Journal of Multivariate Analysis 140 (2015) 60–71

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Rank inversions in the scoring of examinations consisting of several subtests

Simeon M. Berman

Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, United States

ARTICLE INFO

Article history:
Received 10 April 2014
Available online 14 April 2015

AMS subject classifications:
primary 62F07
62H20
secondary 62P15

Keywords:
Correlation matrix
Discordance
Multivariate normal distribution
Ranks
Summary score for subtests

ABSTRACT

A test composed of k subtests is generally scored by assigning weights to the various subtests, defining the summary scores as the weighted sums of the subtest scores, and then ranking the summary scores of the group that was tested. Let $a_1, \ldots, a_k$ represent the weights for the respective subtests and $\xi_1, \ldots, \xi_k$ the subtest scores for a particular individual; then the summary score for this individual is $\sum_{i=1}^k a_i \xi_i$. It has been the practice, particularly in psychological and educational tests, to replace the ith subtest score ξ_i by the standardized subtest score $(\xi_i - \hat{\mu}_i)/s_i$, where $\hat{\mu}_i$ is the average subtest score and s_i is the calculated standard deviation for the group. The use of the standardized subtest scores in the definition of the summary score has the effect of changing the weight a_i to a_i/s_i, $i = 1, \ldots, k$. It is the purpose of this paper to show, under the assumption that the vectors of the subtest scores of those taking the test are i.i.d. normal random vectors in R^k, how the introduction of the standard deviations into the weighted sums affects the set of ranks based on the summary scores. For this purpose we compare the ranks based on the summary scores with weights (a_i/s_i) to those based on the summary scores with weights (a_i/σ_i), where σ_i is the true (unknown) standard deviation of the scores on the ith subtest, $i = 1, \ldots, k$. This represents an extension to a k-part examination of the author's earlier result in the particular case $k = 2$.

1. Introduction

Let $\boldsymbol{\xi}$ be a random vector in R^k representing the scores of an individual on k subtests of a k-part test. The summary score for this individual is represented as a scalar product $\mathbf{a}'\,\boldsymbol{\xi}$, where $\mathbf{a}$ is a vector in R^k whose components (a_i) represent the weights attached to the score ξ_i on the ith subtest. In psychological testing it has been the practice to use the "standard" scores in the place of the "raw" scores (ξ_i) in defining the summary scores: $\mathbf{a}'\,\boldsymbol{\xi} = \sum_{i=1}^k a_i \xi_i$ is replaced by $\sum_{i=1}^k a_i (\xi_i - \mu_i)/\sigma_i$, where $\mu_i = E(\xi_i)$ and $\sigma_i = \textit{standard deviationof } \xi_i$. In practice, (μ_i) and (σ_i) are unknown and are estimated from the sample of scores arising from the group that is taking the test. This sample is represented as n independent copies, $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$ of $\boldsymbol{\xi}$, where $n \geq 1$ is the number in the group. The summary scores are then replaced by the "estimated" summary scores by replacing μ_i and σ_i by their sample estimators $\hat{\mu}_i$ and s_i based on the sample $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$:

$$\hat{\mu}_i = \textit{average of the ith components}$$
$$s_i = \textit{sample standard deviation of the ith components}.$$

Thus the estimated summary score for the typical vector $\boldsymbol{\xi}$ of raw scores is $\sum_{i=1}^k a_i (\xi_i - \hat{\mu}_i)/s_i$.

E-mail address: sberman@cims.nyu.edu.

http://dx.doi.org/10.1016/j.jmva.2015.03.006

The practical effect of replacing the summary scores by the estimated summary scores is discussed in Section 7 in the context of the general procedure of standardizing observations.

The performance of an individual on such a test is generally recorded as the *rank* of the estimated summary score within the group taking the test. The ranks based on the true summary scores are generally different from the ranks based on the estimated summary scores. It is the purpose of this work to provide a statistical measure of the discrepancy between the two systems of rankings.

Let $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ represent an arbitrary pair from the sample of observations $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$, where $\boldsymbol{\xi} = (\xi_i)$ and $\boldsymbol{\eta} = (\eta_i)$.

We will say that the true and estimated summary scores for $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are *discordant* if the order relation between the relative magnitudes of the scores of $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, computed as the true summary scores, is the opposite of that computed as the estimated summary scores:

Either

$$\begin{gathered}\sum_{i=1}^k a_i\,(\xi_i-\mu_i)\Big/\sigma_i > \sum_{i=1}^k a_i\,(\eta_i-\mu_i)\Big/\sigma_i\\ \text{and}\\ \sum_{i=1}^k a_i\left(\xi_i-\hat{\mu}_i\right)\Big/s_i < \sum_{i=1}^k a_i\left(\eta_i-\hat{\mu}_i\right)\Big/s_i\end{gathered} \tag{1.1}$$

or

$$\begin{gathered}\sum_{i=1}^k a_i\left(\xi_i-\hat{\mu}_i\right)\Big/s_i > \sum_{i=1}^k a_i\left(\eta_i-\hat{\mu}_i\right)\Big/s_i\\ \text{and}\\ \sum_{i=1}^k a_i\,(\xi_i-\mu_i)\Big/\sigma_i < \sum_{i=1}^k a_i\,(\eta_i-\mu_i)\Big/\sigma_i.\end{gathered} \tag{1.2}$$

The events described in (1.1) and (1.2) are clearly equivalent to the events

$$\sum_{i=1}^k a_i\,(\xi_i-\eta_i)\Big/s_i < 0 < \sum_{i=1}^k a_i\,(\xi_i-\eta_i)\Big/\sigma_i \tag{1.3}$$

and

$$\sum_{i=1}^k a_i\,(\xi_i-\eta_i)\Big/\sigma_i < 0 < \sum_{i=1}^k a_i\,(\xi_i-\eta_i)\Big/s_i \tag{1.4}$$

respectively.

It is immediately noticeable from (1.3) and (1.4) that the ranks based on the true and estimated summary scores are independent of $\boldsymbol{\mu}$ and $\widehat{\boldsymbol{\mu}}$, respectively; hence we may assume $\boldsymbol{\mu} = \widehat{\boldsymbol{\mu}} = \mathbf{0}$. Furthermore, the probability of discordance (the union of the events (1.3) and (1.4)), calculated under the assumption that $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are normally distributed with common covariance matrix A, may be computed after replacing σ_i by 1, and A by the corresponding correlation matrix. Indeed the joint distribution of the random variables $(\xi_i-\eta_i)/\sigma_i$ and $(\xi_i-\eta_i)/s_i$, $i = 1, \ldots, k$, is invariant under the transformation $\xi_i \to \xi_i/\sigma_i$, $\eta_i \to \eta_i/\sigma_i$, $i = 1, \ldots, k$, and the covariance matrix of the transformed random vectors $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ is identical with the correlation matrix obtained from A. After this transformation the true and estimated summary scores for $\boldsymbol{\xi}$ are of the forms

$$\sum_{i=1}^k a_i\,\xi_i \quad \text{and} \quad \sum_{i=1}^k a_i\,\xi_i/s_i, \tag{1.5}$$

respectively, where s_i is the sample standard deviation when $\sigma_i = 1$, $i = 1, \ldots, k$.

As a measure of the discrepancy between the two sets of ranks based on the true and estimated summary scores we define, for $n \geq 2$,

$$\begin{aligned}\Delta_n = {}& \textit{number of discordant pairs}\\ & (\boldsymbol{\xi}_i, \boldsymbol{\xi}_j), \quad i,j = 1, \ldots, n,\ i \neq j.\end{aligned}$$

Since there are $n(n-1)/2$ pairs $(\boldsymbol{\xi}_i,\ \boldsymbol{\xi}_j)$, and since $(\boldsymbol{\xi}_j)$, $j = 1, \ldots, n$, are assumed to be *i.i.d.*, we have

$$E(\Delta_n) = \frac{1}{2}\,n(n-1)\,P\left(\boldsymbol{\xi}_i \textit{ and } \boldsymbol{\xi}_j \textit{ are discordant}\right), \tag{1.6}$$

for any fixed pair $i,\ j,\ i \neq j$.