



TC/50/28

ORIGINAL : anglais

DATE : 30 janvier 2014

UNION INTERNATIONALE POUR LA PROTECTION DES OBTENTIONS VÉGÉTALES

Genève

COMITÉ TECHNIQUE

**Cinquantième session
Genève, 7-9 avril 2014**

RÉVISION DU DOCUMENT TGP/8 : DEUXIÈME PARTIE : TECHNIQUES UTILISÉES
DANS L'EXAMEN DHS, NOUVELLE SECTION : MÉTHODES STATISTIQUES
APPLICABLES AUX CARACTÈRES OBSERVÉS VISUELLEMENT

Document établi par le Bureau de l'Union

Avertissement : le présent document ne représente pas les principes ou les orientations de l'UPOV

1. L'objet du présent document est de faire rapport sur les faits nouveaux concernant une nouvelle section éventuelle intitulée : "Méthodes statistiques applicables aux caractères observés visuellement" aux fins de son inclusion dans le document TGP/8 : Deuxième partie : Techniques utilisées dans l'examen DHS, au titre d'une future version révisée du document TGP/8.
2. Les abréviations ci-après sont utilisées dans le présent document :

| | |
|----------|--|
| TC : | Comité technique |
| TC-EDC : | Comité de rédaction élargi du Comité technique |
| TWA : | Groupe de travail technique sur les plantes agricoles |
| TWC : | Groupe de travail technique sur les systèmes d'automatisation et les programmes d'ordinateur |
| TWF : | Groupe de travail technique sur les plantes fruitières |
| TWO : | Groupe de travail technique sur les plantes ornementales et les arbres forestiers |
| TWP : | Groupes de travail techniques |
| TWV : | Groupe de travail technique sur les plantes potagères |

INFORMATIONS GÉNÉRALES

3. À sa quarante-huitième session tenue à Genève du 26 au 28 mars 2012, le Comité technique a examiné la proposition portant sur une nouvelle section intitulée : "Méthodes statistiques applicables aux caractères observés visuellement" à incorporer dans le document TGP/8 : Deuxième partie : Techniques utilisées dans l'examen DHS sur la base de l'annexe X du document TC/48/19 Rev. "Révision du document TGP/8 : Protocole d'essai et techniques utilisées dans l'examen de la distinction, de l'homogénéité et de la stabilité", qu'avait préparée un expert du Danemark. Le TC est convenu que la section "Méthodes statistiques applicables aux caractères observés visuellement" devrait être revue avec le concours d'experts DHS du Danemark afin de cibler les conseils destinés aux examinateurs DHS et qu'elle devrait remplacer des modèles statistiques détaillés par une référence générale à des méthodes statistiques appropriées. Le TC est également convenu que les exemples fondés sur la betterave sucrière devraient être remplacés par une plante pour laquelle il y avait des principes directeurs d'examen et que l'exemple du blé devrait être remplacé par un exemple réaliste comme celui que l'on trouvait dans le chanvre ou l'épinard. Le TC est en outre convenu que le TWC devrait étudier les conséquences des décisions pour l'examen DHS car la méthode était un test permettant d'évaluer les différences dans la distribution (emplacement et dispersion). Il est enfin convenu que les conséquences de l'exclusion de certaines variétés de l'essai, dans la mesure où les chiffres étaient

insuffisants pour certaines cellules, devraient faire l'objet d'un examen plus approfondi (voir le paragraphe 61 du document TC/48/22 "Compte rendu des conclusions").

4. À sa quarante-neuvième session tenue à Genève du 18 au 20 mars 2013, le TC a examiné le document TC/49/32.

5. L'annexe II du document TC/49/32 contenait le texte proposé par le rédacteur (M. Kristian Kristensen, (Danemark)) de la nouvelle section : "Méthodes statistiques applicables aux caractères observés visuellement", sur la base des observations formulées par les TWP à leurs sessions en 2012. Les modifications à apporter au texte examiné par les TWP à leurs sessions de 2012 sont indiquées de la manière suivante : les éléments à supprimer sont surlignés et biffés et les éléments à ajouter sont surlignés et soulignés.

6. L'annexe III du document TC/49/32 contenait une copie des informations supplémentaires sur les conséquences des décisions pour l'examen DHS en tant qu'informations générales pour examen lorsque le document TWC/30/29 a été examiné par le TWC, à sa trentième session, tenue à Chisinau (République de Moldova) du 26 au 29 juin 2012 (voir le document TWC/30/19 "Consequences of Decisions for Examination of Distinctness, Uniformity and Stability").

7. À sa quarante-neuvième session, le TC est convenu qu'il ne serait pas approprié de poursuivre l'élaboration d'une section intitulée "Méthodes statistiques applicables aux caractères observés visuellement", à moins que de nouveaux conseils ne soient donnés outre les méthodes déjà fournies dans le document TGP/8. À cet égard, il a demandé au TWC de préciser s'il proposait de modifier une méthode existante ou de fournir une nouvelle méthode additionnelle (voir le paragraphe 72 du document TC/49/41 "Compte rendu des conclusions").

OBSERVATIONS DES GROUPES DE TRAVAIL TECHNIQUES EN 2013

8. À leurs sessions en 2013, les TWO, TWF, TWV, TWC et TWA ont examiné les documents TWO/46/23, TWF/44/23, TWV/47/23, TWC/31/23 Rev. et TWA/42/23 Rev., respectivement.

9. Le TWC est convenu que la méthode proposée dans l'annexe II du document TC/49/32 était une nouvelle méthode, estimant qu'elle était plus utile que le test du khi carré déjà fourni dans le document TGP/8 pour les données à distribution multinomiale telles que les caractères observés visuellement, alors que la méthode COYD pour les données à distribution normale ne convient pas aux données à distribution multinomiale (voir le paragraphe 53 du document TWC/31/32 "Report").

10. Le TWC est convenu qu'il serait utile de mettre au point la méthode pour les données multinomiales et de comparer les décisions prises à l'aide des deux méthodes, à savoir la méthode COYD pour les données à distribution normale et le test du khi carré, sur la base de données réelles de la Finlande et du Royaume-Uni (fléole, trèfle violet et fétuque des prés : port). Le TWC a également pris note que la Finlande avait l'intention d'utiliser la nouvelle méthode pour les données multinomiales une fois que celle-ci aurait été mise en place et que le Royaume-Uni pourrait lui aussi envisager de le faire (voir les paragraphes 54 et 55 du document TWC/31/32 "Report").

11. Le TWA a pris note du mémorandum figurant dans l'annexe I du document TWA/42/23 et des observations du TWC qui précisaient que la méthode proposée au TC, à sa quarante-neuvième session, pour traiter les données à distribution multinomiale était une nouvelle méthode (voir le paragraphe 60 du document TWA/42/31 "Report").

12. Le TWA est convenu avec le TWC qu'il serait avantageux de mettre au point la méthode pour les données multinomiales et de comparer les décisions prises à l'aide des deux méthodes fondées sur des données réelles de la Finlande et du Royaume-Uni (voir le paragraphe 61 du document TWA/42/31 "Report").

13. Les experts des Pays-Bas et de l'Allemagne ont fait part de leur intention d'utiliser la nouvelle méthode pour les données multinomiales une fois que celle-ci aurait été mise en place (voir le paragraphe 62 du document TWA/42/31 "Report").

14. Comme suite au départ à la retraite de M. Kristian Kristensen (Danemark), rédacteur de la nouvelle section proposée, le TWC s'efforcera de trouver un expert doté des compétences nécessaires pour poursuivre l'élaboration de la section.

15. *Le TC est invité*

a) *à approuver l'élaboration d'une nouvelle méthode pour les données à distribution multinomiale;*

b) *à inviter le TWC à comparer la méthode COYD pour les données à distribution normale et le test du khi carré, comme indiqué dans le paragraphe 10 du présent document; et*

c) *à demander au TWC de trouver un expert doté des compétences nécessaires pour servir de rédacteur.*

[L'annexe suit]

MÉ MORANDUM DE L'EXPERT DU DANEMARK

Informations générales

La méthode COYD pour les données à distribution normale a été introduite il y a maintes années à l'UPOV pour faire en sorte que les décisions prises soient plus cohérentes dans les années futures que celles prises avec la méthode classique de 2x1%. Récemment, des travaux additionnels ont été faits pour mettre au point une méthode COYD pour les données à distribution multinomiale tels que les caractères observés visuellement.

Le présent mémorandum fait une analyse des similitudes et des différences qui existent entre la méthode proposée au TC, à sa quarante-neuvième session, pour traiter les données à distribution multinomiale, et deux méthodes existantes décrites dans le document TGP/8 qui sont liées à la méthode proposée (la méthode COYD pour les données à distribution normale et le test χ^2).

Analyse des similitudes et des différences*Similitudes entre la méthode COYD pour les données à distribution normale et la méthode COYD proposées pour les données à distribution multinomiale*

- Ces deux méthodes ont pour objet de faire en sorte que les décisions prises soient cohérentes au cours des années à venir.
- Elles le font en calculant une mesure d'interaction entre les années et les variétés pour ensuite en tenir compte lorsqu'il est procédé à une comparaison des variétés.

Différences entre la méthode COYD pour les données à distribution normale et la méthode COYD proposées pour les données à distribution multinomiale

- Le résultat pour une variété donnée durant une année donnée peut être caractérisé par une seule moyenne pour les données à distribution normale alors que plusieurs valeurs (nombres ou pourcentages) doivent être utilisées pour les données à distribution multinomiale. C'est ainsi par exemple que 5 nombres ou pourcentages sont nécessaires pour caractériser une variété si la note peut prendre 5 valeurs discrètes (à proprement parler, 4 pourcentages seulement sont réellement nécessaires car la somme des 5 pourcentages devrait totaliser 100).
- Pour les données à distribution normale, la variabilité causée par un échantillonnage aléatoire est normalement réputée indépendante des moyennes et elle est par conséquent la même pour toutes les variétés. Cette hypothèse ne s'applique pas aux données à distribution multinomiale car, ici, la variation aléatoire causée par l'échantillonnage est censée dépendre de la valeur moyenne (nombre ou pourcentage) de la note réelle.
- Pour les données à distribution normale, la variabilité causée par une interaction entre les années et les variétés peut se caractériser par une seule valeur pour chaque combinaison d'année et de variété. C'est d'ordinaire aussi le cas pour les données à distribution binomiale et les caractères à distribution multinomiale s'ils sont enregistrés sur une échelle ordinale. Toutefois, dans le cas des données à distribution multinomiale enregistrées sur l'échelle nominale, un plus grand nombre de valeurs est nécessaire pour caractériser l'interaction entre les années et les variétés. C'est ainsi par exemple que, s'il y a 5 notes différentes, 4 fois le nombre de combinaisons d'années et de variétés est nécessaire.
- Pour les données à distribution normale, les variétés peuvent être comparées au moyen de tests t (ou de tests F avec un degré de liberté dans le numérateur). C'est aussi d'ordinaire le cas pour les données à distribution binomiale et les caractères à distribution multinomiale s'ils sont enregistrés sur une échelle ordinale. Toutefois, pour les données à distribution multinomiale qui sont enregistrées sur l'échelle nominale, un test F est nécessaire et ce test aura plus d'un degré de liberté dans le numérateur. C'est ainsi par exemple que, s'il y a 5 notes, il y aura 4 degrés de liberté dans le numérateur des tests F utilisés pour comparer deux variétés.

- La méthode COYD pour les données à distribution multinomiale nécessite une plus grande puissance de calcul que la méthode COYD pour les données à distribution normale, ce qui, compte tenu de la puissance de calcul disponible aujourd'hui, ne devrait pas être un sérieux problème.

Différences entre le test χ^2 et la méthode COYD proposée pour les données à distribution multinomiale

- Le test χ^2 tient uniquement compte de la variabilité causée par un échantillonnage aléatoire lorsqu'il est utilisé pour comparer des variétés en fonction de caractères à distribution multinomiale. Cela peut être considéré similaire à la comparaison des moyennes variétales de caractères à distribution normale au moyen d'une analyse où la variabilité entre les plantes dans chaque parcelle est utilisée dans le calcul d'un test t pour comparer deux variétés. Cela serait censé donner trop de résultats significatifs car la variabilité causée par la variation du sol et d'autres conditions de culture sont ignorées. De même, le test χ^2 est censé donner trop de tests t significatifs si la note dépend de la variation du sol et/ou d'autres conditions de culture.
- Le test χ^2 ne dépend pas de l'échelle des mesures de telle sorte que les données enregistrées sur l'échelle nominale et l'échelle ordinale sont traitées de la même façon et parce que le test χ^2 ignore le classement des notes sur l'échelle ordinale. La nouvelle méthode proposée pour les caractères enregistrés sur l'échelle ordinale tient compte de ce classement. Elle est par conséquent censée être plus efficace si les données sont enregistrées sur l'échelle ordinale que si elles le sont sur l'échelle nominale.

Considération

Pour les données à distribution multinomiale, la méthode la plus couramment utilisée est celle du test χ^2 à des fins d'indépendance dans un tableau de contingence. Toutefois, cette méthode tient uniquement compte de la variation causée par un échantillonnage aléatoire. Cela signifie que les décisions relatives aux caractères, qui dépendent des conditions de culture comme la fertilité du sol et le climat, seront trop libérales et qu'elles ne seront par conséquent pas nécessairement cohérentes dans les années à venir.

La méthode COYD pour les données à distribution normale ne peut pas être utilisée pour les données à distribution multinomiale. C'est pourquoi on estime qu'une nouvelle méthode permettant de mieux s'assurer que les décisions prises sont cohérentes dans les années à venir devrait être élaborée. Les travaux ont été effectués sous le titre "Méthodes statistiques applicables aux caractères observés visuellement" avec pour sous-titres : "La méthode globale pour les caractères ordinaux sur plusieurs années", "La méthode globale pour les caractères nominaux sur plusieurs années" et "La méthode globale pour les caractères binomiaux sur plusieurs années", qui pourraient être regroupés sous le titre : "La méthode globale pour les caractères multinomiaux sur plusieurs années".

On estime qu'il est très important d'élaborer, pour les caractères observés visuellement, une méthode qui donne de meilleurs résultats que celle du test χ^2 et ce, afin d'obtenir des conclusions cohérentes dans les années à venir.

[Fin de l'annexe et du document]