋮	⋮	⋮	⋮
101	6.3	2.3	6.0	2.5	verginica
102	5.8	2.7	5.1	1.9	verginica
⋮	⋮	⋮	⋮	⋮	⋮
150	5.9	3.0	5.1	1.8	verginica

Table 2. Correlation coefficients for each combination of 4 variables

combination	setosa	versicolor	verginica
var.1 & var.2	0.743	0.523	0.457
var.1 & var.3	0.267	0.754	0.864
var.1 & var.4	0.278	0.546	0.281
var.2 & var.3	0.178	0.561	0.401
var.2 & var.4	0.233	0.664	0.538
var.3 & var.4	0.332	0.787	0.322

Table 3. Test results for the equality of three correlation coefficients

combination	FZ		PA		JK		LR	
	statistic	p-value	statistic	p-value	statistic	p-value	statistic	p-value
var.1 & var.2	5.6416	0.05956	5.7707	0.05584	6.5260	0.03827	5.9424	0.05124
var.1 & var.3	26.3582	0.00	23.5691	0.00	20.2049	0.00	26.8109	0.00
var.1 & var.4	3.3317	0.1890	3.4626	0.1771	3.8839	0.1434	3.5235	0.1717
var.2 & var.3	4.8539	0.0883	4.9911	0.08245	5.2528	0.07234	5.1199	0.07731
var.2 & var.4	7.6574	0.02174	7.7243	0.02102	7.7172	0.02110	8.0377	0.01797
var.3 & var.4	16.3952	0.000275	15.6325	0.000403	19.4277	0.00	16.9499	0.000209

Table 4. Permutation p values for the equality of three correlation coefficients

combination	FZ	PA	JK	LR
var.1 & var.2	0.05993	0.05992	0.05879	0.05993
var.1 & var.3	0.0000	0.0000	0.0000	0.0000
var.1 & var.4	0.1881	0.1881	0.1747	0.1881
var.2 & var.3	0.08852	0.08855	0.09840	0.08851
var.2 & var.4	0.02237	0.02238	0.03702	0.02237
var.3 & var.4	0.000319	0.000319	0.000602	0.00032

Table 5와 Table 6은 표본의 크기가 작을 때 퍼뮤테이션 검정의 효과를 알아보기 위해 붓꽃(iris) 데이터 중 그룹별로 일부 데이터($n_1=n_2=n_3=10$)만을 사용하여 얻은 등가성 검정의 결과이다. 표본의 크기가 작음에도 불구하고 정규성을 가정하여 모수 검정(FZ, PA, JK, LR)을 진행하였다. 총 배열의 수 M=(10+10+10)!은 큰 수이므로 L=1,000,000개의 재표본 퍼뮤테이션 검정을 사용하였다. 변수1과 변수2에 대한 상관계수들의 등가성 검정은 모수 검정(FZ, PA, JK, LR)과 퍼뮤테이션 검정에서 모두 동일하게 유의하였으나 다른 변수 조합에서는 모두 유의하지 않은 결과로 나왔다. 또한 FZ를 이용한 퍼뮤테이션 검정의 p값은 FZ 검정의 p값 보다 작았으나(단 변수1과 변수2를 제외하고) 다른 퍼뮤테이션 p값들은 대응되는 각각의 PA, JK, LR 검정의 p값보다 크게 나타났다.

Table 5. Test results for the equality of three correlation coefficients($n_1=n_2=n_3=10$)

combination	FZ		PA		JK		LR	
	statistic	p-value	statistic	p-value	statistic	p-value	statistic	p-value
var.1 & var.2	7.6202	0.02214	7.9887	0.01842	10.7944	0.00453	9.9946	0.00676
var.1 & var.3	3.7627	0.1524	4.5407	0.1033	5.5022	0.06386	5.1505	0.07614
var.1 & var.4	2.7049	0.2586	3.4232	0.1806	4.8710	0.08756	3.7442	0.1538
var.2 & var.3	0.2400	0.8869	0.33901	0.8441	0.3602	0.8352	0.3419	0.8429
var.2 & var.4	1.5958	0.4503	2.1190	0.3466	1.9051	0.3858	2.2373	0.3267
var.3 & var.4	0.3685	0.8317	0.5173	0.7721	0.5583	0.7564	0.5241	0.7695

Table 6. Permutation p values for the equality of three correlation coefficients($n_1=n_2=n_3=10$)

combination	FZ	PA	JK	LR
var.1 & var.2	0.02405	0.02344	0.07309	0.02413
var.1 & var.3	0.1515	0.1523	0,2109	0.1513
var.1 & var.4	0.2515	0.2512	0.2428	0.2516
var.2 & var.3	0.8845	0.8845	0.8813	0.8845
var.2 & var.4	0.4419	0.4417	0.5366	0.4419
var.3 & var.4	0.8309	0.8309	0.8262	0.8309

V. Conclusion

본 연구는 여러개(k)의 독립적인 상관계수들이 서로 동일한지를 검정하기 위해 비모수 통계방법인 퍼뮤테이션 검정을 소개하고 기존의 모수 검정들과 비교한 것이다. 모수 검정은 분포의 정규성이나 큰 표본을 요구하는 반면에, 퍼뮤테이션 검정은 모집단 분포에 의존하지 않고 표본의 크기가 작을 때에도 사용된다는 장점으로 인해 실제로 연구에서 표본 크기의 한계가 있는 심리학, 교육학, 사회학 등의 사회과학 분야와 의학 등에서 유용한 방법으

로 제안되고 있다.

본 논문에서는 카이제곱 분포를 따르는 FZ, PA, JK, LR에 의한 모수검정과 p값 그리고 이와 비교하기 위해 퍼뮤테이션 검정에 의한 p값을 제시하였고 예제 데이터로 iris(붓꽃) 데이터(k=3, $n_1=n_2=n_3=50$)를 사용하였다. 또한 붓꽃(iris) 데이터의 일부만(k=3, $n_1=n_2=n_3=10$)을 사용하여 표본의 크기에 의해 검정에 끼치는 효과를 살펴보았다. 퍼뮤테이션 검정을 실시할 때는 데이터의 퍼뮤테이션에 의해 생성되는 배열의 수가 매우 크므로 L=1,000,000의 재표본 퍼뮤테이션 검정을 하였다.

붓꽃(iris) 데이터 전체의 경우에, 모수 검정 (FZ, PA, JK, LR)과 퍼뮤테이션 검정은 전반적으로 동일한 검정 결과와 비슷한 크기의 p값을 제시하였다. 모수 검정 FZ의 p값들은 PA, JK, LR의 p값들 보다 크고 JK의 p값이 가장 작게 나타나 FZ가 가장 보수적(conservative)이고 JK가 가장 리버럴(liberal)한 모수 검정이라 말 할 수 있어 Fisher의 z 변환을 이용한 검정이 매우 보수적이라고 하는 연구 결과와 일치한다. 그러나 퍼뮤테이션 검정과 비교했을 때 모수 검정(FZ, PA, JK, LR)의 p값은 대부분 퍼뮤테이션 p값보다 작으므로 퍼뮤테이션 검정이 가장 보수적이라고 말 할 수 있다. 붓꽃(iris) 데이터의 일부만을 사용했을 때에도 모수 검정 (FZ, PA, JK, LR)과 퍼뮤테이션 검정은 전반적으로 동일한 검정 결과와 비슷한 크기의 p값을 나타냈다. 마찬가지로 FZ의 p값들이 PA, JK, LR의 p값들 보다 크므로 가장 보수적 검정이라 할 수 있으며, 퍼뮤테이션 검정에서는 퍼뮤테이션 p값들이 대응되는 PA, JK, LR의 p값들 보다 크고 FZ의 p값 보다는 작으므로 퍼뮤테이션 검정은 PA, JK, LR 보다 보수적이고 FZ 보다는 리버럴한 검정이라 말 할 수 있다. 즉 퍼뮤테이션 검정은 표본의 크기에 관계없이 PA, JK, LR 보다 보수적이나, 작은 표본($n_1=n_2=n_3=10$)에서는 FZ보다는 리버럴한 결과를 보여주고 있다. 그러나 주목할 것은 퍼뮤테이션 검정의 p값과 대응되는 모수 검정(FZ, PA, JK, LR)의 p값 사이의 차이가 ($n_1=n_2=n_3=50$)일 때 보다는 ($n_1=n_2=n_3=10$)일 때 크게 나타난다는 것이다. 예를 들어 변수1과 변수4의 조합에서, PA에 의한 p값의 차이 = 0.2512 - 0.1806 = 0.0706 ($n_1=n_2=n_3=10$ 일 때) > 0.05992-0.05584=0.011 ($n_1=n_2=n_3=50$ 일 때) 이다. 이러한 차이는 작은 표본에서는 모수 검정의 가정을 충족하기 어렵고, 표본의 크기가 작을 때 퍼뮤테이션 검정이 모수 검정에 비해 더 정확한 p값을 제공한다는 연구와 동일한 결과를 나타내는 것이다. 표본의 크기가 작음에도 불구하고 무리하게 정규성을 가정한 통계추론은 부정확 할 수밖에 없기 때문에 분포에 의존하지 않는 (distribution-free) 퍼뮤테이션 방법은 좋은 대안이 될 수 있다. 본 논문의 상관계수들의 등가성 검정에서도 퍼뮤테이션 검정은 유효한 통계방법임을 시사하고 있다. 그러나 본 논문은 여러 개의 상관계수들이 상호 독립적이라는 조건하에서 등가성을 다루고 있다는 한계점을 갖고 있다. 따라서 상관계수들이 서로 비독립적인 관계에 있을 때, 예를 들어 어느 직장에서 직원들의 직장 만족도와 연봉의 관계 ($\rho_{직장만족,연봉}$), 직장만족도와 상사에 대한 만족 ($\rho_{직장만족,상사만족}$), 직장만족도와 직장의 복지제도와의 관계($\rho_{직장만족,복지}$) 같이 서로 비독립적인 상관계수들 간의 등가성 검정(H_0 : $\rho_{직장만족,연봉} = \rho_{직장만족,상사만족} = \rho_{직장만족,복지}$) 에서도 퍼뮤테이션 방법을 사용할 수 있는지 그리고 기존의 모수적인 방법에 비해 퍼뮤테이션 검정이 더 유효한 결과를 제시할지는 향후 연구 과제라 하겠다.

REFERENCES

[1] F. N. David, Tables of ordinates and probability Integral of the distribution of the correlation coefficients in small samples, London, Cambridge University Press, 1938.

[2] K. Krishnamoorthy and Y. Xia, Inferences on correlation coefficients: One-sample independent and correlated cases, Journal of Statistical Planning and Inference, 137, 7, 2362-2379, 2007.

[3] I. Olkin and J. D. Finn, Correlation redux, Psychological Bulletin, 118, 155-164, 1995.

[4] K. Pearson, On a method of determining whether a sample of size n supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random, Biometrika, 25, 379-410, 1933.

[5] H. C. Kraemer, On estimation and hypothesis testing problems for correlation coefficients, Psychometrika, 40, 4, 473-485, 1975.

[6] A. Donner and B. Rosner, On inference concerning a common correlation coefficient, Applied Statistics, 29, 69-76, 1980.

[7] S. R. Paul, Estimation of and testing significance for a common correlation coefficient, Communications in Statistics – Theory and Methods, 17, 1, 39-53, 1988.

[8] S. R. Paul, Test for the equality of several correlation coefficients, The Canadian Journak of Statistics, 17, 2, 217-227, 1989.

[9] A. A. Jafari and M. R. Kazemi, Computational approach test for inference about several correlation coefficients: equality and

common, Communications in Statistics – Simulation and Computation, 46, 3, 2043-2056, 2017.
[10] X. Liu, S. Liu and C. X. Ma, Testing equality of correlation coefficients for paired binary data from multiple groups, Journal of Statistical Computation and Simulation, 86, 9, 1686-1696, 2015.
[11] Y. Um, Testing the equality of two linear regression models: Comparison between Chow test and a permutation test, Journal of The Korea Society of Computer and Information, 26, 8, 157-164, 2021.
[12] Y. Um, Permutation est for the equality of several independent Cronbach's alpha coefficients, Journal of The Korea Society of Computer and Information, 24, 6, 159-164, 2019.
[13] Y. Um, Combining independent permutation p-values associated with multi-sample location test data, Journal of The Korea Society of Computer and Information, 25, 7, 175-182, 2020.
[14] Y. Um, Permutation analysis of split-half reliability coefficient, 22, 7, 133-139, 2017.
[15] J. Neyman, Optimal asymptotic tests of composite statistical hypotheses, In: Grenander, V. Q., ed. Probability and Statistics The Harold Cramer Volume. New York: Wiley, p213-234, 1959.
[16] R. A. Fisher, A design of experiment, Oliver & Boyd, Edinburgh, 1935.
[17] P. W. Mielke and K. J. Berry, Permutation methods : A distance function approach, Springer-Verlag, New York. 2001.
[18] R. A. Fisher, The Design of Exprriments, 8th Ed.: Oliver & Boyd: Edinburgh, 1966.

Authors

Yonghwan Um received the B.S. and M.S. in Chemistry from Yonsei University, Korea, in 1981, 1983, respectively, M.S. in Biostatistics from Emory University, in 1990 and Ph.D. in Statistics from University of Florida, U.S.A. in 1995. Dr. Um joined the faculty of the Department of Computational Statistics at Sungkyul University, Anyang, Korea, in 1996. He is currently a Professor in the Department of Industrial and Management Engineering, Sungkyul University. He is interested in reliability measure, data-mining, statistical inference.