

LA SCIENCE STATISTIQUE AU SERVICE DES JEUX PARALYMPIQUES 2024 : L'EXEMPLE DU TIR À L'ARC

Christian Derquenne

Chercheur indépendant - chris.emr@wanadoo.fr

Résumé. Cet article a pour objectif de montrer en quoi la Science Statistique peut répondre aux enjeux issus du domaine du Sport. Nous illustrons cette dimension dans le domaine para-sportif et plus particulièrement axée vers les Jeux Paralympiques de 2024 à Paris. Les travaux présentés entrent dans le cadre du projet Paraperf de l'INSEP et du Mécénat de compétences EDF appuyé par EDF R&D. La discipline sportive présentée dans cet article est le tir à l'arc. Nous analysons les liens entre les phases qualitatives et finales, nous proposons des indicateurs de performance, et nous les utilisons pour construire des typologies d'archers et hiérarchiser leurs contributions au gain en finale.

Mots-clés. Science du Sport, Science Statistique, Jeux Paralympiques, tir à l'arc, indicateurs statistiques de performance individuelle

Abstract. This article aims to show how Statistical Science can respond to issues arising from the field of Sport. We illustrate this dimension in the para-sports field and more particularly focused on the Paralympic Games which will take place in Paris in 2024. The work presented falls within the framework of the Paraperf project of INSEP and the EDF Skills Sponsorship supported by EDF R&D. The sporting discipline presented in this article is archery. We analyze the links between the qualitative and final phases, we propose performance indicators, and we use them to construct typologies of archers and prioritize their contributions to winning in the final.

Keywords. Science of Sport, Statistical Science, Paralympic Games, archery, statistical indicators of individual performance

1 Contexte - objectif

2024 est une année phare pour le multi-sport : les Jeux Olympiques et Paralympiques (JOP) d'été en sont les dignes représentants. Ces deux événements majeurs organisés tous les quatre ans représentent sans doute le plus grand rassemblement de sportives et de sportifs de haut niveau réparti(es) en valides ou en situation de handicap au sein d'une multitude de disciplines les plus diverses. C'est une des raisons de leur succès, mais l'aventure humaine est la véritable raison d'être. En effet, les JOP permettent à la plupart des nations du Monde de se rencontrer avec un esprit de paix et de bonne intelligence. Bien évidemment la performance sportive et son amélioration représentent la clé de voûte des JOP comme toutes les compétitions sportives. Outre, les aspects physique, technique, stratégique, tactique, psychologique, matériel qui composent la performance sportive, la recherche de celle-ci s'appuie de plus en plus sur la Science (physique, chimie, biologie, sociologie, mathématiques, ...). Les britanniques ont notamment fait appel à cette dernière pour augmenter de façon significative le nombre de médailles lors des JOP

de Londres en 2012. Cette pluridisciplinarité de la Science est par conséquent de plus en plus utilisée par des fédérations et des clubs sportifs par le biais d’analyses poussées de la captation vidéo, d’amélioration du matériel, d’entraînements adaptés tenant compte de la charge physique et mentale pour éviter les blessures, ...

Cet article est dédié aux Jeux Paralympiques. En effet la situation de handicap fait partie de notre vie de tous les jours, qu’il soit physique, mental, de naissance ou accidentelle. L’intégration pleine et entière du handicap dans le monde sportif représente un lien humain et social fort en termes de partage et de sens de la vie. L’accès de personnes en situation de handicap à un nombre croissant de disciplines sportives en est la preuve : organisation d’événements sportifs allant des amateurs aux professionnels en rassemblant de plus en plus de types de handicap. Par conséquent, les JP 2024 constituent un enjeu fort en termes de reconnaissance sportive du handicap. Les parathlètes françaises et français sont les premiers acteurs avec leurs coaches pour réussir. L’utilisation de la Science est là pour les aider dans leurs stratégies et leurs décisions. L’INSEP a lancé en 2019 un vaste programme de fundraising comportant un volet mécénat et un volet sponsoring. Ce programme est conçu autour de valeurs humaines, orientées particulièrement vers l’accompagnement des athlètes. C’est dans ce cadre qu’EDF et plus particulièrement sa R&D a offert ses compétences scientifiques dans le cadre du Mécénat de compétences en contribuant au projet Paraperf (Optimisation de la performance Paralympique : de l’identification à l’obtention de la médaille). Celui-ci capitalise sur une approche interdisciplinaire ; il doit optimiser le parcours de chaque athlète français visant un podium aux Jeux Paralympiques de Paris 2024. Ce projet possède trois axes majeurs : les trajectoires de performance des parathlètes français(e)s et de leurs adversaires, les équipements, et les aspects socio-facteurs environnementaux. La présente communication traite du premier axe sous l’angle de la performance sportive dans les compétitions de haut niveau à l’aide d’indicateurs statistiques individuels et collectifs. L’objectif de ce lot est de comprendre et de situer les niveaux de performance des athlètes dans un contexte concurrentiel inter-épreuves. Cela peut aider à déterminer où les écarts relatifs sont les plus importants, ce qui révèle de plus grandes opportunités. En 2023, deux études de ce projet sur la densité de courses para-cyclistes sur route et sur la construction d’indicateurs de performance pour le développé-couché en haltérophilie ont fait l’objet d’un article [Derquenne, 2023]. Dans le présent papier, nous traitons de différentes problématiques sur le tir à l’arc dans la section 2. Nous concluons sur de futures études dans le cadre du projet Paraperf.

2 Etude sur le tir à l’arc pour les Jeux Paralympiques 2024

2.1 Problématiques, données et objectifs

Les compétitions en tir à l’arc se déroule en deux phases : une phase de qualification et une phase de confrontation (phase finale). Dans la phase de qualification, chaque archer tire 72 flèches ou 144 flèches selon la compétition, la somme des points constitue le classement et permet de construire l’arbre des confrontations. Un athlète bien classé va affronter un athlète moins bien classé (parfois, les premiers passent directement le 1er tour en fonction du nombre de participants). Le nombre de rounds varie en fonction du nombre de participants qui peut

être différent en fonction des catégories, du genre et de la compétition. Dans la phase de confrontation, pour chaque set, chaque archère ou archer tire 3 flèches, la somme des 3 tirs est effectuée, alors, le gagnant obtient 2 points, le perdant, 0 et 1 partout en cas d'égalité). Il y a 5 sets maximum (Le premier à 6 points remporte le match). En cas d'égalité au 5ème set : mort subite, ils tirent 1 flèche, celui qui met le plus de points remporte le match.

Les informations sont issues de données scrapées sur le site : <https://worldarchery.sport/fr>, à l'aide de l'API : <https://api.worldarchery.org/v3/API/>. Ces données contiennent toutes les compétitions depuis 2009 : 46 compétitions avec confrontations ; 4 593 confrontations individuelles ; 44 compétitions avec qualifications ; 243 épreuves.

Le principal objectif fixé par la DTN de la fédération du tir à l'arc est d'étudier le lien entre les résultats de la phase de qualification et de la phase de confrontation. Cependant, il s'agit d'un sujet large, il n'y avait pas une demande précise. Par conséquent, toute la liberté est donnée pour compléter cette demande en termes de méthodologies et d'approches statistiques.

Nous avons mis oeuvre plusieurs axes de travail : (i) Lien entre les rangs obtenus en qualification et en confrontation (ii) Construction d'indices de performance individuelle (iii) Typologie des par-athlètes à l'aide des indicateurs de performance (iv) Analyse multidimensionnelle des indicateurs de performance (v) Approches pour prédire l'accès au Podium et en Top 8 (vi) Etude des confrontations à l'aide des indices de performance (vii) Comparaison du classement réel de la compétition et de l'indicateur global de performance individuelle (viii) Classement annuel global des athlètes, des pays et des compétitions en fonction d'indices de performance (ix) Prédiction de performances futures à l'aide performances passées.

Chacune des deux phases a sa propre table de données. *La phase de qualification* possède les variables suivantes : rang du par-athlète en qualification, nombre de points obtenus (72 ou 144 flèches), nombre de 10 et de X (perfect, plein centre de la cible), nombre de X, identifiant du par-athlète, nom du par-athlète, prénom du par-athlète, nationalité du par-athlète, catégorie du handicap (CMO/CWO : arc à poulie ; RMO/RWO : arc classique ; MW1/WM1 : fauteuil - handicap plus important) et le numéro de la compétition. *La phase de confrontation* contient le numéro de la compétition, le nom de la compétition, le lieu de la compétition, le pays de la compétition, la date de début de la compétition, la date de fin de la compétition, le numéro de la phase de confrontation, le libellé de la phase de confrontation, l'identifiant du par-athlète 1, son nom, son prénom, sa nationalité, sa catégorie de handicap, son rang final en confrontation, son rang en qualification, ses résultats des flèches sur les sets, ses résultats des flèches en mort subite, son nombre de points obtenus, son état gagnant/perdant, idem pour le par-athlète 2, son adversaire. Enfin, cette table de données a été transformée de façon à obtenir un résultat individuel pour chaque par-athlète à la place de deux résultats associés à la confrontation des deux par-athlètes.

2.2 Lien entre les rangs obtenus en qualification et en confrontation

L'objectif de cette étude est d'analyser le lien entre les rangs individuels en qualification et en confrontation en répondant à la question : "est-ce que plus un par-athlète est bien classé en phase de qualification, plus il a de chance d'obtenir un bon classement en phase finale ?".

Pour répondre à cette question, nous avons calculé des tables de correspondances ou de passage du rang en qualification au rang final. Pour un même type de handicap, nous avons rassemblé

tous les résultats de chacun des archers aux différentes compétitions auxquelles il a participées. Pour cela, nous avons construit un tri croisé entre les deux rangs.

Par exemple, pour les CWO, nous obtenons la table de correspondances (table 1) dans laquelle les lignes correspondent aux rangs de la finale ; les colonnes aux rangs de qualification. Par exemple, il y a eu 16 archères qui ont été classées 1ères sachant qu’elles ont également été classées 1ères en qualification (cf. 1ère ligne de la 1ère colonne), puis 7 archères ont obtenu la 1ère place en finale, sachant qu’elles étaient classées 2èmes en phase de qualification, (cf. 1ère ligne de la 2ème colonne), etc. Ces résultats permettent de construire un diagramme à bâtons (figure 1) affichant les proportions d’être vainqueur en finale sachant le rang en qualification. Par exemple, elle est égale à 0,39 si l’archère a terminé au premier rang des qualifications, et tombe à 0,20 pour le rang 2, à 0,17 pour le rang 3, etc. Les lignes rouges horizontales pour chaque rang en qualification désignent la proportion observée de celui-ci. Par exemple, il y a une proportion de 0,06 d’archères qui ont obtenu le premier rang en qualification. Nous pouvons tout d’abord comparer empiriquement la proportion de rang 1 en qualification qui ont été vainqueurs et la proportion globale de rang 1 en qualification : 0,39 vs 0,06 ; elle est plus de six fois plus élevée.

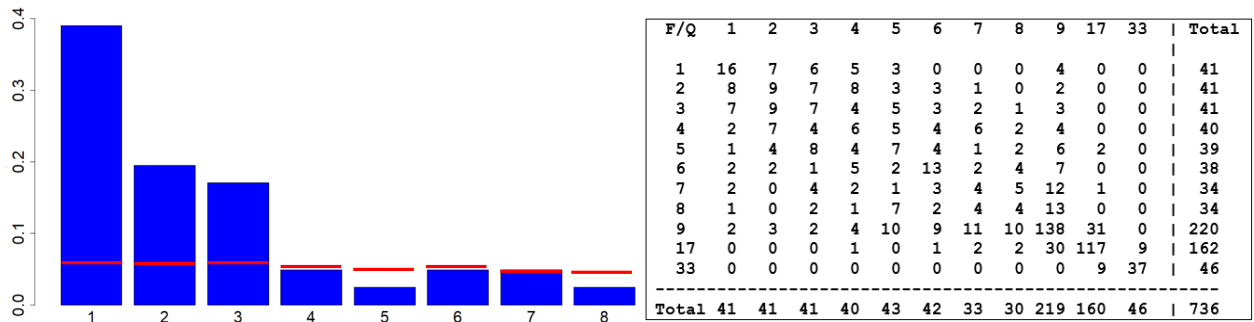


Figure 1 : Proportions de gagner la finale en fonction du rang en qualification Table 1 : Tri croisé des rangs en qualification vs en confrontation

La validation de ces comparaisons de proportions peut être réalisée à l’aide d’un test statistique dans lequel le jeu d’hypothèses est le suivant :

H_0 : Les $n^{(c_r)}$ individus sont tirés au hasard (sans remise) parmi les N individus de la population

H_1 : La proportion de la modalité q_r est ”anormalement” élevée parmi les $n^{(c_r)}$ individus

où N et $n^{(c_r)}$ sont respectivement le nombre total d’individus (ici 736) et le nombre d’archères ayant obtenu le rang r en phase de confrontation (pour $r = 1$, nous avons 41).

Lorsque les marges d’un tableau de contingence sont fixées (les nombres de modalités pour les variables catégorielles sont connues), alors chaque effectif croisé n_{rs} de cette table suit une loi Hypergéométrique $\mathcal{H}(N, n^{(c_r)}, n^{(q_s)})$ sous H_0 , où $n^{(q_s)}$ est le nombre d’archères ayant atteint le rang s en qualification (pour $s = 1$, nous obtenons 41). La p -valeur du test est donnée par $\Pr[\mathcal{H}(N, n^{(c_r)}, n^{(q_s)}) > n_{rs}^{obs} / H_0]$, où n_{rs}^{obs} est le nombre observé d’individus dans la case (r, s) du tri croisé, par exemple 16 dans la table 1 pour la case $s = 1$ et $r = 1$.

Dans [Morineau, 1982], l’auteur propose de construire une valeur-test égale au fractile de la loi normale centrée réduite $\mathcal{N}(0;1)$ au point $1 - p$ -valeur. Plus la p -valeur est faible, plus la

valeur-test est hautement positive et plus une modalité est typique dans sa classe. A l’opposé, une p -valeur élevée indiquera une modalité rare. En pratique, la valeur-test doit dépasser 2 à 3 en valeur absolue pour considérer que la modalité est fréquente ou rare de manière significative.

La table 2 permet d’identifier les rangs de qualification typiques de la place de vainqueur en finale. Les résultats issus des tests confirment les résultats empiriques. En effet, le rang 1 en qualification est très fréquent : 39,02% *vs* 5,57%, la p -valeur est très faible et fournit une valeur-test égale à 6,74 donc significatif. Comme les rangs 2 et 3 en qualification qui obtiennent respectivement des valeurs-test de 3,21 et de 2,72. Les autres rangs ne sont pas significatifs ni de façon fréquente, ni de façon rare.

Rang	Mod/Cla	Global	p.value	v.test
Qr=1	39.02	5.57	1.60e-11	6.74
Qr=2	19.51	5.57	1.33e-03	3.21
Qr=3	17.07	5.57	6.50e-03	2.72
Qr=7	4.88	4.62	8.70e-01	0.16
Qr=6	4.88	5.16	9.96e-01	-0.01
Qr=4	4.88	5.43	9.43e-01	-0.07
Qr=8	2.44	4.62	5.58e-01	-0.59
Qr=5	2.44	5.30	4.46e-01	-0.76

Table 2 : Tests de comparaison de proportions pour les vainqueurs

Les rangs 2, 3, puis 4 et 1 en qualification pour les finalistes sont significativement fréquents, alors que les rangs 5, 2, 3 et 1 le sont pour la 3^{ième} place du podium. Cette corrélation ”positive” de bons rangs en confrontation avec les rangs en qualification se poursuit pour les cinq suivants. Mais ce résultat fort pour la catégorie CWO est plus modulé pour les archers CMO.

2.3 Construction d’indices de performance individuelle

La mesure de performance individuelle fait partie des éléments importants pour les coaches dans les fédérations car ils leur permettent de disposer d’information pour les aider dans leurs décisions. Le principe pour construire des indices de performance individuelle est d’identifier des informations discriminantes, compréhensibles et surtout utilisables par les coaches, dans les données permettant de mesurer certaines capacités sportives du par-athlète. Deux types d’indicateurs ont été construits : statiques et dynamiques. Le premier type est associé aux résultats pour chacune des deux phases (qualification et confrontation), alors que le second mesure l’évolution des résultats entre les deux phases.

2.3.1 Indicateurs statiques

Sept indices statiques ont été construits.

Le premier indice met en avant le nombre de ”X” (perfect) obtenus par chaque archer lors de la phase de qualification (1).

$$I_{qX}(i) = \frac{\sum_{m=1}^{M_q} 1_{[f_m=X]}}{M_q} \quad (1)$$

où M_q est le nombre de flèches tirées, f_m est le score obtenu pour la $m^{ième}$ flèche. $1_{[f_m=X]} = 1$ si la $m^{ième}$ flèche est au centre de la cible (X), sinon $1_{[f_m=X]} = 0$. Par conséquent cet indicateur représente la proportion de ”perfects”. Plus la valeur est élevée, plus l’archer est performant pour atteindre le centre de la cible.

Le deuxième indice repose sur la différence de rangs entre deux compétiteurs. On pose tout d'abord cette différence : $b_{(i/j)} = r_{(q(i))} - r_{(q(j))}$ où $r_{(q(i))}$ et $r_{(q(j))}$ sont respectivement les rangs de qualification des archers i et j . L'objectif de cet indicateur est de bonifier le par-athlète i si son rang est plus élevé que celui de j , s'il a gagné la confrontation (2.1), alors que s'il perd et que son rang est plus faible que son adversaire, alors il est pénalisé (2.4). Par contre, cet indicateur est égal à 0, s'il gagne alors que son rang était meilleur que son adversaire (2.3), ou s'il perd alors que son rang était moins bon que son adversaire (2.2).

$$\text{Si } b_{(i/j)} > 0 \text{ et } W_i = 1 \text{ alors } I_{dr}(i) = b_{(i/j)} \quad (2.1)$$

$$\text{Si } b_{(i/j)} > 0 \text{ et } W_i = 0 \text{ alors } I_{dr}(i) = 0 \quad (2.2)$$

$$\text{Si } b_{(i/j)} < 0 \text{ et } W_i = 1 \text{ alors } I_{dr}(i) = 0 \quad (2.3)$$

$$\text{Si } b_{(i/j)} < 0 \text{ et } W_i = 0 \text{ alors } I_{dr}(i) = -b_{(i/j)} \quad (2.4)$$

où $W_i = 1$ représente le gain du match, alors que la perte est donnée par $W_i = 0$.

Cet indicateur est symétrique par rapport à son adversaire j . En effet, si $I_{dr}(i) = b_{(i/j)}$ alors $I_{dr}(j) = -b_{(i/j)} = b_{(j/i)}$; si $I_{dr}(i) = -b_{(i/j)}$ alors $I_{dr}(j) = b_{(i/j)} = -b_{(j/i)}$ et si $I_{dr}(i) = 0$ alors $I_{dr}(j) = 0$. Par conséquent, une valeur positive indiquera un gain notable de performance par rapport à son adversaire, alors que valeur négative engendrera une sorte de manque à gagner.

Le troisième indice mesure le score moyen des scores obtenus lors de la phase de confrontation par un archer i sur la même compétition, tel que :

$$I_{(sc_c f)}(i) = \frac{\sum_{m=1}^{M_c} f_m}{M_c} \quad (3)$$

avec M_c le nombre de flèches tirées en phase de confrontation (12 par match). Plus la valeur observée de (3) sera grande, plus l'archer sera performant.

Le quatrième indice utilise le rang $r_{(f(i))}$ obtenu par l'archer i lors de la phase finale :

$$I_{rang}(i) = r_{f(i)} \quad (4)$$

Contrairement aux indices introduits précédemment pour lesquels une grande valeur signifie une performance élevée, plus une valeur de (4) est faible, plus la performance de l'archer est forte.

Le cinquième indice mesure le nombre de "X" (perfect) obtenus pour chaque archer lors de la phase de confrontation (5).

$$I_{fX}(i) = \frac{\sum_{m=1}^{M_f} 1_{\{f_m=X\}}}{M_f} \quad (5)$$

où M_f est le nombre de flèches tirées sur l'ensemble des confrontations. Comme pour (1), cet indicateur représente la proportion de "perfects". Plus la valeur est élevée, plus l'archer est performant pour atteindre le centre de la cible.

Le sixième indice correspond à la dispersion des scores obtenus sur les jeux de 12 ou 15 flèches lors de la phase de confrontation. Il permet de mesurer la régularité (l'homogénéité ou la reproductibilité) des tirs de l'archer i .

$$I_{disp}(i) = \sqrt{\frac{\sum_{m=1}^M (f_m - \bar{f})^2}{M-1}} \quad (6)$$

Comme l'indice sur le rang final (4), plus une valeur de (6) est petite, plus le par-athlète est régulier, ce qui peut être signe d'un meilleur contrôle de soi.

Le septième indice correspond au coefficient de variation des scores obtenus sur les jeux de 12 ou 15 flèches lors de la phase de confrontation. Il permet de mesurer la dispersion par rapport à la moyenne des points des tirs de l'archer i .

$$I_{cv}(i) = \frac{I_{disp}(i)}{f} \quad (7)$$

Plus une valeur de (7) est petite, plus le par-athlète est régulier, ce qui peut être signe d'un meilleur contrôle de soi comme dans (6) mais aussi accompagné d'un score élevé.

2.3.2 Indicateurs dynamiques

Les trois indices mesurent l'évolution des résultats de chaque archer entre les 2 phases.

Le premier indice repose sur la différence de rangs entre les deux phases pour chaque archer (8). Si la valeur est positive alors l'archer i a amélioré son classement, si elle est négative, il a déprécié son rang, si elle est nulle, il a fait ce qui était attendu. On peut interpréter cet indice comme une évolution (positive, négative ou nulle) de son classement par rapport à une espérance qui correspond au rang de l'archer en phase de qualification.

$$I_{(ev-rq)}(i) = r_{q(i)} - r_{f(i)} \quad (8)$$

où $r_{q(i)}$ et $r_{f(i)}$ sont respectivement les rangs obtenus en qualification et en confrontation.

Le deuxième indice mesure l'évolution du score moyen (8) entre les deux phases :

$$I_{ev-sc}(i) = I_{sc-cf}(i) - I_{sc-ql}(i) \quad (9)$$

où $I_{(sc-ql)}(i) = \frac{\sum_{m=1}^{M_q} f_m}{M_q}$, avec M_q le nombre de flèches tirées en phase de qualification. Plus une valeur de (9) est grande, plus l'archer a amélioré son score moyen entre les deux phases.

Le troisième indice calcule la différence entre les moyennes des perfects (X) obtenus lors des deux phases. Plus une valeur de (10) est grande, plus l'archer a augmenté sa précision.

$$I_{ev_X}(i) = I_{fX}(i) - I_{qX}(i) \quad (10)$$

Les 10 indices introduits précédemment prennent des valeurs sur des échelles différentes. En effet, les rangs ou leur différence ne sont pas comparables aux scores moyens ou à leur différence, ni à la dispersion des flèches, etc. Par conséquent, une transformation de ces indices est nécessaire pour les rendre comparables et par la suite construire des indicateurs globaux. Pour cela, tous les indices sont transformés afin qu'ils prennent leurs valeurs entre 0 et 1. Plus le résultat de l'indicateur sera proche de 1, plus la performance de l'archer sera élevée. Pour les indicateurs de dispersion et de coefficient de variation intrinsèques, une valeur proche de 1, indiquera la régularité des performances de l'archer. La transformation appliquée est la suivante :

$$I^*(i) = \frac{I(i) - \min_j I(j)}{\max_j I(j) - \min_j I(j)} \quad (11)$$

Ces 10 indices permettent de réaliser de nombreuses études comme indiquées dans la partie 2.1. Nous avons choisi de présenter deux études une typologie des par-athlètes et l'analyse de l'importance de contribution des indices de performance dans le gain du match.

2.4 Typologies des par-athlètes à l'aide des indices de performance

Afin d'enrichir ces analyses, il peut être fructueux de rechercher des structures dans les données en construisant une typologie des par-athlètes à l'aide d'informations les renseignant. Dans le cas présent, nous utiliserons l'ensemble des indicateurs de performance individuelle, mais il serait également possible de séparer les indicateurs statiques et dynamiques. La méthode de classification est le critère de Ward adapté à des données numériques.

L'arbre de classification (dendrogramme) montre plutôt cinq classes (figure 2).

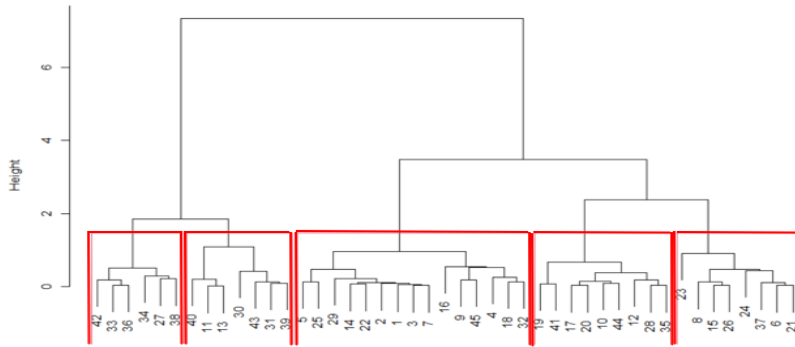


Figure 2 : Arbre de classification des archers

L'interprétation des classes est réalisée à l'aide du concept de la valeur-test permettant de classer l'importance des variables initiales ou supplémentaires et de déterminer le sens de celles-ci pour chaque groupe. Tout d'abord, les résultats suivants (table 3) montrent que les indicateurs les plus structurants pour construire la typologie sont les indices globaux statique et dynamique, le rang final avec des rapports de corrélation de plus 0,75 (plus une valeur est proche de 1, plus la variable est structurante), puis le score moyen final (0,61). Enfin, les trois indicateurs dynamiques possèdent également un poids important dans construction de la classification (0,57).

$$\rho^2(X_j, T) = \frac{\sum_{m=1}^{M_f} n_m (\bar{X}_{jm} - \bar{X}_j)^2}{\sum_{i=1}^n (X_{ji} - \bar{X}_j)^2} \quad (12)$$

où M , n_m , n_m et \bar{X}_j sont respectivement le nombre de classes de la typologie T, le nombre d'individus de la classe m , \bar{X}_{jm} la moyenne de X_j associée et la moyenne de X_j dans l'échantillon.

Indice	$I_{sta/dyn}^*$	I_{rang}^*	I_{scf}^*	$I_{ev_X}^*$	I_{evrg}^*	I_{evrg}^*	I_{fjX}^*	I_{disp}^*	I_{qX}^*	I_{dr}^*
$\rho^2(X_j, T)$	0,765	0,756	0,609	0,577	0,576	0,571	0,561	0,494	0,485	0,425

Table 3 : Importance des indices de performance

Le sens et la force des variables sont fournis par le concept de valeur-test introduit dans le paragraphe 2.2, mais cette fois-ci dans le cas de variables numériques, tel que :

$$vt(x_j) = \frac{\bar{x}_{jm} - \bar{X}_j}{s_{jm}} \quad (13)$$

où $s_{jm}^2 = (n - n_m)/(n - 1)s_j^2$ et s_j^2 est la variance de la variable X_j

Elle s'interprète comme un nombre d'écart-types autour de la moyenne. Plus la valeur-test est positive, (resp. négative), plus la moyenne \bar{X}_{jm} de la variable X_j dans la classe m est éloignée supérieurement (resp. inférieurement) de la moyenne \bar{X}_j de la variable X_j sur l'ensemble de l'échantillon. Dans notre cas, nous avons fixé des valeurs-tests limites à -1,96 ou 1,96.

Pour **la classe 1** (figure 4), le nombre moyen de perfects en phase de qualification est l'indice le plus discriminant et fortement supérieur à la moyenne générale pour l'ensemble des archers sur cette compétition. La valeur-test est égale à 4,19 (4,2 écart-types), la moyenne dans la classe vaut 0,551, alors que la moyenne globale est égale 0,384. Par contre, la moyenne de l'indicateur de l'évolution du nombre moyen de X est plus faible que la moyenne générale associée (0,316 vs 0,515), la valeur-test est égale à -3,72. Les indices sur le rang, du score moyen et de la dispersion en finale possèdent des moyennes plus élevées que les moyennes globales associées. Ce groupe contient des par-athlètes performants pour les indices associés aux valeurs-tests positives.

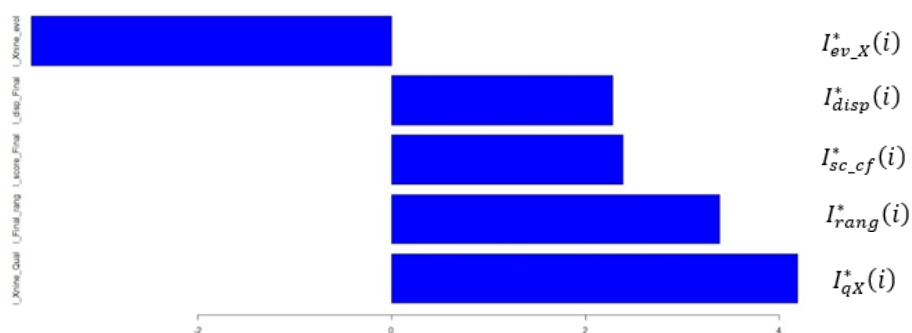


Figure 4 : Représentation des valeurs-tests pour la classe 1

La classe 2 possède des archers très performants sur de nombreux indicateurs, en particulier ceux d'évolution. La différence des rangs entre les deux archers en confrontation, les évolutions du score moyen et des rangs des archers entre les deux phases apparaissent de façon positive dans **la classe 3**. Par contre, le nombre moyen de plein centre dans la cible, lors de la phase qualificative, apparaît comme plus faible que la moyenne générale. Cette classe contient également des archers performants. **La classe 4** possède des par-athlètes moins performants que la moyenne générale pour de nombreux indicateurs. Seul l'indice sur l'évolution des perfects entre les deux phases apparaît de façon positive. Enfin, tous les indices de la **classe 5** possèdent des moyennes plus faibles que les moyennes générales associées. Les par-athlètes sont peu performants.

2.5 Importance des indices de performance individuelle sur la victoire

L'objectif est d'évaluer globalement par compétition, la force d'explication des indicateurs sur le gain du match. Pour cela, nous utilisons la valeur-test. Dans le cas présent les deux catégories à "expliquer" sont la victoire et la défaite. Dans ce cas, les valeurs-tests fournissent l'ordre d'importance des indices de performance individuelle. Elles peuvent être représentées sous forme d'un diagramme à bâtons plus lisible qu'un tableau de résultats (figure 5). La longueur de chaque barre correspond à la valeur-test associée à chaque indice. Elles sont toutes du côté positif. Les lignes verticales du milieu représentent la limite de "significativité" des indices. Toutes les

valeurs-tests qui dépassent la troisième ligne verticale correspondent aux indices significatifs. **I-Final-rang** est l'indice qui agit le plus sur la victoire de façon positive. La moyenne de celui-ci est de 0,844 pour la catégorie des gagnants alors qu'il vaut 0,723 sur l'ensemble de l'échantillon.

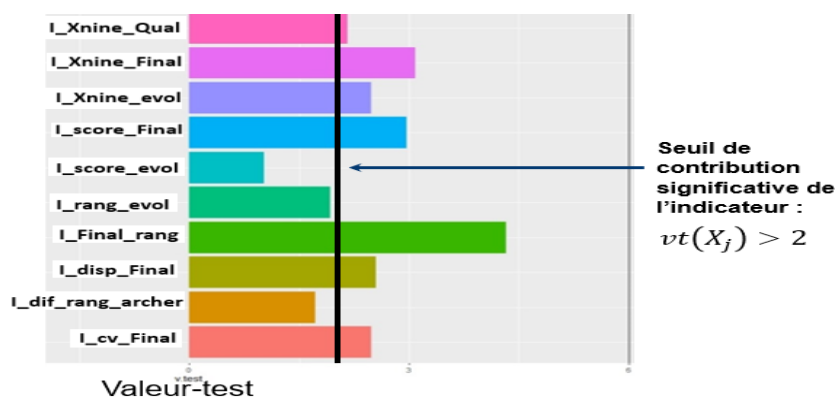


Figure 5 : Importance des indices à l'aide des valeurs-tests

3 Synthèse des résultats et voies futures

Les résultats présentés dans cet article ont montré que pour l'ensemble des compétitions les corrélations entre les rangs des deux phases étaient élevées, quel que soit le type d'handicap. Nous avons construit des indices de performance individuelle. Grâce à ceux-ci, nous avons recherché des structures en mettant oeuvre une analyse typologique pour obtenir des groupes d'archers et identifier leurs caractéristiques, ainsi que hiérarchiser les indicateurs de performance. D'autres analyses, non présentées ici, ont permis de détecter les caractéristiques potentielles pour atteindre le Top 8. Mais l'objectif majeur de cet article, outre la mise en oeuvre de méthodologies statistiques, était de montrer l'importance de la médiation scientifique, en d'autres termes la traduction des résultats obtenus en "langage terrain", ainsi que la mise à disposition d'outils d'aide à la décision à destination des fédérations, des coaches, des athlètes permettant d'adopter des stratégies adaptées aux parathlètes de façon individuelle et à l'égard de la concurrence. Cette chaine vertueuse offre un gain de temps, mais les décisions restent à la main des professionnels du sport. Enfin, nous développerons d'autres indices de performance individuelle et nous analyserons dans le temps l'évolution de différents par-athlètes.

Bibliographie

Derquenne Ch., (2023), Les Jeux Paralympiques 2024 : la Science Statistique en action, *54ièmes Journées de Statistique*, Bruxelles, Belgique.

Morineau A., (1987), Inferential techniques folowing a Multivariate Descriptive Statistical Analysis, *Actes du congrès d'Analyse des Données en Barcelone*.

Projet Paraperf, (2020), INSEP.