

# Visualiser les Premiers Moments d'une Distribution de Performances

Visualizing the First Moments of a Performance Distribution

Yves Guiard

**English Abstract** – Here the focus is the *what*, rather than the *how* of visualization. Human performance scores, strongly constrained by the performers' deliberate effort to extremize (i.e., minimize or maximize) them, should be recognized as a special class of measures. Their distributions are not just skewed, as often thought—they tend to be categorically asymmetric, showing a more or less abrupt front on their good side and a vanishing tail on their poor side. It is proposed that (1) the *best score* (whether sample minimum or maximum) is often better suited than any central-trend summary (mean, median, or mode) to locate a performance distribution (first raw moment) and that (2) to characterize the second moment of a performance distribution (its spread or scale) the *median deviation from best* is often better suited than the traditional spread statistics (variance, IQR, or min-max range), which all take it for granted that the second moment must be centered—a questionable presupposition when it comes to performance scores. Reexamining some pointing data, I show that the gradual relaxation of the error-minimization pressure, which allows the time-minimization effort to strengthen, entails a progressive shift of the best value of movement time down to the point where, the capacity limit being reached, the body of the distribution begins to cluster, thus forming a recognizable front. Essentially the same pattern is observable in the error data as the time-minimization effort is relaxed.

## 1 DÉFINITION DU CONCEPT DE PERFORMANCE

Nous désignerons comme une *performance* toute mesure comportementale censée être *extrémisée* par un agent. Dans le pointage le temps de mouvement doit être *minimisé* ; dans l'épreuve de mémorisation le nombre d'éléments correctement rappelés doit être *maximisé* ; à la bourse le gain financier doit être maximisé. L'extension de la classe d'équivalence ainsi définie est impressionnante (psychologie, économétrie, sciences de l'ingénieur, management, sport, etc.). Ignorant l'effort systématique d'extrémisation qui caractérise les mesures de performances, on les a traditionnellement assimilées, à tort, à des variables aléatoires. Le but de cette présentation est d'illustrer la nature quasi-déterministe des mesures de performance, et d'en tirer quelques implications du point de vue de la statistique et de la visualisation.

## 2 PERFORMANCES QUASI-DÉTERMINISTES

Certes, l'intensité de l'effort varie aléatoirement d'un test au suivant, mais la pression que l'effort fait subir au score est de direction constante (pour le coureur, par exemple, la minimisation du temps n'est pas une option). Ce fait est difficile à concilier avec le présupposé fondamental de la théorie des probabilités qui sous-tend nos statistiques usuelles : la fameuse loi de l'erreur de Moivre, Gauss et Laplace, qui nous dit comment les valeurs obtenues se dispersent de part et d'autre de la 'vraie' valeur, l'espérance.

Dans de nombreux cas l'effort a pour effet de faire buter les scores sur la capacité personnelle du performer. Il est évident qu'Usain Bolt, dont le meilleur temps sur 100 mètres est de 9,58 s, n'aurait jamais pu réaliser 9.00 s. Le paramètre de capacité est susceptible de varier à long terme (sous l'effet du développement, de l'apprentissage, etc.), mais à

l'échelle temporelle de la collecte d'un échantillon statistique la capacité est une limite infranchissable fonctionnant à la fois comme un attracteur global et comme un répulseur local.

On sait que la distribution d'une variable aléatoire est en cloche, avec deux points d'inflexion séparant un corps central concave de deux queues convexes. Les distributions de performances ne suivent pas cette loi. Nous verrons que l'effort croissant du performer fait disparaître d'abord l'un de ces points d'inflexion, donnant naissance à un front de distribution, puis le second, du fait de la compression de la queue. Ceci est aisément interprétable comme un effet d'accumulation des scores à proximité de la limite de capacité. Un score de performance dont la distribution, agglutinée contre une limite fixe, a perdu au moins l'un de ses deux points d'inflexion peut être qualifié de quasi-déterministe.

## 3 DISTRIBUTIONS DE TEMPS ET D'ERREUR

### DANS LE POINTAGE : EFFORT ET CAPACITÉ

On examinera des distributions obtenues par l'auteur dans des tâches de pointage, lesquelles demandent une performance bidimensionnelle où le performer doit minimiser à la fois son temps  $T$  et la valeur absolue de son erreur  $|E|$ . La balance entre ces deux efforts de minimisation contradictoires est manipulée via l'indice de difficulté  $ID$  dans la Section 3.1 et au moyen de consignes verbales dans la Section 3.2.

### 3.1. Distributions de $T$ dans le Paradigme de Fitts

Le paradigme de Fitts [1] impose un niveau variable de précision en jouant sur la distance  $D$  et la largeur  $W$  de la cible, charge au participant de trouver la vitesse maximale dont il est capable à chacun des niveaux de

précision prescrits. Dans [3] nous avons fait varier l'indice de difficulté  $ID = \log_2(2D/W)$  de 2 à 7. La Figure 1 montre comment par la distribution de  $T$  est affectée par la variation de la contrainte de précision.

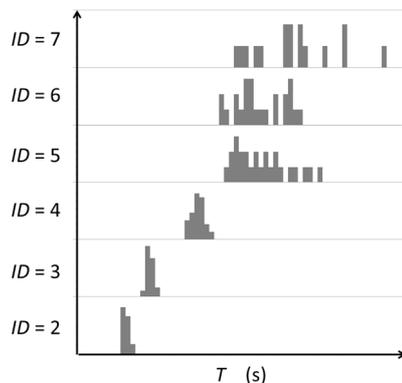


Figure 1. Effet, chez un participant de Mottet et al. (2001), de la réduction de l' $ID$  sur la distribution de  $T$ .

Dans une tâche qui demande à la fois vitesse et précision, réduire l' $ID$  c'est exiger plus de vitesse. Dans la Figure 1 l'effet d'attraction globale se manifeste par une translation progressive vers la gauche de la distribution des valeurs de  $T$ , et l'effet de répulsion locale consiste en une compression progressive de la distribution quand les mesures viennent buter sur un minimum infranchissable.

### 3.2. Distributions de $T$ et de $|E|$ dans le Paradigme de la Double Minimisation

Dans [2] nous avons placé nos participants devant une tablette graphique, leur demandant d'amener leur stylet aussi vite et précisément que possible sur une ligne-cible située à 150 mm d'un point de départ fixe. Ils devaient minimiser à la fois  $T$  et l'erreur  $|E|$ —l'écart final en mm entre la pointe du stylet arrêté et la cible. Nous avons fait varier la balance de l'effort depuis une condition de vitesse max jusqu'à une condition de précision max, en passant par trois conditions intermédiaires.

Les *whisker plots* des Figures 2 et 3 montrent comment se transforment les distributions de  $T$  et de  $|E|$  quand la consigne varie d'un extrême à l'autre. Le pattern est similaire pour les deux variables de performance. Sous l'effet de l'accroissement de l'effort de vitesse, les mesures de  $T$  migrent progressivement vers l'origine du graphique pour venir finalement s'agréger contre le minimum infranchissable, de l'ordre de 100 ms chez le participant considéré (Figure 2). Si le minimum de  $|E|$  est pratiquement insensible à la consigne, c'est qu'il était déterminé par la résolution de l'écran utilisé dans cette expérience. L'accroissement de l'effort de minimisation sur  $|E|$  se traduit donc essentiellement par un tassement progressif des scores contre la butée de capacité (Figure 3).

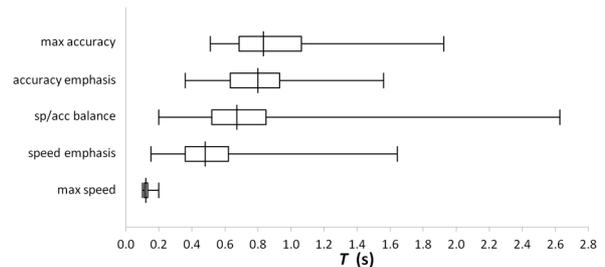


Figure 2. Effet de la manipulation par consigne de la balance vitesse/précision sur la mesure de  $T$ , chez un participant représentatif de Guiard et al. (2011).

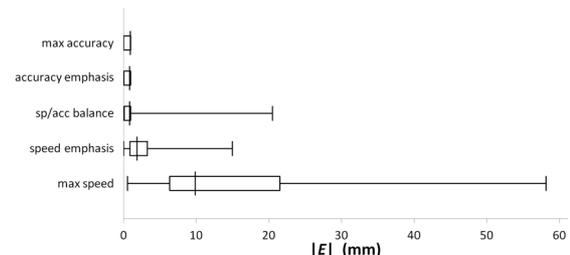


Figure 3. Effet de la manipulation par consigne de la balance vitesse/précision sur la mesure de  $|E|$ , chez un participant représentatif de Guiard et al. (2011).

## 4 CONCLUSIONS

À la différence de l'intensité de l'effort, manipulée de deux manières différentes dans ces études, la limite de capacité est une constante. On critiquera l'approche traditionnelle de la description statistique, tout entière organisée autour des notions de tendance centrale et de dispersion autour de ladite tendance centrale. À partir d'un certain effort de minimisation (ou de maximisation), les distributions de performances deviennent *catégoriquement* asymétriques, un fait qui prive de leur pertinence les techniques usuelles de description statistique. On s'efforcera de montrer l'intérêt d'une approche alternative organisée autour des notions de *meilleure valeur* (minimum ou maximum d'échantillon) et de *déviante médiane* (unidirectionnelle) par rapport à cet extremum.

## 5 REFERENCES

1. Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, 47(6), 381.
2. Guiard, Y., Olafsdottir, H. B., & Perrault, S. T. (2011). Fitts' law as an explicit time/error trade-off. *Proc. of ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 1619-1628). ACM.
3. Mottet, D., Guiard, Y., Ferrand, T., & Bootsma, R. J. (2001). Two-handed performance of a rhythmical Fitts task by individuals and dyads. *Journal of Experimental Psychology: Human Perception and Performance*, 27(6), 1275.