



# PROJET DE NORME INTERNATIONALE ISO/DIS 13528

ISO/TC 69/SC 6

Secrétariat: JISC

Début de vote  
2013-01-22

Vote clos le  
2013-04-22

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION • МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ • ORGANISATION INTERNATIONALE DE NORMALISATION

## Méthodes statistiques utilisées dans les essais d'aptitude par comparaisons interlaboratoires

*Statistical methods for use in proficiency testing by interlaboratory comparisons*

[Révision de la première édition (ISO 13528:2005)]

ICS 03.120.30

**ITeH STANDARD PREVIEW**  
(standards.iteh.ai)  
Full standard:  
<http://standards.iteh.ai/catalog/standards/sist/5ad38656-4adfa-31ab-40de2739/iso-13528-2015>

**Pour accélérer la distribution, le présent document est distribué tel qu'il est parvenu du secrétariat du comité. Le travail de rédaction et de composition de texte sera effectué au Secrétariat central de l'ISO au stade de publication.**

**To expedite distribution, this document is circulated as received from the committee secretariat. ISO Central Secretariat work of editing and text composition will be undertaken at publication stage.**

CE DOCUMENT EST UN PROJET DIFFUSÉ POUR OBSERVATIONS ET APPROBATION. IL EST DONC SUSCEPTIBLE DE MODIFICATION ET NE PEUT ÊTRE CITÉ COMME NORME INTERNATIONALE AVANT SA PUBLICATION EN TANT QUE TELLE.

OUTRE LE FAIT D'ÊTRE EXAMINÉS POUR ÉTABLIR S'ILS SONT ACCEPTABLES À DES FINS INDUSTRIELLES, TECHNOLOGIQUES ET COMMERCIALES, AINSI QUE DU POINT DE VUE DES UTILISATEURS, LES PROJETS DE NORMES INTERNATIONALES DOIVENT PARFOIS ÊTRE CONSIDÉRÉS DU POINT DE VUE DE LEUR POSSIBILITÉ DE DEVENIR DES NORMES POUVANT SERVIR DE RÉFÉRENCE DANS LA RÉGLEMENTATION NATIONALE.

LES DESTINATAIRES DU PRÉSENT PROJET SONT INVITÉS À PRÉSENTER, AVEC LEURS OBSERVATIONS, NOTIFICATION DES DROITS DE PROPRIÉTÉ DONT ILS AURAIENT ÉVENTUELLEMENT CONNAISSANCE ET À FOURNIR UNE DOCUMENTATION EXPLICATIVE.

**iTeh STANDARD PREVIEW**  
(standards.iteh.ai)  
Full standard:  
<https://standards.iteh.ai/catalog/standards/sist/5ad38656-92e5-4ea4-adfa-31ab40de2739/iso-13528-2015>



## DOCUMENT PROTÉGÉ PAR COPYRIGHT

© ISO 2013

Droits de reproduction réservés. Sauf indication contraire, aucune partie de cette publication ne peut être reproduite ni utilisée sous quelque forme que ce soit et par aucun procédé, électronique ou mécanique, y compris la photocopie, l'affichage sur l'internet ou sur un Intranet, sans autorisation écrite préalable. Les demandes d'autorisation peuvent être adressées à l'ISO à l'adresse ci-après ou au comité membre de l'ISO dans le pays du demandeur.

ISO copyright office  
Case postale 56 • CH-1211 Geneva 20  
Tel. + 41 22 749 01 11  
Fax + 41 22 749 09 47  
E-mail [copyright@iso.org](mailto:copyright@iso.org)  
Web [www.iso.org](http://www.iso.org)

Publié en Suisse

## Sommaire

Page

Avant-propos .....	v
0 Introduction.....	vi
0.1 Objectifs des essais d'aptitude.....	vi
0.2 ISO/CEI 17043 .....	vi
0.3 Expertise en statistique .....	vii
0.4 Logiciels .....	vii
1 Domaine d'application .....	1
2 Références normatives .....	1
3 Termes et définitions .....	2
4 Lignes directrices pour le plan statistique de programmes d'essais d'aptitude .....	4
4.1 Introduction.....	4
4.2 Base d'un plan statistique.....	4
4.3 Considérations sur la distribution statistique des résultats .....	5
4.4 Considérations sur les petits nombres de participants.....	6
4.5 Considérations sur les données qualitatives (y compris les propriétés nominales et ordinales) .....	7
4.6 Lignes directrices pour choisir le format de rapport.....	8
5 Lignes directrices pour la revue initiale des entités soumises à l'essai d'aptitude et des résultats.....	9
5.1 Homogénéité et stabilité des entités soumises à l'essai d'aptitude.....	9
5.2 Considérations lorsque des méthodes différentes sont utilisées par les participants .....	10
5.3 Elimination des valeurs aberrantes.....	11
5.4 Examen visuel des données .....	11
5.5 Méthodes statistiques robustes .....	12
5.6 Techniques de détection des valeurs aberrantes pour des résultats individuels .....	12
6 Détermination de la valeur assignée et de son incertitude type.....	13
6.1 Choix de la méthode de détermination de la valeur assignée.....	13
6.2 Détermination de l'incertitude de la valeur assignée .....	13
6.3 Formulation.....	14
6.4 Matériau de référence certifié .....	15
6.5 Résultats provenant d'un laboratoire.....	15
6.6 Résultats issus d'une étude réalisée par des laboratoires experts.....	16
6.7 Résultats de mesure combinés fournis par les participants.....	17
6.8 Comparaison de la valeur assignée avec une valeur de référence indépendante.....	18
7 Détermination des critères d'évaluation des performances.....	18
7.1 Base de l'évaluation .....	18
7.2 Méthode par perception d'experts.....	19
7.3 Par l'expérience acquise lors des campagnes précédentes .....	20
7.4 Détermination à partir d'un modèle général .....	21
7.5 Utilisation des écarts-types de répétabilité et de reproductibilité obtenus lors d'une étude collaborative .....	21
7.6 Détermination à partir de données obtenues lors d'une même campagne d'un programme d'essais d'aptitude .....	22
7.7 Comparaison entre les valeurs de fidélité déduites d'un programme d'essais d'aptitude et les valeurs établies.....	22

8	Calcul des statistiques de performances.....	23
8.1	Considérations générales pour la détermination des performances.....	23
8.2	Estimations des écarts (erreur de mesure).....	23
8.3	Scores $z$ .....	25
8.4	Scores $z'$ .....	26
8.5	Scores zêta ( $\zeta$ ).....	27
8.6	Scores $E_n$ .....	28
8.7	Utilisation des rangs et des rangs en percentiles.....	29
8.8	Interprétation des incertitudes des participants lors des essais .....	29
8.9	Scores composés de performance .....	30
9	Méthodes graphiques de description des scores de performance à partir d'une campagne d'un essai d'aptitude .....	32
9.1	Application .....	32
9.2	Histogrammes de résultats ou de scores de performance .....	32
9.3	Diagrammes en bâtons des scores normalisés .....	33
9.4	Graphique de Youden .....	33
9.5	Graphiques d'écart-types de répétabilité .....	34
9.6	Echantillons fractionnés .....	35
9.7	Méthodes graphiques de combinaison des scores de performance sur plusieurs campagnes d'un programme d'essais d'aptitude .....	36
	Annexe A (normative) Symboles.....	38
	Annexe B (normative) Vérification de l'homogénéité et de la stabilité des échantillons.....	39
	Annexe C (normative) Analyse robuste.....	45
	Annexe D (informative) Exemples illustratifs .....	52
	Bibliographie .....	89

iTeh STANDARD PREVIEW  
 (standards.iteh.ai)  
 Full standard:  
<https://standards.iteh.ai/catalog/standards/sist/22d13528-56-92e5-4ea4-adfa-31ab40de2739/iso-13528-2013>

## Avant-propos

L'ISO (Organisation internationale de normalisation) est une fédération mondiale d'organismes nationaux de normalisation (comités membres de l'ISO). L'élaboration des Normes internationales est en général confiée aux comités techniques de l'ISO. Chaque comité membre intéressé par une étude a le droit de faire partie du comité technique créé à cet effet. Les organisations internationales, gouvernementales et non gouvernementales, en liaison avec l'ISO participent également aux travaux. L'ISO collabore étroitement avec la Commission électrotechnique internationale (CEI) en ce qui concerne la normalisation électrotechnique.

Les Normes internationales sont rédigées conformément aux règles données dans les Directives ISO/CEI, Partie 2.

La tâche principale des comités techniques est d'élaborer les Normes internationales. Les projets de Normes internationales adoptés par les comités techniques sont soumis aux comités membres pour vote. Leur publication comme Normes internationales requiert l'approbation de 75 % au moins des comités membres votants.

L'attention est appelée sur le fait que certains des éléments du présent document peuvent faire l'objet de droits de propriété intellectuelle ou de droits analogues. L'ISO ne saurait être tenue pour responsable de ne pas avoir identifié de tels droits de propriété et averti de leur existence.

L'ISO 13528 a été élaborée par le comité technique ISO/TC 69, *Application des méthodes statistiques*, sous-comité SC 6, *Méthodes et résultats de mesure*. La première édition a été publiée en 2005. La seconde édition apporte des modifications destinées à harmoniser le document avec l'ISO/CEI 17043:2010, qui a remplacé le Guide ISO 43-1. Le titre a été raccourci, car la mention « **comparaison interlaboratoires** » est incluse dans la définition de « **essais d'aptitude** ». La deuxième édition suit une structure révisée afin de mieux refléter le processus de conception, d'analyse et de compte-rendu des programmes d'essais d'aptitude. Elle a également supprimé certaines procédures et ajouté ou modifié certains autres paragraphes afin de mieux refléter l'ISO/CEI 17043 et d'apporter plus de clarté et de corriger des erreurs mineures. De nouveaux paragraphes ont été ajoutés pour les données qualitatives et pour d'autres méthodes robustes.

## 0 Introduction

### 0.1 Objectifs des essais d'aptitude

Les essais d'aptitude impliquent l'utilisation de comparaisons interlaboratoires pour déterminer les performances de participants (qui peuvent être des laboratoires, des organismes de contrôle, ou des individus) relatives à des essais ou à des mesures spécifiques et pour surveiller la continuité des performances des participants. Il existe un certain nombre d'objectifs pour les essais d'aptitude, tels que décrits dans l'Introduction de l'ISO/CEI 17043:2010. Ces objectifs comprennent l'évaluation des performances des laboratoires, l'identification de problèmes dans les laboratoires, la détermination de l'efficacité et la possibilité de comparer les méthodes d'essai ou de mesure, l'amélioration de la confiance des clients des laboratoires, la validation des incertitudes revendiquées et la formation des laboratoires participants.

Le plan et les techniques statistiques appliqués doivent être appropriés à ou aux objectifs définis. Cette Norme présente des techniques statistiques pour les objectifs décrits, qui sont conformes à d'autres normes internationales, notamment à celles du TC69/SC6, et en particulier à la série de normes ISO 5725 portant sur l'*Exactitude : justesse et fidélité*. Les techniques sont également destinées à refléter les techniques préconisées par d'autres normes internationales, le cas échéant, et à être en cohérence avec le Guide ISO 98 et avec le JCGM 200 (VIM).

La définition de l'essai d'aptitude donnée dans l'ISO/CEI 17043 est reprise dans l'ISO 13528, avec des notes décrivant les différents types d'essais d'aptitude et l'étendue de la complexité des conceptions pouvant être utilisées. La présente norme ne peut pas couvrir spécifiquement l'ensemble des conceptions, des objectifs, des matrices et des mesurandes. Les techniques présentées dans l'ISO 13528 sont destinées à être appliquées de façon étendue, notamment pour des programmes d'essais d'aptitude nouvellement établis. On s'attend à ce que les techniques statistiques utilisées dans le cadre d'un programme d'essais d'aptitude particulier évoluent au fur et à mesure que le programme gagnera en maturité ; et les scores, les critères d'évaluation et les méthodes graphiques seront affinées afin de mieux répondre aux besoins spécifiques d'un groupe cible de participants, d'organismes d'accréditation et d'organismes réglementaires.

L'ISO 13528 applique également des lignes directrices issues d'un protocole harmonisé pour les essais d'aptitude de laboratoires d'analyses chimiques [20], mais elle est destinée à être utilisée avec toutes les méthodes de mesure. La présente révision de l'ISO 13528:2005 contient la plupart des informations issues de la première édition, complétées autant que nécessaire par les documents précédemment cités en référence, par le domaine d'application étendu de l'ISO/CEI 17043, qui comprend des essais d'aptitude pour le contrôle, et l'Annexe B qui comprend des éléments d'appréciation pour les résultats qualitatifs et ordinaux. Certaines procédures issues de la première édition n'ont pas été reprises dans la présente édition, car l'expérience a démontré que de telles techniques ne sont plus considérées comme appropriées.

### 0.2 ISO/CEI 17043

La présente norme soutiendra la mise en œuvre de l'ISO/CEI 17043 qui décrit différents types de programmes d'essais d'aptitude et donne des lignes directrices pour l'organisation et la conception de programmes d'essais d'aptitude. L'Annexe C de l'ISO/CEI 17043 donne des lignes directrices pour la sélection et l'utilisation de programmes d'essais d'aptitude par les laboratoires, les organismes de contrôle, les organismes d'accréditation et autres parties intéressées. Il convient de consulter l'ISO/CEI 17043 pour avoir des informations détaillées sur ces questions, car les informations qui y figurent ne sont pas reprises ici.

L'Annexe B de l'ISO/CEI 17043 fournit une description succincte des méthodes statistiques générales utilisées dans les programmes d'essais d'aptitude. La présente Norme internationale est destinée à être complémentaire de l'Annexe B, en fournissant des indications détaillées ne figurant pas dans ce document sur des méthodes particulières d'essais d'aptitude.

### 0.3 Expertise en statistique

L'ISO/CEI 17043:2010 exige que, pour être compétent, un organisateur d'essais d'aptitude doit avoir accès à l'expertise statistique (4.4.1.4) et doit autoriser du personnel spécifique à mener l'analyse statistique (4.2.4). Ni l'ISO/CEI 17043 ni le présent document ne peuvent définir plus en détail en quoi consiste cette expertise nécessaire. Pour certaines applications, un niveau élevé en statistiques est utile, mais généralement les besoins en expertise peuvent être assurés par des personnes ayant une expertise technique dans d'autres domaines, si celles-ci connaissent les concepts et les techniques statistiques de base. Si un consultant est chargé du plan et/ou de l'analyse statistique, il est très important que cette personne ait une expérience dans le domaine des comparaisons interlaboratoires, même si elle a un niveau élevé dans le domaine des statistiques. La formation classique en statistiques avancées n'inclut pas l'acquisition de connaissances dans le domaine des comparaisons interlaboratoires, ni les causes d'erreur de mesure dans les essais d'aptitude. Les lignes directrices données dans la présente Norme internationale ne peuvent pas fournir toute l'expertise nécessaire pour prendre en compte toutes les applications, et ne peuvent pas remplacer l'expérience acquise lors de la pratique des comparaisons interlaboratoires.

### 0.4 Logiciels

Les logiciels nécessaires pour l'analyse statistique peuvent varier de façon importante, allant de l'absence totale de logiciel (par exemple, concernant des programmes d'essais d'aptitude pour des participants individuels, utilisant des valeurs de référence connues et des critères d'évaluation de l'adéquation à l'objectif), jusqu'à des logiciels statistiques sophistiqués utilisés dans le cadre de programmes d'essais d'aptitude avec des nombres importants de participants et de mesurandes, ou utilisant des méthodes robustes compliquées. La plupart des techniques mentionnées dans la présente Norme internationale peuvent être mises en œuvre par des tableurs classiques, éventuellement avec une analyse personnalisée pour un programme ou une analyse spéciale ; certaines techniques nécessitent des logiciels gratuits (au moment de la publication). Dans tous les cas, l'utilisateur est tenu de vérifier l'exactitude de leurs calculs, en particulier lorsque des programmes spéciaux ont été entrés par l'utilisateur. Toutefois, même lorsque les techniques traitées dans la présente Norme internationale sont appropriées et correctement mises en œuvre par des logiciels adéquats, elles ne peuvent pas être appliquées sans la surveillance expérimentée d'une personne disposant d'une expertise technique et statistique suffisante pour identifier et rechercher des anomalies susceptibles de se produire lors de toute campagne d'essais d'aptitude.

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

Full standard:  
<https://standards.iteh.ai/catalog/standards/sist/5ad38656-92e5-4ea4-adfa-31ab40de2739/iso-13528-2015>



# Méthodes statistiques utilisées dans les essais d'aptitude par comparaisons interlaboratoires

## 1 Domaine d'application

La présente Norme internationale fournit des descriptions détaillées de méthodes statistiques à utiliser par les organisateurs pour concevoir des programmes d'essais d'aptitude et pour analyser les données obtenues à partir de tels programmes ; elle donne également des recommandations sur l'utilisation pratique de techniques statistiques par les participants à de tels programmes et par les organismes d'accréditation.

La présente Norme internationale peut être appliquée pour démontrer que les résultats de mesure obtenus par des laboratoires, des organismes de contrôle et des personnes remplissent les critères spécifiés pour des performances acceptables.

La présente Norme internationale est applicable à des essais d'aptitude dans lesquels les résultats consignés sont des observations quantitatives ou qualitatives sur des entités soumises à l'essai.

**NOTE** Les procédures énoncées dans la présente norme peuvent être également applicables pour l'évaluation d'un avis d'expert lorsque les avis ou les jugements sont consignés sous une forme qui peut être comparée de manière objective avec une valeur de référence indépendante ou avec une statistique combinée de résultats de participants. Par exemple, lorsque des entités soumises à essai sont classées en catégories connues par contrôle – ou lorsque l'on cherche à déterminer par contrôle si des entités soumises à essai proviennent ou non de la même source d'origine – et lorsque les résultats de la classification sont comparés de manière objective.

## 2 Références normatives

Les documents de référence suivants sont indispensables pour l'application du présent document. Pour les références datées, seule l'édition citée s'applique. Pour les références non datées, la dernière édition du document de référence (y compris les éventuels amendements) s'applique.

ISO 3534-1, Statistique — Vocabulaire et symboles — Partie 1 : Termes statistiques généraux et termes utilisés en calcul des probabilités.

ISO 3534-2, Statistique — Vocabulaire et symboles — Partie 2 : Statistique appliquée.

ISO 5725-1, Exactitude (justesse et fidélité) des résultats et méthodes de mesure — Partie 1 : Principes généraux et définitions.

ISO/CEI 17043, Evaluation de la conformité — Exigences générales concernant les essais d'aptitude.

JCGM 200:2012, Vocabulaire international de métrologie — Concepts fondamentaux et généraux et termes associés (VIM), 3ème édition.

ISO/CEI Guide 98-3, Incertitude de mesure — Partie 3 : Guide pour l'expression de l'incertitude de mesure (GUM:1995).

### 3 Termes et définitions

Pour les besoins du présent document, les termes et définitions donnés dans l'ISO 3534-1, l'ISO 3534-2, l'ISO 5725-1, l'ISO/CEI 17043, le JCGM 200, l'ISO Guide 34, ainsi que les suivants s'appliquent.

**3.1 comparaison interlaboratoires**  
organisation, exécution et évaluation d'essais ou de mesures sur la même entité ou sur des entités similaires par deux laboratoires ou plus dans des conditions prédéterminées

**3.2 essai d'aptitude**  
évaluation de la performance d'un participant par rapport à des critères préétablis au moyen de comparaisons interlaboratoires

NOTE Pour les besoins de la présente Norme internationale, le terme « essai d'aptitude » est considéré dans son sens le plus large et il inclut sans s'y limiter :

- a) les programmes quantitatifs, dans lesquels l'objectif est de quantifier un ou plusieurs mesurandes de l'entité soumise à l'essai d'aptitude ;
- b) les programmes qualitatifs, dans lesquels l'objectif est d'identifier ou de décrire une ou plusieurs caractéristiques de l'entité soumise à l'essai d'aptitude ;
- c) les programmes séquentiels, dans lesquels une ou plusieurs entités soumises à l'essai d'aptitude sont distribuées séquentiellement pour procéder à l'essai ou au mesurage, et reviennent par intervalles à l'organisateur d'essais d'aptitude ;
- d) les programmes simultanés, dans lesquels les entités soumises à l'essai d'aptitude sont réparties en vue d'essais ou de mesurages concurrents dans une période de temps définie ;
- e) les exercices de situation unique, dans lesquels les entités soumises à l'essai d'aptitude sont fournies à une seule occasion ;
- f) les programmes continus, dans lesquels les entités soumises à l'essai d'aptitude sont fournies à intervalles réguliers ;
- g) les échantillonnages, dans lequel des échantillons sont prélevés en vue d'une analyse ultérieure ; et
- h) les transformations et interprétations de données, dans lesquelles des ensembles de données ou d'autres informations sont fournis et les informations sont traitées pour en effectuer une interprétation (ou un autre résultat).

**3.3 valeur assignée**  
valeur attribuée à une propriété particulière d'une entité soumise à l'essai d'aptitude

**3.4 écart-type pour une évaluation de l'aptitude**  
mesure de la dispersion utilisée dans l'évaluation des résultats d'un essai d'aptitude, fondée sur les informations disponibles

NOTE 1 Cela peut être interprété comme un écart-type cible pour une population de laboratoires compétents pour réaliser une procédure de mesure particulière.

NOTE 2 L'écart-type pour l'évaluation de l'aptitude ne s'applique qu'aux résultats sur une échelle linéaire.

NOTE 3 Les programmes d'essais d'aptitude n'évaluent pas tous l'aptitude en se fondant sur la dispersion des résultats.

NOTE 4 Les notes relatives à la définition dans le présent document ne sont pas identiques aux celles de la définition donnée dans l'ISO/CEI 17043.

**3.5****score z**

mesure normalisée de la performance, calculée à partir du résultat du participant, de la valeur assignée et de l'écart-type pour l'évaluation de l'aptitude

**3.6****valeur aberrante**

observation dans un ensemble de données, semblant incohérente avec le reste de cet ensemble

NOTE 1 Une valeur aberrante peut provenir d'une population différente ou résulter d'un enregistrement incorrect ou d'une autre erreur grossière.

NOTE 2 De nombreux programmes utilisent le terme « valeur aberrante » pour désigner un résultat qui génère un signal d'action. Il ne s'agit pas là de l'utilisation prévue du terme. Bien que les valeurs aberrantes génèrent habituellement des signaux d'action, il est possible d'avoir des signaux d'action issus de résultats qui ne sont pas considérés comme des valeurs aberrantes.

**3.7****participant**

laboratoire, organisme ou particulier, recevant les entités soumises à l'essai d'aptitude et soumettant les résultats à l'organisateur d'essais d'aptitude

**3.8****entité soumise à l'essai d'aptitude**

échantillon, produit, artefact, matériau de référence, élément d'un matériel, étalon, ensemble de données ou autres informations utilisées pour un essai d'aptitude

NOTE Dans la plupart des cas, les entités soumises à l'essai d'aptitude sont conformes à la définition de « matériau de référence » donnée dans le Guide ISO 34 (voir 3.11 ci-dessous)

**3.9****organisateur d'essais d'aptitude**

organisme ayant la responsabilité de toutes les tâches d'élaboration et d'exécution d'un programme d'essais d'aptitude

**3.10****programme d'essais d'aptitude**

essai d'aptitude conçu et exécuté en une ou plusieurs campagnes dans un domaine spécifié d'essai, de mesurage, d'étalonnage ou d'inspection

NOTE Un programme d'essais d'aptitude peut recouvrir un type particulier d'essai, d'étalonnage, d'inspection ou un certain nombre d'essais, d'étalonnages ou d'inspections sur des entités soumises à l'essai d'aptitude.

**3.11****matériau de référence (MR)**

matériau suffisamment homogène et stable en ce qui concerne des propriétés spécifiées, qui a été préparé pour être adapté à son utilisation prévue pour un mesurage (Guide ISO 34:2009)

NOTE 1 MR est un terme générique.

NOTE 2 Les propriétés peuvent être quantitatives ou qualitatives (par exemple, identité de substances ou d'espèces).

NOTE 3 Les utilisations peuvent comprendre l'étalonnage d'un appareil de mesure, l'évaluation d'une méthode de mesurage, l'attribution de valeurs à d'autres matériaux et la maîtrise de la qualité.

### 3.12

#### **matériau de référence certifié (MRC)**

matériau de référence caractérisé par une procédure métrologiquement valide pour une ou plusieurs propriétés spécifiées, accompagné d'un certificat fournissant la valeur de la propriété spécifiée, son incertitude associée, et une déclaration de traçabilité métrologique (Guide ISO 34:2009)

NOTE Le concept de valeur comprend des attributs qualitatifs tels qu'identité ou séquence. Les incertitudes associées à ces attributs peuvent être exprimées en probabilités.

## **4 Lignes directrices pour le plan statistique de programmes d'essais d'aptitude**

### **4.1 Introduction**

Les essais d'aptitude portent sur l'évaluation des performances des participants et, en tant que tels, ils ne traitent pas spécifiquement de biais ou de fidélité (bien qu'ils puissent être évalués avec des plans spécifiques). Les performances des participants sont considérées à travers l'évaluation statistique de leurs résultats suite aux mesurages ou aux interprétations qu'ils effectuent sur les entités soumises à l'essai d'aptitude. Les performances sont souvent exprimées sous forme de scores permettant une interprétation cohérente sur une gamme de mesurandes et pouvant permettre la comparaison, sur un même pied d'égalité, des résultats relatifs à différents mesurandes. Les scores de performance sont généralement obtenus en comparant la différence entre les résultats fournis par les participants et une valeur assignée avec une dispersion admissible ou avec une mesure de l'incertitude de la différence. L'examen des scores de performances sur de nombreux mesurandes ou sur de nombreuses campagnes d'un programme d'essais d'aptitude peut fournir des informations permettant de savoir si des laboratoires démontrent la présence ou non d'effets systématiques cohérents (« biais ») ou d'une fidélité médiocre à long terme.

Les paragraphes suivants donnent des indications sur la conception des programmes d'essais d'aptitude et sur les traitements statistiques des résultats, y compris le calcul et l'interprétation de divers scores de performances.

### **4.2 Base d'un plan statistique**

**4.2.1** Selon l'ISO/CEI 17043, paragraphe 4.4.4.1, le plan statistique « doit être élaboré pour répondre aux objectifs du programme, sur la base de la nature des données (quantitatives ou qualitatives, y compris ordinales et nominales), des hypothèses statistiques, de la nature des erreurs et du nombre attendu de résultats ». Par conséquent, des programmes avec des objectifs différents et des sources d'erreur différentes peuvent avoir des plans différents.

NOTE Par exemple, des programmes avec les principales approches suivantes peuvent avoir des plans statistiques différents :

- un programme pour comparer les résultats des participants avec des résultats combinés issus d'un groupe au cours de la même campagne, et les limites déterminées par la variabilité des résultats des participants ;
- un programme pour comparer un résultat de participant avec une valeur de référence prédéterminée, dans des limites spécifiées avant le début de la campagne ;
- un programme pour comparer les résultats des participants leurs incertitudes de mesure .

Des programmes d'essais d'aptitude ayant d'autres objectifs (peut-être secondaires) peuvent également nécessiter des techniques statistiques spécifiques, comme dans le cas où un objectif consiste à comparer les performances de différentes méthodes de mesure.

**4.2.2** Divers types de données sont utilisés dans les essais d'aptitude, y compris des données quantitatives, nominales (catégoriques) et ordinales. Parmi les variables continues, certains résultats peuvent se situer sur une échelle différentielle ou rationnelle (une échelle avec un zéro arbitraire), ou sur une échelle relative ou linéaire (une échelle avec un zéro vrai). Pour certaines mesures sur une échelle continue, seul un ensemble discret et discontinu de valeurs peut être réalisé ; toutefois, dans de nombreux cas, ces résultats peuvent être traités comme d'autres variables continues.

**4.2.3** Les programmes peuvent être utilisés pour d'autres objectifs, en plus de ceux mentionnés ci-dessus, comme évoqué dans le paragraphe 0.1 et dans l'ISO/CEI 17043. Le plan doit être approprié pour toutes les utilisations déterminées.

### 4.3 Considérations sur la distribution statistique des résultats

#### 4.3.1 Hypothèse statistique pour la distribution

**4.3.1.1** Il doit y avoir une cohérence entre le plan statistique et les méthodes d'analyse statistiques des données d'une part et les hypothèses statistiques concernant les données (voir l'ISO/CEI 17043, 4.4.4.2) d'autre part. Pour les essais d'aptitude, les techniques d'analyse les plus courantes supposent qu'un ensemble de résultats fourni par des participants compétents sera unimodal, raisonnablement symétrique et approximativement distribué selon une loi normale. Il existe une hypothèse supplémentaire courante selon laquelle la distribution des résultats fournis par tous les participants est mélangée avec (ou « contaminée par ») des résultats issus d'une population de valeurs erronées et de valeurs aberrantes. Habituellement, l'interprétation des scores s'appuie sur l'hypothèse de normalité, mais uniquement pour la distribution sous-jacente supposée pour des participants compétents.

**4.3.1.2** En général, il n'est pas nécessaire de vérifier que les résultats suivent une distribution normale, mais il est important de vérifier, visuellement, la symétrie approximative. Si cela n'est pas possible, il convient alors que l'organisateur d'essais d'aptitude utilise des techniques robustes par rapport à l'asymétrie (voir Annexe C).

**4.3.1.3** Lorsque la distribution prévue pour le programme d'essais d'aptitude n'est pas assez proche d'une distribution normale symétrique (avec contamination par des valeurs aberrantes), il convient que l'organisateur d'essais d'aptitude choisisse des méthodes d'analyse de données qui tiennent dûment compte de l'asymétrie prévue et qui soient insensibles aux valeurs aberrantes, ainsi que des méthodes d'attribution de scores qui tiennent aussi dûment compte de la distribution prévue des résultats fournis par des participants compétents. Cela peut comprendre :

- une transformation pour obtenir une symétrie approximative ;
- des méthodes d'estimation résistantes à l'asymétrie ;
- des méthodes d'estimation qui incorporent des hypothèses de distribution (par exemple, adéquation du maximum de vraisemblance avec des hypothèses de distribution adaptées et, si nécessaire, élimination des valeurs aberrantes).

**NOTE** Par exemple, étant donné que des résultats fondés sur une dilution, comme par exemple pour les comptages microbiologiques quantitatifs ou les techniques d'immunodosages, sont souvent distribués selon la distribution normale logarithmique, une transformation logarithmique est donc appropriée en tant que première étape de l'analyse. Un autre exemple qui peut être cité concerne les comptages de petits nombres de particules ; ceux-ci peuvent être distribués selon une distribution de Poisson et, par conséquent, l'intervalle d'évaluation serait déterminé à l'aide d'une table de distribution de probabilités de Poisson, en se fondant sur le comptage moyen pour le groupe de participants.

**4.3.1.4** De nombreux mesurandes d'étalonnage suivent des distributions statistiques définies qui sont décrites dans la procédure de mesure ; il convient que ces distributions soient prises en compte dans tout protocole d'évaluation.