

研究員 の眼

審査の差の定量化

審査のブレはどれくらい？

保険研究部 主席研究員 篠原 拓也
(03)3512-1823 tshino@nli-research.co.jp

コンクールやコンテストには、審査がつきものだ。音楽やバレエのコンクールでは、演奏や踊りの技術力や芸術性が審査される。スポーツの体操競技やフィギュアスケート競技では、技の難度や芸術的要素が評価される。漫才のコンテストでは、ネタの面白さ、熟練度、ウケ(親和性)等が審査される。

こうした審査は、演技手のパフォーマンスに対して、複数の審査員が点数を付けて、それをもとに順位を決めることが一般的だ。審査員が点数を付ける際には、審査員間のブレを減らすために、採点基準が細かく設定されていることが多い。だが、演技手のパフォーマンスから受ける印象は審査員ごとに異なるため、審査員の付ける点数には差が生じがちだ。その結果、審査員によって演技手の順位が異なる場合もある。

こうした審査員ごとの審査の差が生じることは、ある程度想定されたものといえる。複数の審査員の眼を通じて、審査を安定、充実させようという狙いだ。ただ、差が大きくなると、審査の一貫性が問われることになりかねない。今回は、審査の差を統計的にどう定量化するべきか、見ていこう。

◇ 3人の演技手のパフォーマンスを2人の審査員が順位付けする場合

バレエコンクールでも漫才コンテストでもよいが、まず、演技手がa、b、cの3人、審査員がsとtの2人、という単純な場合を考えてみる。それぞれの審査員は、3人の演技手のパフォーマンスを見て、1位～3位の順位を付けることとする。このとき、2人の審査員の審査の差は、どのように数字で表せるだろうか？

3人の演技手の順位を[1位, 2位, 3位]の順番に表す(同順位は付けないものとする(以下、同様))とすると、順位付けは、(1) [a, b, c]、(2) [a, c, b]、(3) [b, a, c]、(4) [b, c, a]、(5) [c, a, b]、(6) [c, b, a]の6通りある。この(1)～(6)のうちのどれか1つを、2人の審査員sとtがそれぞれ決めることにな

る。したがって、2人の審査員の順位付けのパターンは、全部で36通り($=6 \times 6$)となる。

この36通りについて、sとtの違いを示す指標をなにか設定したい。ここで、各演技手にsとtが付けた順位を掛け算して、それを合計した指標Vを考えてみる。

例えば、sが(1) [a, b, c]、tが(5) [c, a, b]と順位付けした場合、aについては $1 \times 2 = 2$ 、bについては $2 \times 3 = 6$ 、cについては $3 \times 1 = 3$ となり、それらを合計して、 $V = 2 + 6 + 3 = 11$ といった具合いだ。

この指標Vは、sとtの順位が揃っていると大きくなり、異なっていると小さくなる。したがって、Vを使えば2人の審査員の審査の差を数字で表すことができる。

◇ 演技手の数が4人、5人、…と増えていく場合

この指標Vは、演技手の数が3人よりも増えていっても、順位が揃っていると大きくなり、異なっていると小さくなるという性質は変わらないので、指標として有効に機能しそうだ。

ただし、ここで1つ問題がある。そもそも演技手の数が増えるとVは大きくなるので、Vの大きさだけを見ても審査の差の程度が把握しづらいことだ。

なにか、工夫をする必要がある。そこで考えられるのが、Vのとりうる値の最大値、最小値をもとにVを修正することだ。具体的には、最大値と最小値の中間の値を求めて、それをVから差し引く。そして、その差し引いた値を、最大値と最小値の差で割り算する。さらに、割り算の結果を2倍する。このように修正した指数を、ギリシャ文字の ρ (ロー)で表すことにしよう。

ρ の計算で、最後に2倍するのは、相関係数のように-1~1の値をとるようにするためだ。sとtの順位が完全に一致している場合は $\rho = 1$ 、全く逆に異なっている場合は $\rho = -1$ となる。 ρ の値が1に近いのか、それとも-1に近いかを見ることによって、審査の差の程度が把握できるようになる。

◇ 演技手の数が5人の場合の例で実際に計算してみる

それでは、実際に計算してみる。演技手の数がa、b、c、d、eの5人の場合を例にとってみよう。

sが[a, b, c, d, e]、tが[b, a, d, c, e]と順位を付けたとする。このとき、 $V = 1 \times 2 + 2 \times 1 + 3 \times 4 + 4 \times 3 + 5 \times 5 = 53$ となる。Vの最大値は $55 (= 1 \times 1 + 2 \times 2 + 3 \times 3 + 4 \times 4 + 5 \times 5)$ 、最小値は $35 (= 1 \times 5 + 2 \times 4 + 3 \times 3 + 4 \times 2 + 5 \times 1)$ だから、最大値と最小値の中間の値は45、最大値と最小値の差は20となる。

したがって、 $\rho = (53 - 45) \div 20 \times 2 = 0.8$ と計算できる。 ρ が1に近いプラスの値なので、sとtの

審査は大体揃っている。審査のブレはそれほど大きくないと言える。

◇ ρ の正体は…

ここで、修正した指標 ρ について、考えてみよう。そもそも、なぜ「 ρ (ロー)」という名前にしたのか？ 鋭い読者は「なにかありそうだ」と思ったに違いない。

実は、この ρ は、統計学で「スピアマンの順位相関係数」と言われているものにあたる。なにかのコンテストの順位のように、整数値しかとり得ないようなデータ(離散データ)をもとに、統計的な分析を行う場合によく用いられる指標だ。この指標は ρ という記号で表されることが多い。

それでは、 ρ の一般的な計算式を求めてみよう。演じ手の数が n 人(n は自然数)の場合だ。

V の最大値は、 $1 \times 1 + 2 \times 2 + \dots + n \times n = n(n+1)(2n+1)/6$

V の最小値は、 $1 \times n + 2 \times (n-1) + \dots + n \times 1 = n(n+1)(n+2)/6$ となる。 (※)

(※)

$$\begin{aligned} & 1 \times n + 2 \times (n-1) + \dots + n \times 1 \\ &= 1 \times (n+1) + 2 \times (n+1) + \dots + n \times (n+1) - (1 \times 1 + 2 \times 2 + \dots + n \times n) \\ &= n(n+1)/2 \times (n+1) - n(n+1)(2n+1)/6 \\ &= n(n+1)(n+2)/6 \end{aligned}$$

したがって、最大値と最小値の中間の値は $n(n+1)^2/4$ ($= \{n(n+1)(2n+1)/6 + n(n+1)(n+2)/6\} \div 2$)、最大値と最小値の差は $n(n+1)(n-1)/6$ ($= n(n+1)(2n+1)/6 - n(n+1)(n+2)/6$) となる。

その結果、 ρ の一般的な計算式は、

$$\rho = (V - n(n+1)^2/4) \div \{n(n+1)(n-1)/6\} \times 2 \quad (*)$$

となる。ちょっと、ごちゃごちゃしているが、それほど複雑な計算式ではない。

◇ たまには計算式の意味を振り返ってみる

スピアマンの順位相関係数をインターネット上のサイトで検索すると、次のような形で表示されていることが多い。

$$\rho = 1 - 6 \times \sum D^2 / (n^3 - n)$$

($\sum D^2$ は、 n 人の演じ手について、それぞれ s と t の順位の差を 2 乗して、その合計をとったもの)

これは、もちろん正しいし、計算をするにはこの式のほうが簡単だ。実際に、(*)の式と同じものとなっている。(その確認は、稿末に(参考)として付ける。)

だが、「なぜこんな計算式で表されるのか?」という点については、最終的な形を見るだけでは理解が難しいだろう。

統計学では、実務に用いるために、さまざまな指標や統計量の計算式が簡潔に整備されていることがある。それはそれで、かなり役に立つものだ。

だが、「この計算式は一体どのように導出されたのか?」という、統計的な拠り所が切り離されてしまう恐れもある。

計算式を用いて統計処理を行う際、たまには、その意味を振り返ってみるのもよいかもしれない。

(参考文献)

「推計学のすすめ ― 決定と計画の科学」佐藤信著（講談社，ブルーバックス B-116，1968 年）

“Spearman’s rank correlation coefficient” (Wikipedia)

(参考) $1-6 \times \Sigma D^2 / (n^3-n) = (V - n(n+1)^2/4) \div \{n(n+1)(n-1)/6\} \times 2$ であることの確認

i 番目の演じ手($i=1 \sim n$)に対して、審査員 s が付けた順位を s_i 、審査員 t が付けた順位を t_i とする。V は、演じ手に s と t が付けた順位を掛け算して、それを合計した指標だから、つぎのように表すことができる。

$$V = s_1 \times t_1 + s_2 \times t_2 + \cdots + s_n \times t_n = \Sigma (s_i \times t_i) \quad (\Sigma \text{ は、} i=1 \sim n \text{ に対する和})$$

確認する式の左辺を展開していくと、右辺の形が導出される。

$$\begin{aligned} & 1-6 \times \Sigma D^2 / (n^3-n) \\ &= 1-6 \times \Sigma (s_i - t_i)^2 / (n^3-n) \quad (\Sigma \text{ は、} i=1 \sim n \text{ に対する和}) \\ &= 1-6 \times \{ \Sigma s_i^2 + \Sigma t_i^2 - 2 \Sigma (s_i \times t_i) \} / (n^3-n) \\ &= 1-6 \times \{ n(n+1)(2n+1)/6 + n(n+1)(2n+1)/6 - 2V \} / (n^3-n) \\ &= 1-2 \times n(n+1)(2n+1) / \{n(n+1)(n-1)\} + 12V / \{n(n+1)(n-1)\} \\ &= 1-2 \times (2n+1) / (n-1) + 12V / \{n(n+1)(n-1)\} \\ &= \{ (n-1) - (4n+2) \} / (n-1) + 12V / \{n(n+1)(n-1)\} \\ &= -3 \times (n+1) / (n-1) + 12V / \{n(n+1)(n-1)\} \\ &= -3 \times n(n+1)^2 / \{n(n+1)(n-1)\} + 12V / \{n(n+1)(n-1)\} \\ &= \{ 12V - 3 \times n(n+1)^2 \} / \{n(n+1)(n-1)\} \\ &= (V - n(n+1)^2/4) \div \{n(n+1)(n-1)/6\} \times 2 \end{aligned}$$

本資料記載のデータは各種の情報源から入手・加工したものであり、その正確性と完全性を保証するものではありません。
また、本資料は情報提供が目的であり、記載の意見や予測は、いかなる契約の締結や解約を勧誘するものではありません。