

ZWECKMÄSSIGE BEWERTUNGEN IM SPORT

nach Anne S. Hawkins
Originaltitel in "Teaching Statistics" Vol. 8 (1986) Nr. 1:
Practical Sports Judgments
Bearbeitung: M. Borovcnik

Zusammenfassung: Bewertungssysteme spielen im Sport eine große Rolle. In manchen Sportarten kommt es nicht auf den objektiv feststellbaren besten Wert einer Serie von Einzelleistungen an, sondern man muß aus einer Reihe von subjektiven Bewertungen von Juroren eine einzelne Kennzahl ermitteln, die zur Ermittlung der Rangliste herangezogen wird. Die Konstruktion und Diskussion von solchen Kennzahlen kann ein ertrageiches Projekt der Beschreibenden Statistik darstellen. Dies wird im folgenden am Bewertungssystem Turniertanz demonstriert.

0. Vorbemerkungen

Im Rahmen der Beschreibenden Statistik hat folgende Mathematisierungsaufgabe in der Zwischenzeit einen festen Platz gewonnen: Gegeben sind zwei Datenlisten A und B. Diese gilt es zu vergleichen. Wie geht man vor? Welche Kennziffer soll die jeweilige Datenliste für einen solchen Vergleich repräsentieren? Eingekleidet wird das oft in einen Sportkontext: Zwei Klassen A und B feiern ein Sportfest. Die Datenlisten sind die Klassenergebnisse im Weitsprung. Welche Klasse ist die bessere und soll den Klassenpreis erhalten? Die in Lehrbüchern vorgeschlagenen verschiedenen Mathematisierungen erwecken leicht den Anschein, daß sie völlig aus der Luft gegriffen sind: Modus der Datenlisten z. B. ist so eine Kennziffer. Nun kann gerade der unterrichtliche Reiz darin bestehen, daß Schüler ihre eigenen Regeln zur Ermittlung des Mannschaftssiegers aufstellen dürfen. Wenn sie auf den Geschmack kommen, dann werden sie sicherlich eine Vielfalt von "Regeln" erfinden. Zurück bleibt eventuell der Belgeschmack eines "eitlen Perlenspiels". Anne S. Hawkins überrascht aber im englischen Originalartikel "Practical Sports Judgments" mit einem Beispiel für eine tatsächlich angewendete "abstruse" Regelung zur Ermittlung eines Siegers.

1. Bewertungssysteme - ein Beispiel aus dem Turniertanz

Subjektive Bewertungssysteme im Sport

In verschiedenen Sportarten, z. B. Eiskunstlauf, beim Tauchen oder beim Turnen, müssen Preisrichter die Teilnehmer *subjektiv* bewerten. Die Art und Weise, in der ihre Bewertungen in Zahlen ausgedrückt und weiterhin behandelt werden, unterscheidet sich von Sportart zu Sportart. In jedem dieser Bewertungssysteme jedoch ist es notwendig, die Bewertung mehrerer Juroren in eine einzige zusammenfassende Kennzahl zu verschmelzen, die dann dazu benutzt wird, die Konkurrenten in eine Rangfolge zu bringen. Man hofft, die zusammenfassende Zahl erscheint für die Mehrzahl der Personen in der Mehrzahl der Fälle als fair.

In Sportkreisen gibt es mitunter heftige Debatten, wenn eine spezielle Liste von Einzelbewertungen zu einem "unfairen" Resultat führt. Interessant dabei ist, daß in solchen Debatten Sportanhänger involviert sind, die keineswegs behaupten, Statistiker oder Mathematiker zu sein, die daher Intuitionen und Information über das Problemfeld einbringen müssen, um die tiefliegenden Merkmale der Daten und ihrer Bearbeitung zu diskutieren, die sich also nicht auf offensichtlich "automatisch" ablaufende Methoden zur Analyse, die man in den schlechtesten Beispielen aus Statistik-Rezeptbüchern findet, abstützen können. In dieser Hinsicht können sie mit unseren Schülern verglichen werden (die erst noch zu lernen haben, wie man "kocht").

Bewertungssystem beim Turniertanz

Im folgenden betrachten wir die Welt des Turniertanzes als Beispiel. Die Konkurrenten stellen sich einem Bewertungssystem, das unter dem Namen *Skating System* bekannt ist. Weiter entwickelt durch eine Vielzahl von Jahren, benutzt man in diesem System das Kriterium der Majorität der Präferenzen unter den Rangfolgen der Juroren, um "Fairness" zu gewährleisten. Leider sind Majoritätspräferenzmuster oft alles andere als

klar und es bedarf komplizierter Zusatzregeln, um Konflikte zu lösen und "faire" Resultate abzuleiten. Gelegentlich ist diese "Fairness" suspekt, wie bei den folgenden realen Wettbewerbsergebnissen. Nach dem Prinzip der Majorität der Präferenzen nimmt Paar Nummer 1 den ersten Platz in diesem Finale mit sechs Paaren ein. Da es fünf Juroren gibt, ist drei die Majoritätszahl. Der Einfachheit halber werden nur die Placierungen für die Paare Nummer 1 und 2 angegeben.

Placierungen der Juroren - Rumba

Juror	A	B	C	D	E	Zahl der ersten Plätze
Paar Nr.1	3	4	1	1	1	3 Majorität → 1.
Paar Nr.2	4	1	2	4	2	1 → 2.

Im Jive, einem anderen Tanz desselben Wettbewerbs, hat keines der Paare eine Majorität an ersten Plätzen. Das *Skating System* verlangt in so einem Fall, daß das Paar mit der (größeren) Majorität an *ersten und zweiten* Plätzen den Sieg davonträgt.

Placierungen der Juroren - Jive

Juror	A	B	C	D	E	Zahl der 1./2. Plätze
Paar Nr.1	4	3	1	1	2	3 Majorität → 2.
Paar Nr.2	2	1	2	5	1	4 größere Majorität → 1.

Ermittlung des Turniersiegers - Kombination der Ergebnisse aus Einzeltänzen

Der Bewerb besteht aus zwei Tänzen, die Bewertung für *beide* Tänze müssen zusammengelegt werden. Üblicherweise werden die Ränge addiert, die in jedem einzelnen Tanz erreicht werden, die Endränge werden auf der Basis aufsteigender Summen ermittelt. In diesem Fall jedoch haben die Paare Nr. 1 und 2 je-

weils drei Punkte (1. und 2. bzw. umgekehrt 2. und 1.) und das *Skating System* muß erneut zur Hilfe herangezogen werden, um die Pattstellung aufzulösen. Dies wird so bewerkstelligt:

Man tut so, als ob alle Placierungen von verschiedenen Juroren bei ein und demselben Tanz abgegeben worden wären, d.h., als als ob es 10 Juroren gäbe und die Majorität 6 wäre. Wie immer, werden dann nur Majoritätsplacierungen betrachtet.

Juror	Placierungen										Zahl der 1. Plätze	Zahl der 1./2. Plätze
	Rumba					Jive						
	A	B	C	D	E	A	B	C	D	E		
Paar Nr. 1	3	4	1	1	1	4	3	1	1	2	5	6 → 2.
Paar Nr. 2	4	1	2	4	2	2	1	2	5	1	3	7 → 1.

Nun nimmt das Paar Nr. 2 den ersten Gesamtrang ein, weil es die größere Majorität an ersten und zweiten Placierungen hat, obwohl nach vielen anderen Kriterien Paar Nr. 1 die "besseren" Wertungen aufweist (z. B. mehr erste Placierungen, weniger niedrige Placierungen, mehr Juroren, die es dem Paar Nr. 2 vorziehen etc.).

Hätte im obigen Beispiel Paar Nr. 2 einen ersten oder zweiten Platz weniger bekommen, hätte es auch nach der Zahl der ersten und zweiten Plätze eine Pattstellung geben (beide 6, beides wären Majoritäten). In diesem Fall verlangt das *Skating System*, daß die *Summe* aller Placierungen, die zu den Majoritäten beitragen, in Betracht gezogen wird: Paar Nr. 1 würde dann fünf erste und einen zweiten Platz, also 7 Punkte zusammen haben. Paar Nr. 2 (mit einem ersten oder zweiten Platz weniger als in der obigen Tabelle) würde so auf eine Gesamtsumme von neun oder zehn kommen, was größer ist als bei Paar Nr. 1. Paar Nr. 1 würde nun den ersten Gesamtrang einnehmen, weil es (gemessen an der niedrigeren Gesamtsumme) eine bessere Zusammensetzung an Placierungen für seine Majorität aufzuweisen hat.

Problematik und Sinn von Bewertungssystemen

Das Beispiel weiter oben illustriert eine Vielzahl von Regeln, an die man sich zu halten hat, um zwischen den Konkurrenten zu entscheiden. Im Prinzip sind diese Regeln leicht aufzufassen und relativ leicht anzuwenden (Abzählen oder Summieren von Placierungen bzw. Ordnen). Das Problem liegt darin, daß solch einfache Regeln Anomalien, zweifelhafte Fälle und Widersprüche produzieren, und, wie beim Beispiel des *Skating Systems*, dazu neigen, mit weiteren einfachen Regeln zur Auflösung solcher Schwierigkeiten überladen zu sein. Wie man auch an diesem Beispiel erkennen kann, werden nicht alle Schwierigkeiten auch immer gelöst.

Solche rätselhaften Punktezahlungen ziehen viel Aufmerksamkeit auf sich, besonders wenn Einkommen und Karriere von Berufssportlern auf dem Spiel stehen. Ebenso wie in [2] besteht eine Notwendigkeit, mehrere Daten auf *eine* Kennzahl zu verdichten, sodaß man mit Hilfe dieser Kennzahl verschiedene Datenlisten dem Rang nach anordnen kann.

2. Diskussion von Bewertungssystemen im Unterricht

Erfassung von Fairness - eine offene Mathematisierungsaufgabe

Diese Art von Problemen ist sehr wertvoll und anregend für Schüler und kann in Form eines Projekts bearbeitet werden. Der *intuitiven* Problemerkennung und Bewertung soll dabei ein entsprechender Rang eingeräumt werden. Zum ersten ist das Problem an sich schon interessant wegen der Fragen rund um die Idee der "Fairness", die es aufwirft. Schüler werden herausgefordert, Wege der Datenbearbeitung zu suchen, sodaß das Ergebnis mit den festgelegten Kriterien von "Fairness" in Einklang zu bringen ist. Solche Ideen haben über den Sport hinaus Bedeutung, etwa in wirtschaftlichen Entscheidungen oder in der Meinungsforschung.

Zweitens, Schüler werden ermutigt, Daten kritisch unter die Lupe zu nehmen, z. B., auf intuitiver Ebene die Beziehungen zwischen einzelnen Anordnungen von Placierungen und dem Ge-

samtergebnis zu erforschen. Dies gibt Anlaß zu stimulierenden Vermutungen über erklärende Faktoren in quantitativen zusammenfassenden Kennzahlen. Die Schüler müssen "faire" Methoden der Datenverschmelzung unterstellen, sie müssen nach Gegenargumenten suchen, unterstützt durch alternative Anordnung der Daten, und sie müssen nach zufriedenstellenden Lösungen suchen. Zum Dritten, dieses Thema stellt ein Problem, welches einer *operationalen* Lösung bedarf. Es gibt nicht *eine* richtige Antwort, dies ermutigt Schüler, Verfahren bzw. Kriterien zu erfinden, die überprüfen helfen, ob eine Lösung "gut" ist oder "besser" als eine andere Lösung ist (im Gegensatz zur naiven Einschätzung von "korrekt"). Darüberhinaus müssen sie herausfinden, in welchen Fällen und wie oft die vertretene Lösung zusammenzubrechen droht. Alle diese Denkansätze sind eng verbunden mit kognitiven Fertigkeiten, welche wir bei unseren Statistikschülern entwickeln wollen.

Im allgemeinen werden die Methoden, die ein Statistiker anwenden würde, um das Problem anzupacken, weit jenseits der Fähigkeiten der Schüler sein. Natürlich werden die Schüler jene Techniken anwenden, die sie schon gelernt haben oder die der Lehrer in diesem Zusammenhang einführt. Die wirkliche Schönheit dieses Themengebietes jedoch liegt darin, daß die Schüler sich auf mehr intuitive Methode abstützen müssen - eine gute Sache für sich, wenn man statistisches Denken ausbilden möchte. Es ist ganz wichtig, Schüler von der Vorstellung zu befreien, daß es *eine und nur eine* Technik gibt, eine vorgegebene Datenliste zu behandeln. Wir möchten sie nachdrücklich mit der Idee bekanntmachen, daß man eine durch Kenntnisse abgestützte Wahl zwischen verschiedenen Methoden zu treffen hat.

Grobe Anomalien von Bewertungssystemen - Kriterien für Bewertungssysteme

In diesem Themengebiet ist das beste, was man für einen vorgeschlagenen Lösungsansatz erhoffen kann, daß die angenommene Verschmelzungsmethode die Häufigkeit einiger der schlimmsten Paradoxa und Anomalien reduziert. Im *Skating System* z. B. können folgende drei grobe Anomalien auftreten:

Intransitivität: Man spricht von intransitiven Bewertungssystemen, wenn folgendes passieren kann: Paar Nr. 1 wird Paar Nr. 2 vorgezogen, Paar Nr. 2 wird Paar Nr. 3 vorgezogen, trotzdem wird das Paar Nr. 3 dem Paar Nr. 1 vorgezogen.

Irrelevanz: Von Irrelevanz spricht man, wenn folgendes passieren kann: Paar Nr. 1 schlägt Paar Nr. 2, wenn Paar Nr. 3 mitgezählt wird. Falls aber Paar Nr. 3 eliminiert wird, schlägt Paar Nr. 2 das Paar Nr. 1.

Diktatur: Davon spricht man, falls es möglich ist, daß ein einzelner Richter mehr Einfluß auf das Gesamtergebnis hat als es seinem fairen Anteil entspricht. Das soll bedeuten, daß die Rangfolge dieses Preisrichters mit der ganzen oder beinahe der ganzen Rangfolge des Gesamtergebnisses übereinstimmt, während das für die anderen Juroren keineswegs zutrifft, d. h. deren Rangfolge können sehr deutlich vom Ergebnis abweichen. Unterstellt man, daß eines der Hauptkriterien des *Skating Systems* ist, daß das Gesamtergebnis das Majoritätspräferenzmuster widerspiegelt, so hat dieses letztere Problem die schwerwiegendsten Konsequenzen für die "Fairness" des Systems.

Operations-Charakteristiken zur Beurteilung von Bewertungssystemen

Schüler werden mit Vorteil *operationale* und überprüfbare/meßbare Kriterien von "Fairness" aufstellen. Z. B. müssen sie einen Weg herausfinden, die Frage "Wie 'unfair' ist unfair?" zu beantworten. Sie werden Regeln aufstellen müssen, nach denen sie die Daten bearbeiten, sodaß "faire" Gesamtergebnisse herauskommen. Wie wir am *Skating System*-Beispiel gesehen haben, müssen solche Regeln nicht sehr anspruchsvoll oder von rechnerischer Komplexität sein. In der realen Welt müssen die Regeln ja auch einfach und leicht anzuwenden sein. Die Operationscharakteristiken dieser Regeln sollten getestet werden gegen den Hintergrund der aufgestellten Kriterien von "Fairness". Im Unterricht kann man durch Simulationsstudien verschiedene Bewertungsstrategien miteinander vergleichen [Hier wurden einige Auslassungen am Original vorgenommen].

Alles in allem kann das ganze Unterfangen ein stimulierendes Projekt darstellen, in welchem Schüler in einer integrierten, problemlöse-orientierten Art verschiedenste Fertigkeiten verwenden können. Die Betonung liegt dabei auf der Anwendung ihrer vorhandenen Fertigkeiten und ihres Wissens auf Common-Sense-Basis und nicht darauf, daß man mit dem Finger auf der "richtigen" Seite des Statistik-Lehrbuchs ist.

Literatur

- [1] Dawson, A.: The Skating System. Official Board of Ballroom Dancing Ltd. (1963,1980)
- [2] Hawkins, A. S.: Stretching the imagination. In: Teaching Statistics 7 (1985), S. 6 - 11.

Anne S. Hawkins
Institute of Education
University of London
Bedford Way
London WC 1 HOAL
GROSSBRITANNIEN