Mesure et évaluation en éducation

L'incertitude des seuils statistiques et l'établissement d'une norme de qualification sûre

Louis Laurencelle

Volume 25, Number 2-3, 2002

URI: https://id.erudit.org/iderudit/1088319ar
DOI: https://doi.org/10.7202/1088319ar

See table of contents

Publisher(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (print)
2368-2000 (digital)

Explore this journal

Cite this article

Laurencelle, L. (2002). L'incertitude des seuils statistiques et l'établissement d'une norme de qualification sûre. *Mesure et évaluation en éducation*, *25*(2-3), 19–33. https://doi.org/10.7202/1088319ar

Article abstract

In the course of a qualifying examination or aptitude test, a standard is sometimes set to help decide for the selection or refusal of a candidate. Now, the said standard, empirically determined, corresponds to a statistical threshold and is marked with uncertainty. We propose a method for determining a «safe standard» which accounts for this uncertainty, thanks to which the decision taken on each person is made to observe a given rate of error.

L'incertitude des seuils statistiques et l'établissement d'une norme de qualification sûre

Louis Laurencelle
Université du Québec à Trois-Rivières

MOTS CLÉS : Limites de tolérance normales, évaluation normative, norme sûre, sélection

Lors d'un examen ou d'un concours, on utilise parfois une norme de qualification afin de décider en faveur de la sélection ou du refus de la personne testée. Or, la dite norme, établie la plupart du temps de manière empirique, correspond à un seuil statistique et contient de l'incertitude. Nous proposons un procédé qui tienne compte de cette incertitude afin de déterminer une « norme sûre », c'est-à-dire une méthode grâce à laquelle la décision prise sur chaque personne respecte un taux d'erreur contrôlé.

KEY WORDS : Normal tolerance limits, normative evaluation, safe standard, selection

In the course of a qualifying examination or aptitude test, a standard is sometimes set to help decide for the selection or refusal of a candidate. Now, the said standard, empirically determined, corresponds to a statistical threshold and is marked with uncertainty. We propose a method for determining a « safe standard » which accounts for this uncertainty, thanks to which the decision taken on each person is made to observe a given rate of error.

Note de l'auteur : Toute correspondance peut être adressée à Louis Laurencelle, Département des sciences de l'activité physique, Université du Québec à Trois-Rivières, case postale 500, Trois-Rivières (Québec) G9A 5H7. Téléphone : 819-376-5011 (poste 3794) ; télécopieur : 819-376-5092 ; courriel : Louis_Laurencelle@uqtr.ca

L'examen de qualification est une étape incontournable dans plus d'un contexte, telles la sélection de personnel dans un poste de travail exigeant des compétences particulières, l'admission à une étape de sélection plus avancée, la réussite d'un test. Par exemple, on retiendra les candidats qui, à l'examen, ont pu accomplir correctement les tâches proposées ou ceux, dans d'autres cas, qui ont obtenu un score atteignant un seuil prescrit. Les procédures de qualification qui s'appuient sur un seuil, ou une *norme de réussite*, sont généralement de trois types. En premier lieu, la norme peut être appliquée en vérifiant la présence ou l'absence de caractéristiques préétablies, par exemple à l'aide d'une liste de vérification : c'est le cas de l'examen clinique, dans lequel le clinicien contrôle l'absence de tout symptôme, jauge les réflexes, etc. La norme consiste parfois en une valeur de performance qu'il faut atteindre et qui est prescrite sans autre justification : c'est le cas habituel des tests scolaires, pour lesquels la note de 60 % est traditionnellement imposée pour « passer ». Le troisième type de norme concerne celles émanant d'une base statistique et justifiées en regard d'une population supposée : pour « réussir » le test ou être retenu comme candidat qualifié, il faut obtenir un score X qui atteigne ou dépasse un seuil L_S, la personne étant alors réputée faire partie de la portion supérieure de la population. C'est ce type de norme qui est appliqué pour la sélection de personnel, pour le recrutement d'étudiants dans les programmes d'études universitaires à accès restreint, voire pour le repérage pathognomonique de cas problèmes, le seuil appliqué se situant alors dans la portion inférieure de la population.

La norme de population a donc pour but de définir le niveau de performance ou de qualification à partir duquel un candidat peut prétendre faire partie de la portion supérieure de population. Soit L_S, la limite supérieure qui fournit cette norme : tout candidat obtenant au test un score X tel que X déborde L_S fait partie de la portion supérieure. La justification et les diverses applications de ce type de norme n'entrent pas dans le cadre du présent article : notons cependant que cette norme est équitable et aveugle, puisqu'elle ne dépend pas d'une théorie ou d'une analyse conceptuelle du domaine d'application, et que la part d'arbitraire y est réduite sinon éliminée, comparativement à une norme numérique imposée *ex cathedra*.

Sur la base des données connues, on exigera par exemple un QI d'au moins 130 pour être admis au baccalauréat en psychologie, ou bien un niveau de $\mathring{V}O_2$ max de plus de 45,5 ml.kg^{-1}.min^{-1} d'O_2 pour participer à l'équipe élite de marathon canadienne, etc. La sélection, on le conçoit, vise à assurer un niveau de réussite ou de performance éventuelle satisfaisant pour les candidats reçus.

Définition d'une norme empirique sûre

La norme L_S caractérise et délimite la fraction supérieure f de la population concernée, selon :

(1) $\Pr\{ X \geq L_S \} = f.$

Cette norme L_S, toutefois, n'est pas connue, non plus que n'a été mesurée toute la population[1]. Pour établir la norme, l'usage courant consiste à se baser sur un échantillon de la population, comportant n éléments, et à élaborer une estimation de la norme L_S sur cette base. Supposons qu'on ait en mains un estimateur L_S' calculé à partir des observations (X_1, X_2,..., X_n) de l'échantillon. La plupart des estimateurs, on le sait, sont consistants, en ce sens qu'ils tendent vers la valeur (réelle) du paramètre à mesure que la taille d'échantillon n augmente, de sorte que :

(2) $\Pr\{ X \geq L_S' \} \rightarrow f$ si $n \rightarrow \infty$.

Nonobstant cette tendance asymptotique, il reste que la valeur d'un estimateur basé sur $n = 25$ éléments, ou sur $n = 100$ ou 200, va fluctuer considérablement, et il y a lieu de se demander si la norme empirique ainsi obtenue est sûre.

Comment définir une « norme sûre », la « sûreté d'une norme » ? En premier lieu, convenons d'utiliser un estimateur particulier d'une norme de réussite, soit la valeur située à un écart standardisé au-dessus de la moyenne,

(3) $L_S' = \overline{X} + z_{1-f}s,$

lequel estimateur a pour cible la valeur paramétrique, et inconnue, correspondante :

(4) $L_S = \mu + z_{1-f}\sigma\ ;$

l'écart standardisé z_{1-f} est défini plus loin. En raison de la fluctuation due à l'échantillonnage et l'imprécision des statistiques $\overline{X}$ et s, l'estimateur L_S' va osciller autour de L_S, la vraie norme. En établissant notre norme empirique, nous souhaitons que tout individu non qualifié, celui pour lequel $X < L_S$, ne soit pas retenu par le critère $X \geq L_S'$, et ce *avec une bonne probabilité* : c'est cette probabilité qui définit la sûreté de la norme.

Pour établir cette probabilité, considérons quelqu'un tout juste non qualifié, c'est-à-dire quelqu'un qui mériterait le score $X_{NQ} = L_S - \varepsilon$, $\varepsilon > 0$. Cette personne serait rejetée par la norme exacte, de sorte que :

(5a) $\Pr\{ X_{NQ} \geq L_S \} = 0.$

D'autre part, la personne tout juste qualifiée, avec $X_Q = L_S + \varepsilon$, sera sûrement retenue, soit :

(5b) $\Pr\{ X_Q \geq L_S \} = 1$.

À la limite, pour la personne ayant un niveau de qualification limitrophe, $\varepsilon \to 0$ et $X_{NQ} \to L_S$, la probabilité de rétention oscillera entre 0 et 1, de sorte que :

(5c) $\Pr\{ X_{NQ} \geq L_S \} \to 1/2$ si $\varepsilon \to 0$.

Dans le cas de l'utilisation de la norme exacte L_S, nous voyons que la marge de variation entre le rejet (5a) et la sélection (5b) est très fine et sûre, puisque, par exemple, quelqu'un tout juste non qualifié est certain d'être rejeté. La norme empirique L_S', quant à elle, est incertaine et la marge de variation entre sélection et rejet a de l'épaisseur. Il est néanmoins possible et approprié de définir une norme sûre Λ_S telle que :

(6) $\max \Pr\{ X \geq \Lambda_S \mid X < L_S \} = \alpha$.

Cette norme empirique[2], basée par exemple sur les statistiques échantillonnales $\overline{X}$ et s, assure que quelqu'un non qualifié sera retenu selon une probabilité d'au plus α, ou rejeté selon une probabilité d'au moins $1 - \alpha$. Pour obtenir $\alpha = 0$ et retrouver le critère de sûreté absolue donné en (5a), il faudrait bien entendu un échantillon de taille infinie.

Modèle et calcul d'une norme empirique sûre

Le seuil empirique L_S', défini en (3), fluctue avec l'échantillon dont il provient. Supposons que nous ayons affaire à des mesures d'un test de capacités intellectuelles dans une population homogène, que les scores se répartissent selon une distribution normale, et que ses paramètres, qui en fait sont inconnus, égalent $\mu = 100$ pour la moyenne et $\sigma = 15$ pour l'écart-type : ce sont là des caractéristiques typiques du QI. Nous ignorons donc la valeur des paramètres μ et σ. Cependant, à des fins de recrutement de candidats exceptionnels, nous souhaitons sélectionner des personnes faisant partie des 5 % les plus douées de la population.

La norme paramétrique visée. En connaissant les valeurs paramétriques μ et σ, il nous serait possible de définir une norme rigoureuse comme (4). Selon la distribution normale à valeurs standardisées (z), de moyenne 0 et d'écart-type 1, on retrouve la fraction $f = 0{,}05$ supérieure au-delà du score $z_{0,95} \approx 1{,}645$, de sorte que la norme absolue serait $L_S = \mu + z_{0,95}\,\sigma = 100 + 1{,}645{\times}15 \approx 124{,}7$, dans notre cas.

Un exemple de norme empirique obtenue. À défaut de connaître la vraie norme L_S = 124,7, nous procédons en échantillonnant et en recrutant n = 25 personnes, dont les scores divers ont pour moyenne $\overline{X}$ = 97,8 et pour écart-type s = 13,6. La norme approximative obtenue, selon (3)[3], serait donc $L_S' = \overline{X} + z_{0,95}\, s = 97{,}8 + 1{,}645 \times 13{,}6 \approx 120{,}2$. Que vaut cette norme ?

En fait, pour fabriquer la norme empirique L_S', nous devons remplacer la valeur ponctuelle du paramètre μ (= 100) par celle d'un estimateur statistique, la moyenne $\overline{X}$, qui elle fluctue de façon plus ou moins importante. La moyenne $\overline{X}$, en fait, se distribue elle-même selon la loi normale, avec pour moyenne μ (= 100) et écart-type $\sigma/\sqrt{n}$ (= 3), tel que l'illustre la figure 1 pour la moyenne de 25 QI.

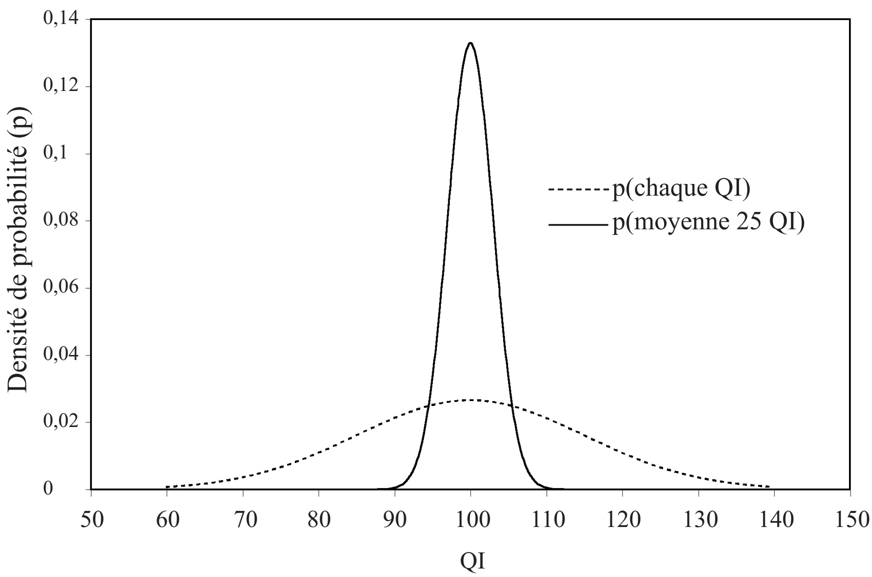

Figure 1. ***Distribution du QI moyen de 25 personnes et du QI individuel***

La norme empirique remplace aussi la valeur ponctuelle du paramètre σ (= 15) par celle de l'écart-type s estimé dans l'échantillon. Ce dernier fluctue aussi, cette fois selon la loi du χ (Khi), la racine carrée du Khi-deux, avec $n - 1$ degrés de liberté : en fait, la statistique s se distribue comme $\sigma\, \chi_{n-1}/\sqrt{n\text{-}1}$. La figure 2 illustre cette variation pour l'écart-type de 25 QI.

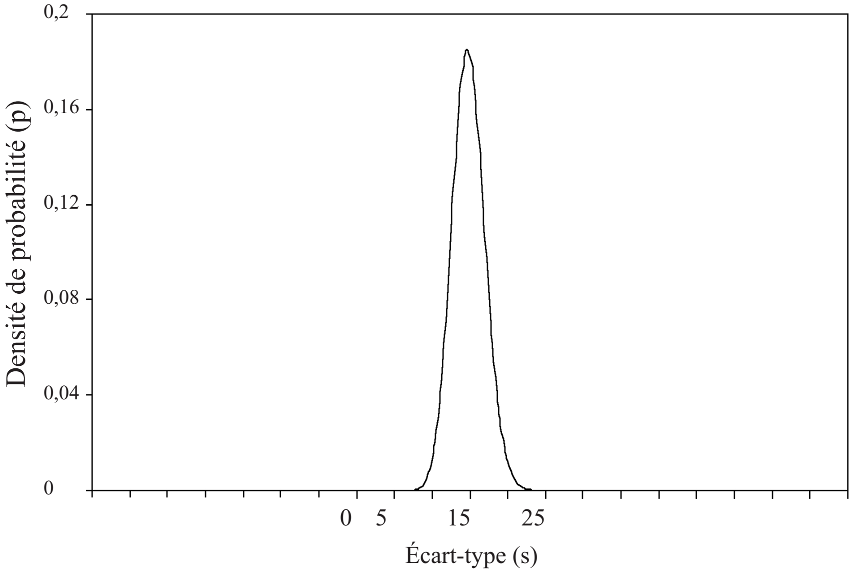

Figure 2. ***Distribution de l'écart-type du QI de 25 personnes (Noter que l'axe des abscisses a été défini à la même échelle que celle utilisée à la figure 1)***

Puisque, dans la fabrication d'une norme empirique pour estimer $\mu + z_{1-f}\sigma$, les ingrédients $\overline{X}$ et s fluctuent, la norme résultante, $L_S' = + z_{1-f}s$, fluctue elle aussi. Laurencelle (2000) montre qu'elle est biaisée, son espérance (ou moyenne) n'égalant pas tout à fait L_S, et que sa variance dépend principalement du facteur z_{1-f}, soit :

(7a) $\mathrm{E}\{ L_S' \} \approx L_S - \sigma\, z_{1-f} / (4n–4)$,

(7b) $\mathrm{var}\{ L_S' \} \approx \dfrac{\sigma^2}{n} [\, 1 + z^2_{1-f} \times (4n^2–5n) / 8(n–1)^2 \,] \} \approx \dfrac{\sigma^2}{n} [\, 1 + z^2_{1-f} / 2 \,]$.

On peut garantir la sûreté de la norme, ici une norme supérieure (ou relative à la fraction supérieure de la population), de deux façons distinctes : soit en contrôlant et minimisant le risque de sélectionner quelqu'un non qualifié : c'est la norme *exigeante* ; soit en minimisant le risque de rejeter quelqu'un qualifié : c'est la norme *permissive*. Le critère (6), qui minimise le risque d'accepter quelqu'un non qualifié, fournit la norme exigeante.

La probabilité définie par l'expression (6) est évidemment maximale lorsque $X \to L_S$ ou $X = L_S - \varepsilon$, ε étant une quantité positive infinitésimale. Posons donc, carrément, $X = L_S$ et cherchons à évaluer la fonction de probabilité :

(8) $P(\overline{X}, s, t) = \Pr\{ L_S \geq \overline{X} + t\, s \}$

pour toutes les valeurs de $\overline{X}$ et de s possibles, selon leurs distributions respectives (telles que représentées aux figures 1 et 2), soit :

(9) $$P(t) = \int_X \int_S \Pr\{ L_S \geq \overline{X} + t\,s \}\, d\overline{X}\, ds\,.$$

La documentation suggère différentes méthodes pour évaluer cette double intégrale (Johnson, Kotz & Balakrishnan, 1994, 1995 ; Kendall & Stuart, 1977 ; Laurencelle, 2000) : voir l'annexe au présent article. Cela fait, il suffit de faire *avancer t* depuis sa position de départ, z_{1-f}, vers le haut : à mesure que t croît, le seuil empirique devient de plus en plus sévère, à probabilité $P(t)$ décroissante, jusqu'à ce que cette dernière atteigne le seuil voulu α, ou jusqu'à ce que la probabilité de sélectionner quelqu'un tout juste non qualifié égale α. Ce procédé fournit ainsi la norme exigeante λ_E, telle que :

(10) $$\max \Pr\{ X \geq \overline{X} + \lambda_E\, s \mid X < L_S \} = \alpha.$$

Comme on voit, la quantité λ_E est exprimée en valeur standardisée et doit être combinée aux statistiques de l'échantillon pour donner le seuil applicable, $\Lambda_S = \overline{X} + \lambda_E s$. Elle est basée ici sur le modèle normal et dépend essentiellement de deux quantités : la taille n et la fraction de sélection f. On peut donc écrire $\lambda_E(n, f)$.

La norme sûre, $\lambda_E(n, f)$ ou λ_E, que nous présentons ici dans un contexte de sélection et de qualification, est connue dans la documentation statistique sous le vocable de « limite de tolérance » ; différentes réalisations de ce concept ont cours en industrie, en particulier dans le secteur dit du contrôle de la qualité (voir par exemple Burr, 1979). Odeh et Owen (1980) ont publié plusieurs jeux de tables relatives à différentes définitions des limites de tolérance, quelques-unes seulement s'appliquant au présent contexte. Laurencelle (2000) présente aussi des tables : nous reproduisons ici, au tableau 1, une partie de son tableau 5, qui donne les limites de tolérance supérieures de mode exigeant λ_E, pour les fractions f = 0,50, 0,40, 0,30, 0,20, 0,10, 0,05 et 0,01 et le niveau de sûreté $1 - \alpha = 0{,}95$. Nous proposons aussi, au tableau 2, des limites pareilles de mode permissif λ_P, *i.e.* telles que quelqu'un qualifié (*i.e.* appartenant à la fraction supérieure f) ait une probabilité d'au plus α d'être rejeté.

Tableau 1

Norme exigeante unilatérale, avec niveau de sûreté de 1 – α = 0,95

	f	*0,5*	*0,4*	*0,3*	*0,2*	*0,1*	*0,05*	*0,01*
n	4	1,177	1,672	2,265	3,026	4,162	5,144	7,042
	5	0,953	1,370	1,861	2,483	3,407	4,203	5,741
	6	0,823	1,199	1,638	2,191	3,006	3,708	5,062
	7	0,734	1,087	1,495	2,005	2,755	3,399	4,642
	8	0,670	1,006	1,393	1,875	2,582	3,187	4,354
	9	0,620	0,945	1,317	1,779	2,454	3,031	4,143
	10	0,580	0,896	1,257	1,703	2,355	2,911	3,981
	11	0,546	0,856	1,208	1,643	2,275	2,815	3,852
	12	0,518	0,823	1,168	1,593	2,210	2,736	3,747
	13	0,494	0,794	1,133	1,551	2,155	2,671	3,659
	14	0,473	0,770	1,104	1,514	2,109	2,614	3,585
	15	0,455	0,748	1,078	1,483	2,068	2,566	3,520
	16	0,438	0,729	1,056	1,455	2,033	2,524	3,464
	18	0,410	0,696	1,017	1,409	1,974	2,453	3,370
	20	0,387	0,669	0,986	1,371	1,926	2,396	3,295
	25	0,342	0,619	0,927	1,302	1,838	2,292	3,158
	49	0,240	0,506	0,798	1,150	1,650	2,070	2,869
	64	0,209	0,472	0,761	1,107	1,597	2,008	2,789
	81	0,185	0,446	0,732	1,074	1,557	1,962	2,730
	100	0,166	0,426	0,710	1,049	1,527	1,927	2,684
	150	0,135	0,393	0,674	1,008	1,478	1,870	2,611
	200	0,117	0,374	0,653	0,984	1,450	1,837	2,570
	250	0,104	0,361	0,639	0,969	1,431	1,815	2,542
	300	0,095	0,351	0,629	0,957	1,417	1,800	2,522
	400	0,082	0,338	0,614	0,941	1,398	1,778	2,494
	500	0,074	0,329	0,605	0,930	1,385	1,763	2,475
	750	0,060	0,315	0,590	0,913	1,365	1,741	2,447
	1000	0,052	0,307	0,581	0,904	1,354	1,727	2,430
	∞	0,000	0,253	0,524	0,842	1,282	1,645	2,326

Tableau 2

Norme permissive unilatérale, avec niveau de sûreté de 1 – α = 0,95

	f	*0,5*	*0,4*	*0,3*	*0,2*	*0,1*	*0,05*	*0,01*
n	4	–1,177	–0,746	–0,355	0,021	0,444	0,743	1,246
	5	–0,953	–0,583	–0,237	0,110	0,519	0,818	1,331
	6	–0,823	–0,483	–0,160	0,173	0,575	0,875	1,396
	7	–0,734	–0,413	–0,103	0,220	0,619	0,920	1,449
	8	–0,670	–0,361	–0,060	0,258	0,655	0,958	1,493
	9	–0,620	–0,319	–0,025	0,290	0,686	0,990	1,530
	10	–0,580	–0,286	0,004	0,316	0,712	1,017	1,563
	11	–0,546	–0,258	0,029	0,339	0,734	1,041	1,591
	12	–0,518	–0,233	0,050	0,359	0,754	1,062	1,616
	13	–0,494	–0,213	0,069	0,376	0,772	1,081	1,638
	14	–0,473	–0,194	0,086	0,392	0,788	1,098	1,658
	15	–0,455	–0,178	0,100	0,406	0,802	1,114	1,677
	16	–0,438	–0,163	0,114	0,419	0,815	1,128	1,694
	18	–0,410	–0,138	0,137	0,441	0,839	1,153	1,724
	20	–0,387	–0,117	0,156	0,460	0,858	1,175	1,749
	25	–0,342	–0,077	0,194	0,497	0,898	1,217	1,801
	49	–0,240	0,018	0,286	0,590	0,997	1,327	1,933
	64	–0,209	0,048	0,315	0,619	1,030	1,362	1,976
	81	–0,185	0,071	0,337	0,643	1,056	1,391	2,011
	100	–0,166	0,089	0,356	0,661	1,077	1,414	2,040
	150	–0,135	0,119	0,386	0,693	1,112	1,454	2,089
	200	–0,117	0,137	0,404	0,712	1,133	1,478	2,118
	250	–0,104	0,149	0,416	0,725	1,148	1,494	2,139
	300	–0,095	0,158	0,426	0,735	1,159	1,507	2,154
	400	–0,082	0,171	0,439	0,749	1,175	1,525	2,176
	500	–0,074	0,179	0,447	0,758	1,186	1,537	2,191
	750	–0,060	0,193	0,461	0,773	1,203	1,556	2,215
	1000	–0,052	0,201	0,470	0,782	1,213	1,567	2,230
	∞	0,000	0,253	0,524	0,842	1,282	1,645	2,326

Il est possible aussi de concevoir des normes sûres *bilatérales*, qui s'inspirent du même raisonnement. En mode exigeant par exemple, la norme bilatérale standardisée $\pm\lambda_E$ est telle que quelqu'un n'appartenant pas à l'une des fractions extrêmes de la population, soit la fraction f / 2 inférieure ou la fraction f / 2 supérieure, ait une probabilité d'au plus α d'y être classé, soit[4] :

$$\text{(11)} \qquad \max \Pr\{X \geq \overline{X} + \lambda_E s \text{ ou } X \leq \overline{X} - \lambda_E s \mid \mu - \lambda_E \sigma < X < \mu + \lambda_E \sigma\} = \alpha .$$

Le risque de rétention étant doublé, soit dans les portions supérieure ou inférieure de la population, il s'ensuit que la norme bilatérale est légèrement plus sévère que son pendant unilatéral. Les tableaux 3 et 4 fournissent quelques valeurs représentatives. Des tables plus complètes doivent paraître bientôt[5]. Le contexte d'application de telles normes reste encore peu développé.

Tableau 3

Norme exigeante bilatérale, avec niveau de sûreté de 1 – α = 0,95

	f	*0,5*	*0,6*	*0,7*	*0,8*	*0,9*	*0,95*	*0,99*
n	4	1,263	1,640	2,136	2,793	3,801	4,691	6,454
	5	1,084	1,416	1,830	2,371	3,192	3,916	5,349
	6	0,996	1,300	1,670	2,149	2,875	3,514	4,776
	7	0,944	1,228	1,571	2,012	2,680	3,266	4,424
	8	0,909	1,179	1,503	1,919	2,547	3,098	4,185
	9	0,884	1,143	1,453	1,850	2,449	2,975	4,011
	10	0,865	1,116	1,415	1,798	2,375	2,881	3,877
	11	0,849	1,094	1,385	1,757	2,316	2,806	3,772
	12	0,837	1,076	1,360	1,723	2,268	2,746	3,686
	13	0,827	1,061	1,339	1,695	2,228	2,695	3,615
	14	0,818	1,048	1,322	1,671	2,194	2,653	3,554
	15	0,810	1,038	1,307	1,650	2,165	2,616	3,502
	16	0,804	1,028	1,294	1,632	2,140	2,584	3,457
	18	0,793	1,012	1,272	1,603	2,098	2,531	3,382
	20	0,784	1,000	1,255	1,579	2,064	2,488	3,322
	25	0,768	0,977	1,223	1,536	2,003	2,411	3,213
	49	0,735	0,929	1,157	1,446	1,876	2,251	2,987
	64	0,726	0,916	1,139	1,421	1,841	2,208	2,925
	81	0,719	0,906	1,126	1,403	1,816	2,175	2,880
	100	0,714	0,899	1,115	1,389	1,796	2,150	2,845
	150	0,706	0,887	1,099	1,367	1,765	2,111	2,789
	200	0,701	0,880	1,090	1,354	1,747	2,089	2,758
	250	0,698	0,876	1,084	1,346	1,736	2,074	2,737

300	0,696	0,873	1,079	1,340	1,727	2,063	2,722
400	0,693	0,868	1,073	1,332	1,715	2,049	2,701
500	0,691	0,865	1,069	1,326	1,707	2,039	2,687
750	0,688	0,861	1,063	1,317	1,695	2,023	2,665
1000	0,686	0,858	1,059	1,312	1,688	2,015	2,653
∞	0,000	0,253	0,524	0,842	1,282	1,645	2,326

Tableau 4

Norme permissive bilatérale, avec niveau de sûreté de 1 – α = 0,95

	f	*0,5*	*0,4*	*0,3*	*0,2*	*0,1*	*0,05*	*0,01*
n	4	0,122	0,154	0,193	0,307	0,601	0,836	1,259
	5	0,139	0,176	0,219	0,390	0,687	0,924	1,360
	6	0,153	0,194	0,242	0,455	0,752	0,992	1,438
	7	0,166	0,209	0,292	0,507	0,805	1,047	1,501
	8	0,177	0,223	0,334	0,550	0,848	1,093	1,553
	9	0,187	0,235	0,370	0,585	0,884	1,131	1,598
	10	0,196	0,246	0,400	0,615	0,916	1,165	1,637
	11	0,204	0,257	0,427	0,642	0,943	1,194	1,671
	12	0,211	0,271	0,450	0,665	0,967	1,220	1,701
	13	0,219	0,292	0,471	0,685	0,989	1,243	1,728
	14	0,225	0,310	0,489	0,704	1,009	1,264	1,753
	15	0,231	0,327	0,505	0,721	1,026	1,283	1,775
	16	0,237	0,342	0,520	0,736	1,043	1,301	1,795
	18	0,248	0,369	0,547	0,763	1,071	1,332	1,831
	20	0,258	0,391	0,569	0,786	1,096	1,359	1,863
	25	0,279	0,435	0,613	0,832	1,145	1,412	1,925
	49	0,386	0,544	0,724	0,947	1,272	1,549	2,087
	64	0,420	0,579	0,760	0,985	1,313	1,595	2,141
	81	0,447	0,606	0,789	1,015	1,347	1,631	2,184
	100	0,469	0,628	0,812	1,040	1,374	1,661	2,219
	150	0,505	0,666	0,851	1,081	1,420	1,712	2,279
	200	0,527	0,688	0,874	1,106	1,448	1,743	2,316
	250	0,542	0,704	0,891	1,124	1,468	1,764	2,342
	300	0,553	0,715	0,903	1,137	1,482	1,780	2,361
	400	0,569	0,732	0,920	1,156	1,503	1,803	2,388
	500	0,580	0,743	0,932	1,168	1,517	1,819	2,407
	750	0,597	0,761	0,950	1,188	1,540	1,844	2,437
	1000	0,607	0,771	0,962	1,201	1,553	1,859	2,455
	∞	0,000	0,253	0,524	0,842	1,282	1,645	2,326

Le lecteur imaginatif peut déjà entrevoir des variantes simples de l'usage de ces normes sûres. Par exemple, dans un contexte donné, on peut vouloir retenir toute la population sauf sa fraction inférieure f, cela en s'assurant que les candidats rejetés soient vraiment dignes de l'être. Il s'agit alors de minimiser la probabilité de rejeter quelqu'un logé dans la fraction supérieure $1-f$, c'est-à-dire de s'assurer (avec une probabilité d'au moins $1 - \alpha$) que les candidats refoulés dans la fraction inférieure f devaient l'être. Après quelque réflexion, la norme qu'il convient d'appliquer au cas présent est la norme exigeante $-\lambda_E(n, f)$, donc la limite inférieure $\Lambda_I = \overline{X} - \lambda_E(n, f) \times s$.

« Norme sûre » signifie aussi décision ou sanction qu'on peut appuyer et argumenter, notamment dans un litige en cour de justice. L'utilisation de tests et de mesures pour la sélection, la qualification et la requalification de personnel a entraîné des plaintes pour rejet ou congédiement arbitraires, ou pour disqualification infondée. La base des plaintes peut être diverse et avoir trait, par exemple, à la validité (ou pertinence) des mesures employées. Le rejet de quelqu'un parce que sa mesure n'atteint pas une norme prescrite constitue une autre base, qui a donné lieu récemment à de retentissantes répercussions légales[6]. Non seulement le présent point de vue sur l'incertitude des seuils statistiques, démontrée par les équations (7a) et (7b), conforte en partie le bien-fondé de ces contestations, mais il permet de concevoir un système de qualification qui tienne compte explicitement de cette incertitude, qui s'en protège et qui en établisse la portée devant le candidat, l'employeur et la tierce partie.

Annexe

Norme exigeante unilatérale (λ_E). La double intégrale (9) concerne les statistiques $\overline{X}$ et s, qui sont mutuellement indépendantes sous l'hypothèse d'une distribution normale de X. Pour des valeurs données de $\overline{X}$ et de t, $\Pr\{ L_S \geq +t\,s \}$ devient $P(\overline{X}, t) = \Pr\{ s \leq (L_S - \overline{X}) \,/\, t \} = \Pr\{ (n-1)s^2/\sigma^2 \leq (n-1)(L_S - \overline{X})^2/(\sigma^2 t^2)\}$ $= \Pr\{ \chi^2 \leq (n-1)(L_S - \overline{X})^2/(\sigma^2 t^2)\}$, expression qui s'évalue aisément comme une intégrale du Khi-deux (χ^2) avec $n - 1$ degrés de liberté. Il faut ensuite évaluer $P(\overline{X}, t)$ pour tout le domaine de $\overline{X}$, soit :

$$(12) \qquad P(t) = \int \mathrm{P}(\overline{X}, t) f(\overline{X})\, d\overline{X},$$

où $f(\mathrm{X})$ est la densité d'une moyenne de n données normales, soit $e^{-n\overline{X}^2/2} / \sqrt{2\pi/n}$. Il reste à trouver par repérage la norme $t = \lambda_E$ telle que $P(t) = \alpha$. Johnson et al. (1994, 1995)[7] proposent pour ce cas une méthode plus directe, qui exploite la variable non centrale t de Student, soit $t'(\delta)$, où δ est le paramètre de non-centralité. Il s'agit alors de trouver λ_E telle que :

(13) $\Pr\{t'_{n\text{-}1}(\delta) < \lambda_E \sqrt{n}\} = 1 - \alpha$,

en utilisant $\delta = z_{1-f}\sqrt{n}$.

Norme permissive unilatérale (λ_P). Il s'agit ici, pour une personne qualifiée par $X \geq L_S$, d'assurer qu'elle soit retenue avec une probabilité d'au moins α, soit symboliquement :

(14) $\min \Pr\{ X \geq \overline{X} + \lambda_P\, s \mid X \geq L_S \} = 1 - \alpha$.

Le minimum cherché étant atteint lorsque $X \to L_S$, nous voyons que la valeur cherchée λ_P est de même espèce que λ_E ci-dessus, en remplaçant simplement $1 - \alpha$ par α.

Norme exigeante bilatérale ($\pm\lambda_E$). Pour une personne n'occupant pas les fractions extrêmes, soit $L_I < X < L_S$, où $L_S = -L_I = \mu + z_{1-f/2}\,\sigma$, il faut contrôler la probabilité qu'elle se voie qualifiée par le critère empirique $X \geq L_S'$ ou bien par $X \leq L_I'$. L'expression (11), plus haut, est rendue maximale lorsque $X = L_S$ ou L_I. En passant en valeurs standardisées ($\mu = 0$, $\sigma = 1$), l'expression (11) devient :

(15a) $\max \Pr\{ z_{1-f} \geq \overline{X} + ts \text{ ou } -z_{1-f} \leq \overline{X} - ts \} = \alpha$

ou, plus simplement :

(15b) $\Pr\{ z_{1-f} \geq |\overline{X}| + ts \} = \alpha$.

Cette expression peut être traitée comme ci-dessus, en exploitant d'abord la fonction de répartition du Khi-deux afin d'obtenir une fonction $P(|\overline{X}|, t)$, puis la probabilité $P(t)$ en intégrant la précédente sur le domaine de $|\overline{X}|$, avec la densité $e^{-n\overline{X}^2/2} / \sqrt{\pi/(2n)}$. La valeur de t qui donne l'égalité $P(t) = \alpha$ définit la norme bilatérale exigeante, $\pm\lambda_E = \pm\, t$, qui assure que la personne classée dans l'une des deux fractions extrêmes, de taille $f/2$, de la population y appartient vraiment.

Norme permissive bilatérale ($\pm\lambda_P$). Pour une personne qui loge dans la fraction $f/2$ supérieure ou la fraction $f/2$ inférieure de la population, la norme permissive assure qu'elle soit classée « extrême » avec une probabilité suffisante. Symboliquement,

(16) $\min \Pr\{X \geq \overline{X} + \lambda_P\, s \text{ ou } X \leq \overline{X} - \lambda_P\, s \mid X \geq \mu + z_{1-f/2}\,\sigma \text{ ou}$
$X \leq \mu - z_{1-f/2}\,\sigma \} = 1 - \alpha$.

Le même raisonnement que celui proposé pour la norme permissive unilatérale s'applique ici, de sorte qu'on obtient les normes permissives $\pm\lambda_P$ par le procédé indiqué pour les normes exigeantes $\pm\lambda_E$, en égalisant toutefois sur $1 - \alpha$ plutôt que sur α.

Calculs approximatifs. La valeur d'une norme sûre, basée comme ici sur le modèle normal, peut être grossièrement estimée par une approximation normale exploitant les moments (7a) et (7b). Prenons l'exemple d'une norme unilatérale exigeante basée sur la moyenne et l'écart-type de $n = 200$ données. La fraction concernée serait de $f = 0{,}10$, avec une sûreté $1 - \alpha = 0{,}95$. La valeur (standard) cherchée est estimée par :

$$(17) \qquad \lambda_E(f, n, \alpha) \approx z_{1-f} + z_{1-f}/(4n{-}4) + z_{1-\alpha}\sqrt{\frac{1 + z_{1-f}^2/2}{n}}$$

$$\approx 1{,}282 + 1{,}282/796 + 1{,}645\times\sqrt{(1+1{,}282^2/2)/200}$$

$$\approx 1{,}441\ ,$$

alors que la valeur exacte, au tableau 1, est de 1,450. L'estimation de la norme permissive λ_P pour les mêmes circonstances, obtenue en utilisant $-z_{1-\alpha}$ plutôt que $z_{1-\alpha}$, fournit 1,126 comparativement à 1,133.

NOTES

1. Nous considérons ici la sélection d'éléments faisant partie de la portion supérieure de la population, en mode de sélection exigeant. D'autres stratégies et d'autres modes sont possibles et intéressants. Nous y revenons brièvement, plus loin.
2. D'autres estimateurs de normes sont possibles, notamment ceux basés sur les statistiques d'ordre de l'échantillon normatif. Voir par exemple Laurencelle (1998, 2000).
3. Au contraire de la variance d'échantillon s^2 et de la moyenne $\overline{X}$, l'écart-type s est un estimateur biaisé de son paramètre σ, le biais étant à peu près de $-\sigma / (4n{-}4)$ pour une variable normale. Ce biais joue dans l'estimateur (3) et il est concevable d'y porter remède. Cependant, le biais de l'estimateur s est englouti dans sa variabilité, l'erreur-type de s étant d'ordre $\sigma/\sqrt{2n}$, la correction se révélant donc peu efficace.
4. Laurencelle (2000) présente des « limites de tolérance » bilatérales telles que celles retrouvées chez les autres auteurs. Cependant ces limites ont une portée statistique différente (elles sont telles que la probabilité de capturer une fraction d'au moins $1-f$ de la population est contrôlée) et ne conviennent pas au contexte présent, pour lequel la qualification tient au fait d'appartenir ou non à une fraction extrême, soit la fraction $f/2$ supérieure, soit la fraction $f/2$ inférieure, de la population. Il appert que les auteurs, et Laurencelle (1998, 2000) lui-même, sont passés à côté de cette définition. Par contraste avec « l'intervalle de tolérance » flottant retrouvé dans la documentation, nous proposons ici un « intervalle ancré », qui prend pour référence une hypothétique borne fixe (L_S) dans la population. Les développements en annexe précisent complètement cet intervalle ancré.
5. Dans *Tables psychométriques expliquées et appliquées*, à paraître du même auteur.

6. La Cour suprême du Canada, présidée par l'honorable juge Antonio Lamer, a jeté un regard sévère sur l'utilisation de tests à des fins de qualification de personnel, notamment en ce qui touche de possibles biais discriminatoires. Référence est faite ici à la cause de la Colombie-Britannique (Public Service Employee Relations Commission) *c.* BCGSEU, parue dans les *Rapports de la Cour suprême du Canada*, 1999, vol. 3, p. 3-107.
7. Les formules données dans la référence contiennent de petites erreurs, notamment la formule 13.107 (tome 1, p. 142), qui devrait se lire : $t'_{n-1,\alpha}(-\sqrt{n}\mathrm{U}_\gamma)/\sqrt{n}$.

RÉFÉRENCES

Burr, I.W. (1979). *Elementary quality control*. New York : Marcel Dekker.

Johnson, N.L., Kotz, S. & Balakrishnan, N. (1994, 1995). *Continuous univariate distributions* (2 tomes, 2e édition). New York : Wiley.

Kendall, M.G. & Stuart, A. (1977). *The advanced theory of statistics. Volume 1 : Distribution theory* (4e édition). New York : Macmillan.

Laurencelle, L. (1998). *Théorie et techniques de la mesure instrumentale*. Sainte-Foy : Presses de l'Université du Québec.

Laurencelle, L. (2000). L'incertitude des seuils statistiques et les limites de tolérance, avec des applications en psychométrie. *Lettres Statistiques*, *11*, 1-29.

Odeh, R.E. & Owen, D.B. (1980). *Tables for normal tolerance limits, sampling plans, and screening*. New York : Marcel Dekker.