Concorso pubblico, per esami, per la copertura a tempo indeterminato e pieno di n. 1 posto di categoria D, posizione economica D1, Area Tecnica, Tecnico - Scientifica ed elaborazione dati – Profilo professionale Esperto Statistico. Il posto è riservato prioritariamente in favore dei componenti delle FF.AA. ai sensi del D.Lgs. 15 marzo 2010, n. 66 (codice concorso: **PTA.Dtec.23.07**).

VERBALE n. 5

La Commissione Esaminatrice (di seguito denominata anche solo "Commissione") della procedura indicata in epigrafe, nominata con D.D. n. 846 del 18/10/2023, si riunisce presso l'Aula Biblioteca in data 11 marzo 2024 alle ore 9.30; la Commissione risulta così composta:

Dott.ssa Teresa ROMEI Direttore Generale Università degli Studi di Foggia - PRESIDENTE;

Prof. Michele RUTA Professore I fascia s.s.d. ING-INF/05 Politecnico di Bari - COMPONENTE;

Dott.ssa Angela Maria D'UGGENTO Ricercatore Universitario s.s.d. SECSS/01 Università degli Studi di Bari-- COMPONENTE;

Dott. Dimitri PATELLA Cat. EP Centro Servizi di Ateneo per la didattica Politecnico di Bari - SEGRETARIO.

Il Presidente, constatata la regolare costituzione della Commissione e la presenza di tutti i Componenti, dichiara aperta la seduta.

Il Presidente ricorda che nel verbale n. 1 della presente procedura concorsuale, sono stati stabiliti i criteri per la valutazione della prova orale.

Vengono predisposte n. 4 schede numerate (essendo 2 i candidati ammessi alla prova orale) contenenti ognuna n. 3 quesiti sugli argomenti riportati nel bando di concorso, di cui uno volto ad accertare le competenze informatiche e un ulteriore foglio (con medesima numerazione) contenente un testo di lingua inglese da leggere e tradurre, ai fini dell'accertamento della conoscenza della lingua.

Ciascuna scheda, siglata da tutti i componenti della Commissione, viene inserita e chiusa in una busta a sua volta siglata sui lembi di chiusura dal Presidente e da tutti i componenti della Commissione e tutte le buste vengono tenute in custodia dal Segretario della Commissione stessa.

Terminate tali operazioni alle ore 10:35, la Commissione consente l'accesso all'aula da parte dei candidati.

I candidati presenti vengono quindi identificati tramite esibizione di un documento di identità e sottoscrizione sul foglio di presenza (allegato n. 1).

La Commissione prende atto che sono presenti n. 2 candidati.

I candidati vengono invitati a spegnere il telefono cellulare, smart watch, orologio o apparecchiature simili e a conservarli, unitamente ad altri oggetti non forniti per lo svolgimento della prova, nelle borse.

Il Presidente avvisa i presenti che la seduta è pubblica e che dovranno uscire dall'aula alla fine dell'esposizione per permettere alla Commissione di assegnare la votazione a porte chiuse.

Alle ore 10:45 la Commissione invita la candidata Porziana CAIATI a sorteggiare la busta contenente i quesiti della prova orale. Risulta estratta la scheda n. 1, di cui all'allegato n. 2 del presente verbale.

Alle ore 10:58 termina la prova, tutti i presenti vengono invitati a uscire dall'aula.

La Commissione valuta la prova orale della candidata e unanime attribuisce il punteggio di 21/30.

La candidata ha superato la prova orale avendo ottenuto un punteggio non inferiore a 21/30.

Alle ore 11.05 la Commissione invita la candidata Vittoria Claudia DE NICOLO' a sorteggiare la busta contenente i quesiti della prova orale. Risulta estratta la scheda n.4, di cui all'allegato n. 3 del presente verbale.

Alle ore 11.20 termina la prova, tutti i presenti vengono invitati a uscire dall'aula.

La Commissione valuta la prova orale della candidata e unanime attribuisce il punteggio di 26/30.

La candidata ha superato la prova orale avendo ottenuto un punteggio non inferiore a 21/30.

Si allegano al presente verbale le schede non estratte contrassegnate con n.2 e n.3 che vengono lette (allegati n. 4 e n. 5)

Alle ore 11.25 la Commissione avendo terminato le operazioni della prova orale redige il seguente riepilogo della votazione ottenuta dai candidati:

Cognome	Nome	Voto prova orale
CAIATI	Porziana	21/30
DE NICOLO'	Vittoria Claudia	26/30

La Commissione, alla luce del punteggio conseguito all'esito della prova orale e riportato nel presente verbale, nonché dei punteggi attribuiti alla prova scritta, redige la seguente tabella riepilogativa:

Cognome	Nome	Voto 1a prova scritta	Voto 2a prova scritta	Voto prova orale	Punteggio totale
CAIATI	Porziana	23	21	21	65
DE NICOLO'	Vittoria Claudia	27	24	26	77

La Commissione predispone la seguente graduatoria finale di merito:

	Cognome e Nome	Valutazione 1a prova scritta	Valutazione 2a prova scritta	Valutazione orale	Totale
1	DE NICOLO' Vittoria Claudia	27	24	26	77
2	CAIATI Porziana	23	21	21	65

La Commissione, dopo la formulazione della graduatoria finale, ritiene conclusi i lavori e trasmette il verbale al Responsabile del procedimento per gli adempimenti di competenza.

La seduta termina alle ore 11:45.

Letto, approvato e sottoscritto.

Dott.ssa Teresa ROMEI – PRESIDENTE

Prof. Michele RUTA – COMPONENTE

Dott.ssa Angela Maria D'UGGENTO – COMPONENTE

Dott. Dimitri PATELLA – SEGRETARIO

Concorso pubblico, per esami, per la copertura a tempo indeterminato e pieno di n. 1 posto di categoria D, posizione economica D1, Area Tecnica, Tecnico - Scientifica ed elaborazione dati – Profilo professionale Esperto Statistico. Il posto è riservato prioritariamente in favore dei componenti delle FF.AA. ai sensi del D.Lgs. 15 marzo 2010, n. 66 (codice concorso: **PTA.Dtec.23.07**)

Foglio presenze prova orale 11/03/2024

	COGNOME	NOME	LUOGO DI NASCITA	DATA DI NASCITA	DOCUMENTO	FIRMA
1	Caiati	Porziana	[redacted]	[redacted]	CI n [redacted] COMUNE DI [redacted] SCAD. [redacted]	[redacted]
3	De Nicolò	Vittoria Claudia	[redacted]	[redacted]3	CI n [redacted] COMUNE D [redacted] SCAD. [redacted]	[redacted]

ALL. 1

All. 2

Scheda n.1

1) Il candidato illustri gli organi di una Università statale ai sensi dell'art. 2, comma 1, della legge 240/2010, soffermandosi, in particolare, sulle funzioni del Direttore.
2) Il candidato illustri un indice per la valutazione della bontà di adattamento di un modello teorico.
3) Costruzione e utilizzo tabelle pivot.

[redacted]

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 27 Number 7, April 2022 ISSN 1531-7714

An Introduction to the Analysis of Ranked Response Data

Holmes Finch
Ball State University

Researchers in many disciplines work with ranking data. This data type is unique in that it is often deterministic in nature (the ranks of items *k*-1 determine the rank of item *k*), and the difference in a pair of rank scores separated by *k* units is equivalent regardless of the actual values of the two ranks in the pair. Given its unique qualities, there are specific statistical analyses and models designed for use with ranking data. The purpose of this manuscript is to demonstrate a strategy for analyzing ranking data from sample description through the modeling of relative ranks and inference regarding differences in ranking patterns between groups. An example dataset of university faculty ratings of job characteristics was used to demonstrate these various methods, and the ways in which they can be tied together to obtain a comprehensive understanding of a ranking dataset. The analyses were carried out using libraries from the R software package, and the code for this purpose is included in the appendix to the manuscript.

Introduction

Ranking data arises from situations in which a finite number of entities, such as sports teams, product brands, political candidates, television programs, or job qualities, are ranked relative to one another. There are many examples of ranking data in an array of academic disciplines, including education (Acuna-Soto, Liern, & Perez-Gladish, 2021) psychology (Regenwetter, et al., 2007), health care (Hackert, et al., 2019; Bothung, et al., 2015; Craig, et al., 2009), quality of life (Peiro-Palomino & Picazo-Tadeo, 2017), sociology (Harakawa, 2021), market research (Kamishima & Akaho, 2006), and political science (Moors & Vermunt, 2007; Gormley & Murphy, 2008). The breadth of these examples demonstrates the great utility of rankings as a tool for understanding human behavior and other scientific phenomena. Throughout this manuscript, the entities being ranked will be referred to as items.

The mechanism for rankings can come in the form of a sample of raters, television viewers, voters, or professional sports competitions. Whichever mechanism is used to rank the items, this type of data share some common qualities. By its very nature, ranking data has a deterministic quality that is not found in most other data situations. Determinism in this context refers to the fact that given the first *k*-1 of *k* rankings, the k^{th} item can only take a specific value. For example, if we know that among a set of 4 tennis players, Novak Djokovic is ranked first, Rafael Nadal second, and Roger Federer third, Andy Murray must be ranked fourth. It should be noted that this deterministic quality is not present if ties are allowed. In that case, it is possible for two or more of the items to have the same rank, and thus the ranking pattern of items *k*-1 does not dictate the ranking of item *k*. In addition to the deterministic nature of the scores, a second signal feature of ranking data is that typically the difference in scores between any pair of items with adjacent rankings is equivalent, regardless of the actual values. For example, the difference between rankings 4 and 5 is equal to the difference between rankings 1 and 2. A third unique quality of ranking data is with respect to their correspondence with the set of

[redacted]

ALL. 3

Scheda n. 4

1) Il candidato illustri il Piano Integrato di Attività e Organizzazione (PIAO).
2) Il candidato illustri le varie tipologie di rapporti statistici e, in particolare, l'uso dei numeri indice
3) Descrivere la funzione logica "SE" di Excel.

[redacted]

4 ALL. 3

We can go through the same steps for the PLMC with highest degree as the covariate. The χ^2_Δ and df_Δ are calculated below.

$$\chi^2_\Delta = 447.69 - 426.93 = 20.56$$

$$df_\Delta = 610 - 600 = 10.$$

The *p*-value for this test statistic is 0.02, based on the following R command.

```
pchisq(q=20.56,df=10,lower.tail=FALSE)
```

The AIC for the model with grad degree was 446.93, which was smaller than the AIC for the model without covariates (457.69). Taken together, these results indicate that the PLMC with highest degree yielded better fit to the data than did the PLM with no covariates. This finding confirms the statistical significance of the relationship between degree and contract length, which was described above. Finally, in order to further investigate the relationships between participant covariates and item worth, a PLT was fit to the data using the `pltree` function from the `PlackettLuce` R library. As described above, the PLT is particularly effective for exploring interactions of the covariates with regard to the item worth parameters. For this example, the PLT model did not find any statistically significant splits with regard to either of the covariates. Therefore, the resulting tree was simply a single node including all of the participants. The worth estimates yielded by the tree were very close to those provided by the PLM as displayed in Table 4.

Synthesis of results

Now that the results from the various analyses have been described, it is important to synthesize them in order to obtain a more complete picture of the rankings considered in this study. Based upon both the raw sample means, the centrality of its position in the UMDS plot, and the PLM worth estimates, it is clear that respondents valued the salary paid by their employer most highly, followed by the health care insurance coverage that they received. They ranked the travel budget as being least important. In addition, the hypothesis tests associated with the PLM revealed that salary was the single most important job quality of those included in this study. In sum, respondents valued salary as the most important job quality, followed by health care coverage, and they valued travel budget least among the traits that they ranked.

The results of the UMDS revealed that respondents who valued salary highly also tended to value health care highly as well. In other words, the two job qualities that were most highly ranked individually were also ranked highly by the same participants. In addition, the UMDS results revealed that rankings of contract terms, chair support, and workload were loosely associated with one another such that higher ranks for one were associated with higher ranks for the others. In contrast, individuals who ranked travel budget more highly tended to give lower ranks to contract terms, chair support, and workload. With respect to qualities of the respondents themselves, the results presented above showed that individuals with a higher terminal degree were more likely to give higher ranks to the terms of the contract. Otherwise, none of the demographic information associated with the respondents was related to their ranking behavior.

Taken together, we can see that the respondents tended to value salary and health care coverage the most, that rankings on these two job qualities were positively correlated with one another, and that between the two salary was significantly more important to the respondents than was health care. In addition, these were the only two job qualities that were likely to be ranked first by most respondents. The rankings of other aspects of the job, including contract terms, chair support, and workload were positively associated with one another, though not as strongly as was the case for salary and health care. Contract term rankings were also positively related to level of the terminal degree of the study respondent. The scores given to travel budget were not related to rankings given to any other job quality, and indeed the travel budget was viewed as the least important from among those included in this study.

Conclusion

The goal of this manuscript was to describe a strategy for analyzing ranking data, and to demonstrate the utilization of that strategy with an existing dataset. Ranking data presents special challenges to researchers, not least because the scores provided by members of the sample are partially deterministic. In other words, when an individual is asked to rank a set of 6 items from most to least favorable, the rankings of the first

DOI: https://doi.org/10.7275/tgkh-qk47

[redacted]

All. 4

Scheda n.2

1) Il candidato illustri le fasi in cui si articola il ciclo di gestione della performance ai sensi del d.lgs. n. 150/2009.
2) Il candidato illustri le caratteristiche della curva di Gauss e il suo utilizzo
3) Le caratteristiche degli Open Data.

ALL. 4

2

permutations of the data. Specifically, common analytic approaches such as histograms or analysis of variance (ANOVA) are not appropriate for use with ranking data because the set of all possible permutations from which the ranks are drawn do not have a natural linear ordering (Fischer, et al., 2019; Alvo & Yu, 2014). Therefore these commonly used techniques will not yield meaningful results and alternative methods, such as those described in this paper, are needed. As described above, all items are ranked by all raters. However, this design is not always used, and in some cases raters are asked to rank only a subset of the *k* items. For example, individuals may be asked to rank their three top candidates for office from a set of 10 in an election. This data structure presents the researcher with unique data analysis challenges, and though interesting, will not be addressed in this manuscript.

Study purpose

The purpose of the current work is to describe and to demonstrate a strategy for analyzing a set of ranking data, from the initial description of the sample through inferential models for characterizing the ranking patterns and investigating relationships between one or more covariates and these patterns. The goal in this demonstration is to provide researchers with a complete example for how to consider ranking data from an analytic perspective, and how to synthesize the results from these multiple techniques in order to gain a full picture of the ranked phenomena being studied. The data analyses include a description of the rankings, as well as model based explorations of the rankings, and investigations of relationships between the rankings and substantively relevant covariates. The example analyses were conducted using the R software package, with an eye to providing the reader with the tools necessary to successfully investigate their own ranking data. Therefore, the R code for conducting these analyses appears in the appendix and the example data are available as supplementary materials to the manuscript.

Sample description

A first step in most data analyses involves an exploration of the sample using descriptive statistics. This is certainly true of ranking data for which we are interested in the mean ranks of the items, as well as the pairwise comparisons of the items and the distribution of ranks for each of the items. The mean rank for item i (m_i) is defined as

$$m_i = \frac{\sum_{j=1}^{t!} n_j v_j(i)}{n} \quad (1)$$

Where

v_j =All possible rankings of the *t* objects

$xxv_j(i)$ =Rank score given to object *i* in ranking *j*

n_j =Observed frequency of ranking *j*

$n = \sum_{j=1}^{t} n_j$

A lower value for m_i indicates that the item is more favored by the members of the sample; i.e., has received a higher ranking with 1 being most favorable. For example, if item 1 has a mean rank of 2.4 and item 2 has a mean rank of 3.9, we would conclude that item 1 was typically ranked higher than item 2.

Another useful description of the sample is the frequency of pairwise comparisons of the item rankings. In other words, how frequently was item A preferred over item B? Table 1 includes a pairwise matrix for a simple example of 3 items that were ranked by 10 individuals. In this example, we can see that item 1 was ranked above item 2 five times, and above item 3 three times. In contrast, item 2 was ranked above item 1 8 times, and above item 3 10 times. Another way in which the rankings can be described is based on the marginal frequency of each rank for each of the items. These results can be presented in a marginal ranking matrix, as in Table 2. For this hypothetical example, item 2 most frequently received a top ranking, followed by item 1, and then item 3. Item 3 was most frequently the lowest ranked.

In addition to describing the sample in terms of central tendency and relative ranking, we may also want to ascertain whether the pattern of rankings is random in nature. One way to do that is to test the null hypothesis that the mean rank is equal to $\frac{(t+1)}{2}$ for *t* ranked items. For the three ranked items, the mean under the null hypothesis of a random ranking would be $\frac{(3+1)}{2} = 2$. In other words, if the rankings provided by the members of the sample had no systematic pattern (i.e., were random in nature), then the mean

https://scholarworks.umass.edu/pare/vol27/iss1/7
DOI: https://doi.org/10.7275/tgkh-qk47

ALL. 5

Scheda n. 3

1) Il candidato illustri i contenuti e il procedimento di adozione/aggiornamento del Sistema di misurazione e valutazione della performance di un Ateneo Statale.
2) Il candidato illustri le principali rappresentazioni grafiche da utilizzarsi in base al tipo di dato e si soffermi sull'istogramma.
3) Come scrivere delle formule in un testo in Word.

respondents. UMDS was fit to the data using the Euclidean distance, as well as the Kendall and Hamming distance measures. Results for all three approaches were quite similar, and only those for the Euclidean distance are reported below. The PLM was fit to the data using the `PlackettLuce` function from the R `PlackettLuce` library (Turner, Kosmidis, Firth, & van Eten, 2021b) in conjunction with the `prefmod` library (Hatzinger & Maier, 2017), with quasi-standard errors for the worth parameter estimates obtained using the `qvcalc` R library (Firth, 2020). The PLMC with both experience and highest degree serving as covariates was fit to the data using the `rol` function from the `pmr` R library. Finally, a PLT was used to investigate the possibility of interactions between highest degree and years of teaching experience in terms of the ranking behavior. This tree model was employed using the `pltree` library from the `PlackettLuce` R library.

Results

Sample description

The mean ranks for the six items appear in Table 5. Salary was the highest ranked job quality on average, followed by health care. The least favored (lowest sample means) items were travel budget and workload. Table 5 also includes the pairwise rank comparisons for the set of items. Recall that these values reflect the number of times that the row item was ranked higher than the column item. For example, Salary was ranked higher than contract by 30 of the 41 study participants. From these results, we can confirm that salary was the most popular (highest ranked) job quality, with pairwise comparison values ranging between 30 and 38 when compared to the other items; i.e., it received a higher rank than each of the other qualities from between 30 and 38 of the study participants. In contrast, travel budget was not ranked higher than any of the other items by a majority of the respondents. It performed best compared to workload, against which it was given a higher rank by 13 individuals.

The marginal frequencies, which appear at the bottom of Table 3, provide more evidence regarding the most and least popular items. Salary received the highest rank 24 times, and the second highest rank 7 times, and was never the lowest ranked item. Health care was the highest ranked item for 4 respondents, and the second highest for an additional 19 respondents. In contrast to these popular items, the travel budget was the least valued by study participants, with 31 of them ranking it either lowest or next to lowest. Workload yielded a bimodal distribution of ranks with 11 individuals placing it third, and 13 placing it fifth.

In order to assess whether the pattern of ranks departed from what we would expect were they completely random, the Chi-square test was used, as described above. The mean rank under the null hypothesis for this calculation was $\frac{t+1}{2} = \frac{6+1}{2} = 3.5$. The Chi-square statistic for this problem was 78.99, with degrees of freedom of 5 (6-1), yielding a *p*-value less than 0.001. Thus, if □ =0.05, we would reject the null hypothesis and conclude that there was a nonrandom pattern to the ranks provided by the participants. In other words, we would conclude that in the population some of the job qualities are ranked as more important than are others.

UMDS

In order to gain insights into how the ranked items are related to one another, UMDS with 2 dimensions was fit to the data using the `smacofRect` function from the `smacof` R library. The plot was created using the `mdpref` function from the `pmr` R library. This model explained approximately 55% of the variance in the rankings. Figure 1 displays the locations of the 6 items and 41 respondents on dimensions 1 and 2. First, we note that salary is most central with respect to the study participants, which reflects that it was the highest ranked of the items by many individuals. In contrast, travel budget and workload lay furthest from the cloud of participant points, which is expected given that they were the lowest ranked items for most raters. The locations of health care, contract and chair support relative to the participants indicates their midlevel rankings as also shown in Table 5.

Based on the distribution of job categories in Figure 1, dimension 1 appears to reflect the contrast between workload and contract, such that those who ranked workload relatively more highly were also more likely to rank contract terms relatively lower. In addition, dimension 1 also reveals that ranks for salary and health care were closely related to one another; i.e., those who ranked salary highly also tended to rank health care highly. The second dimension displays the