

LE FRANÇAIS DES AFFAIRES
RENCH FOR BUSINESS

Comment veiller à l'intégrité des résultats aux tests au moyen de l'analyse des données ?

Dominique Casanova



CCI PARIS ILE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?

PLAN DE LA COMMUNICATION

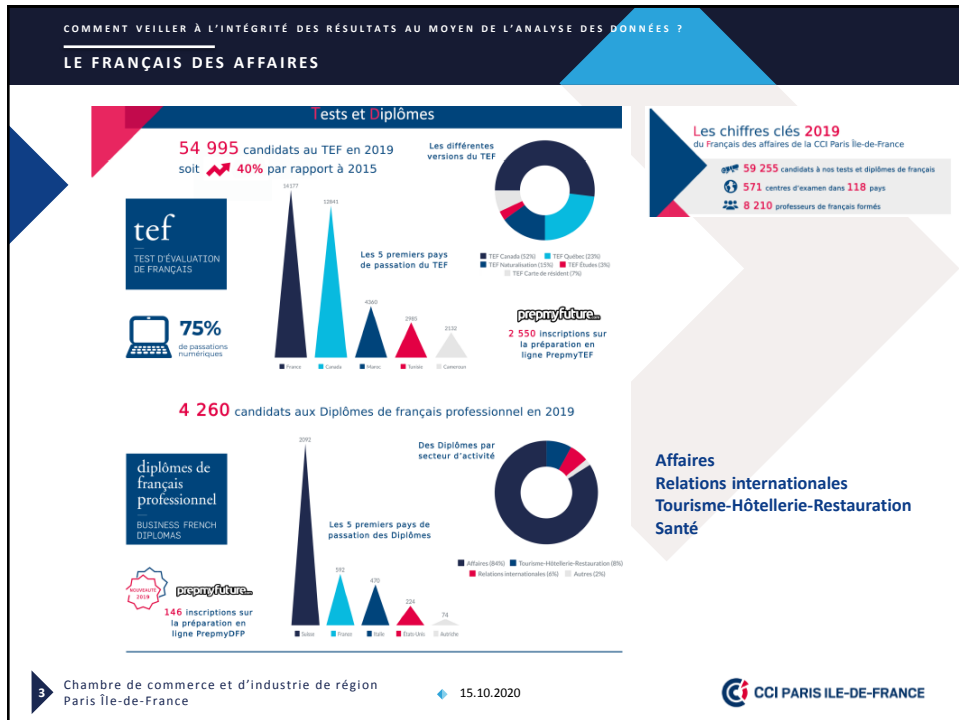
- Le français des affaires
- Les menaces qui pèsent sur l'intégrité des résultats aux tests
- Analyse des patrons de réponse
- Détecter la prise de connaissance préalable d'une partie du matériel de test par des candidats
- Identifier les items concernés
- Perspectives d'amélioration

2

Chambre de commerce et d'industrie de région
Paris Île-de-France

15.10.2020

CCI PARIS ILE-DE-FRANCE



COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?

PRINCIPALES MENACES PESANT SUR L'INTÉGRITÉ DES RÉSULTATS AU TEF

Menaces sur l'intégrité des résultats

- **Anomalies dans le traitement des sessions**
indépendante de la volonté du candidat
→ procédures de contrôle
- **Fraude individuelle ou organisée**
→ procédures de surveillance
→ procédures de contrôle
- **Prise de connaissance par le candidat d'une partie du matériel de test avant la session**, dans le cadre de son effort de préparation au test

Risque plus élevé dans un contexte de sessions à la demande

5

Chambre de commerce et d'industrie de région
Paris Ile-de-France

15.10.2020

CCI PARIS ILE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?

PRINCIPALES MENACES PESANT SUR L'INTÉGRITÉ DES RÉSULTATS AU TEF

Menaces sur l'intégrité des résultats (sessions à la demande)

- **Anomalies dans le traitement des sessions**
indépendante de la volonté du candidat
→ procédures de contrôle
- **Fraude individuelle ou organisée**
→ procédures de surveillance
→ procédures de contrôle
- **Prise de connaissance par le candidat d'une partie du matériel de test avant la session**, dans le cadre de son effort de préparation au test

+ VIGIES
(résultats atypiques)

Risque plus élevé dans un contexte de sessions à la demande

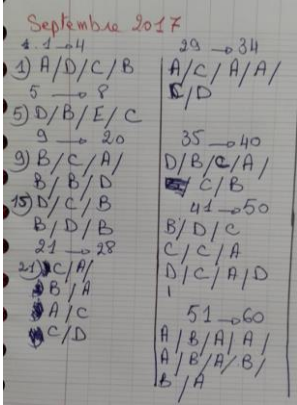
6

Chambre de commerce et d'industrie de région
Paris Ile-de-France

15.10.2020

CCI PARIS ILE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
LE DRAME DES RÉSEAUX SOCIAUX





Brudr Batchmand
Salutations. 🇫🇷
Êtes-vous un candidat du TEF? ..
TEFAQ!
1- Nous fournissons des fuites de questions et réponses sur tout examen à venir.
2-Avez-vous écrit et n'êtes pas sûr ou n'avez pas obtenu les partitions / groupes souhaités? Nous pouvons mettre à niveau vos partitions selon vos souhaits.
3-Êtes-vous un travailleur et vous n'avez pas le temps de passer les examens? nous vous fournissons des certificats 100% authentiques et vérifiés sans passer les examens ... pour plus d'information



Chambre de commerce et d'industrie de région Paris Ile-de-France



15.10.2020



CCI PARIS ILE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
LE DRAME DES RÉSEAUX SOCIAUX

Solutions

- Renouveler la banque d'items
- Multiplier les versions de test (parallèles)
- Arrêter la distribution de tests au format papier
- **Stratégie d'inactivation des items**
- **Détecter les cas suspects**



Chambre de commerce et d'industrie de région Paris Ile-de-France



15.10.2020



CCI PARIS ILE-DE-FRANCE

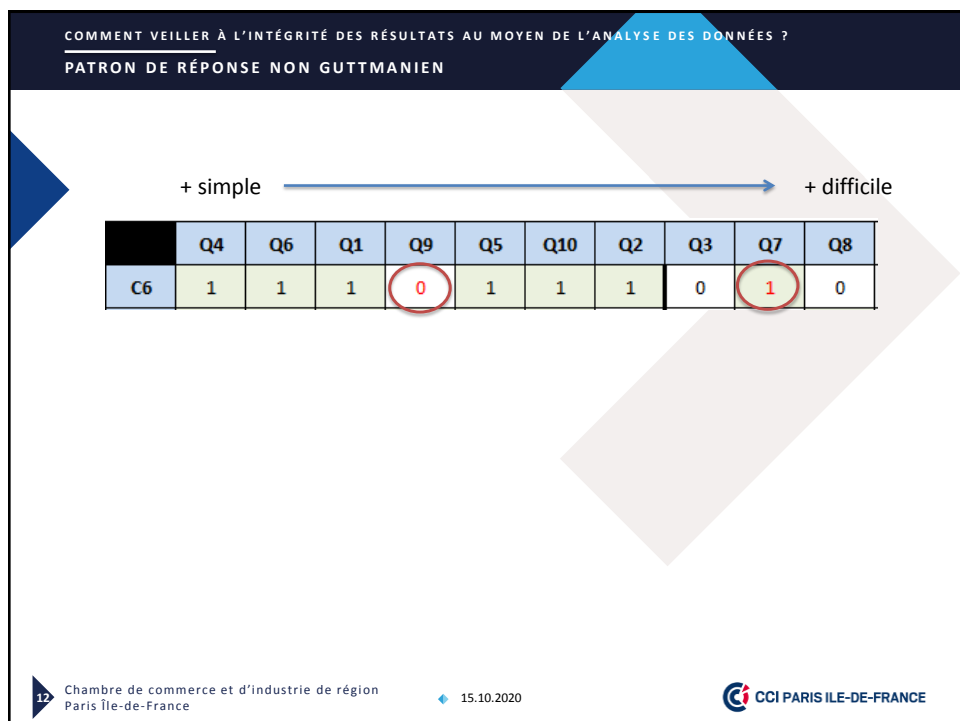
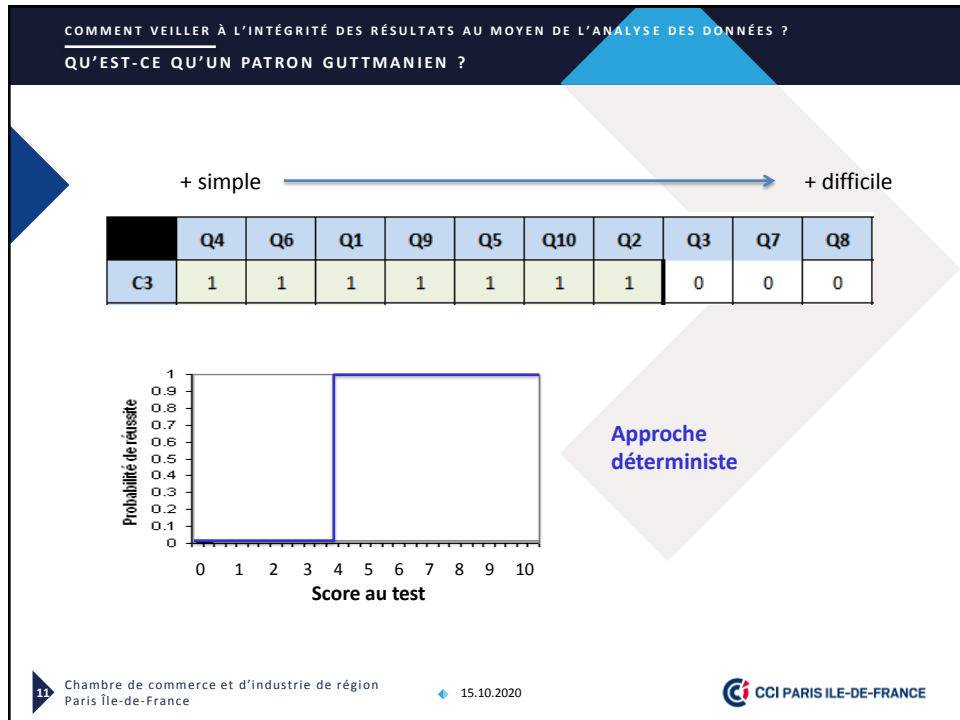
L'analyse des patrons de réponse

9 • Chambre de commerce et d'industrie de région Paris Île-de-France ➤ 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
QU'EST-CE QU'UN PATRON DE RÉPONSE ?

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
C3	1	1	0	1	1	1	0	0	1	1

10 • Chambre de commerce et d'industrie de région Paris Île-de-France ◆ 15.10.2020 CCI PARIS ÎLE-DE-FRANCE



COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
LES RAISONS D'UN ÉCHEC TROUBLANT

+ simple → + difficile

	Q4	Q6	Q1	Q9	Q5	Q10	Q2	Q3	Q7	Q8
C6	1	1	1	0	1	1	1	0	1	0

Oubli de réponse ?
Étourderie dans la sélection de la réponse ?
Déconcentration ?
Parcours d'apprentissage ?

13 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

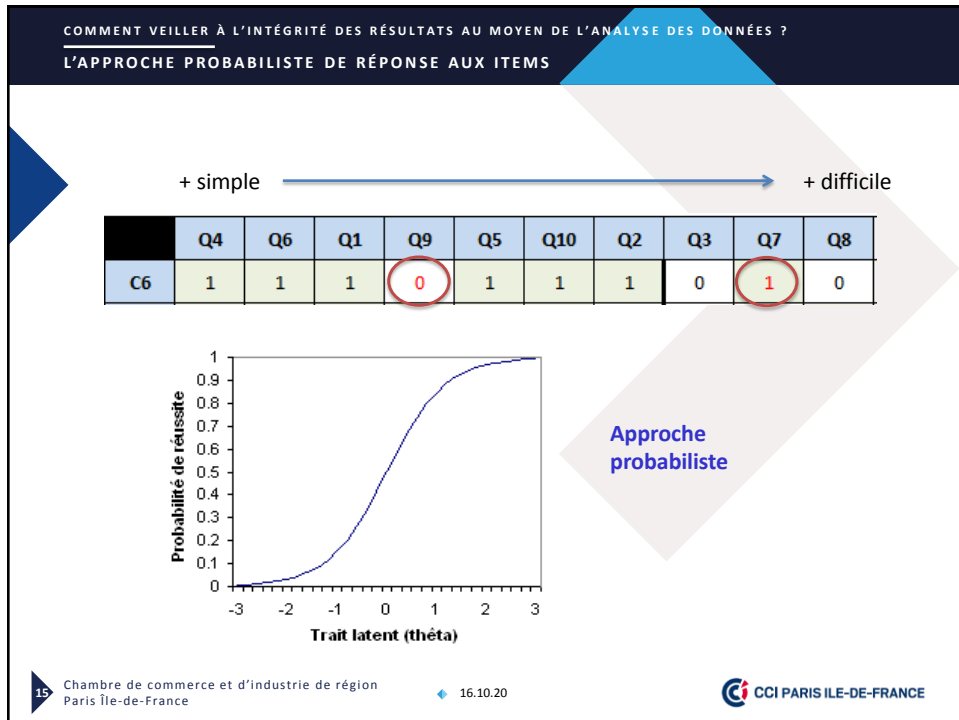
COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
LES RAISONS D'UNE RÉUSSITE TROUBLANTE

+ simple → + difficile

	Q4	Q6	Q1	Q9	Q5	Q10	Q2	Q3	Q7	Q8
C6	1	1	1	0	1	1	1	0	1	0

Réponse au hasard ?
Chance ?
Copie sur voisin ?
Connaissance préalable de l'item ?
Parcours d'apprentissage ?

14 Chambre de commerce et d'industrie de région Paris Île-de-France 16.10.20 CCI PARIS ÎLE-DE-FRANCE



Méthodes de détection de patrons de réponse aberrants

16 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?

PATRONS DE RÉPONSE D'UN ENSEMBLE DE CANDIDATS

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
C1	1	0	0	1	1	1	1	1	1	0
C2	1	0	0	1	1	1	0	0	0	1
C3	1	1	0	1	1	1	0	0	1	1
C4	1	1	1	1	1	1	1	0	1	1
C5	1	1	1	1	0	1	0	0	0	0
C6	1	0	1	1	1	1	1	0	0	1
C7	1	0	0	1	1	1	1	0	1	0
C8	1	1	0	1	0	1	0	0	1	0
C9	1	0	1	1	0	1	0	0	1	0
C10	0	0	0	1	0	1	0	0	0	0

17 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?

ORDONNANCEMENT DES PATRONS DE RÉPONSE

+ simple → + difficile

	Q4	Q6	Q1	Q5	Q9	Q2	Q3	Q7	Q10	Q8	Total
C4	1	1	1	1	1	1	1	1	1	0	9
C1	1	1	1	1	1	0	0	1	0	1	7
C3	1	1	1	1	1	1	0	0	1	0	7
C6	1	1	1	1	0	0	1	1	1	0	7
C7	1	1	1	1	1	0	0	1	0	0	6
C2	1	1	1	1	0	0	0	0	1	0	5
C5	1	1	1	0	0	1	1	0	0	0	5
C8	1	1	1	0	1	1	0	0	0	0	5
C9	1	1	1	0	1	0	1	0	0	0	5
C10	1	1	0	0	0	0	0	0	0	0	2
Total	10	10	9	6	6	4	4	4	4	1	

18 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
APPROCHE « GUTTMANNIENNE »

Approche « Guttmanienne »

	Q4	Q6	Q1	Q9	Q5	Q10	Q2	Q3	Q7	Q8	Total
C4	1	1	1	1	1	1	1	1	1	0	9
C3	1	1	1	1	1	1	1	0	0	0	7
C6	1	1	1	0	1	1	1	0	1	0	7
C1	1	1	1	1	1	1	0	0	0	1	7
C7	1	1	1	1	1	0	0	0	1	0	6
C8	1	1	1	1	0	0	1	0	0	0	5
C9	1	1	1	1	0	0	0	1	0	0	5
C2	1	1	1	0	1	1	0	0	0	0	5
C5	1	1	1	0	0	0	1	1	0	0	5
C10	1	1	0	0	0	0	0	0	0	0	2
Total	10	10	9	6	6	5	5	3	3	1	

19 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
CAUTION INDEX

x_n Patron de réponse du candidat
 p Patron de réponse total du groupe
 x_n^* Patron de réponse guttmannien pour un tel score

$$C.Sato = 1 - \frac{cov(x_n, p)}{cov(x_n^*, p)}$$

0 si le patron de réponses est guttmannien
D'autant plus grand que le patron de réponses s'écarte d'un patron guttmannien

Sato, T. (1975) *The construction and interpretation of S-P tables*. Tokyo: Meiji Tosho.

20 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?

MÉTHODE C*

	Q4	Q6	Q1	Q9	Q5	Q10	Q2	Q3	Q7	Q8	Total	C*
C4	1	1	1	1	1	1	1	1	1	0	9	0,000
C3	1	1	1	1	1	1	1	0	0	0	7	0,000
C6	1	1	1	0	1	1	1	0	1	0	7	0,144
C1	1	1	1	1	1	1	0	0	0	1	7	0,192
C7	1	1	1	1	1	0	0	0	1	0	6	0,089
C8	1	1	1	1	0	0	1	0	0	0	5	0,042
C9	1	1	1	1	0	0	0	1	0	0	5	0,125
C2	1	1	1	0	1	1	0	0	0	0	5	0,042
C5	1	1	1	0	0	0	1	1	0	0	5	0,167
C10	1	1	0	0	0	0	0	0	0	0	2	0,000
Total	10	10	9	6	6	5	5	3	3	1		

21 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?

APPROCHE PARAMÉTRIQUE (LZ)

Approche « probabiliste »

Probabilité de réponse à un item pour une compétence donnée :

$$P_i(\theta) = P(x_i = 1 | \theta, b_i) = \frac{\exp[\theta - b_i]}{1 + \exp[\theta - b_i]}$$

22 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
APPROCHE PARAMÉTRIQUE (LZ)

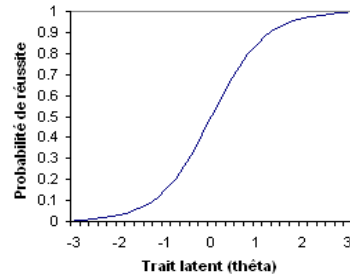
Approche « probabiliste »

Probabilité de réponse à un item
pour une compétence donnée :

$$P_i(\theta) = P(x_i = 1 | \theta, b_i) = \frac{\exp[\theta - b_i]}{1 + \exp[\theta - b_i]}$$

Vraisemblance d'un patron de réponse
pour une compétence donnée :

$$L(\theta) = \prod_{i=1}^n p_i(\theta)^{x_i} (1 - p_i(\theta))^{1-x_i}$$



COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
APPROCHE PARAMÉTRIQUE (LZ)

Approche « probabiliste »

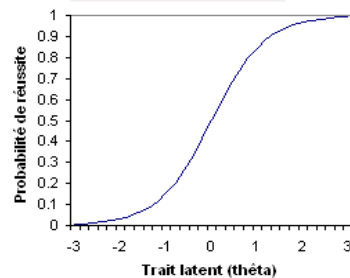
Probabilité de réponse à un item
pour une compétence donnée :

$$P_i(\theta) = P(x_i = 1 | \theta, b_i) = \frac{\exp[\theta - b_i]}{1 + \exp[\theta - b_i]}$$

Vraisemblance d'un patron de réponse
pour une compétence donnée :

$$L(\theta) = \prod_{i=1}^n p_i(\theta)^{x_i} (1 - p_i(\theta))^{1-x_i}$$

$$l_0(\theta) = \log L(\theta) = \sum_{i=1}^I \{x_i \log P_i(\theta) + (1 - x_i) \log Q_i(\theta)\}$$



$$l_z = \frac{l_0 - E(l_0)}{V(l_0)^{1/2}}$$

Drasgow, F., Levine, M. V., and Williams, E. A. (1985) Appropriateness measurement with polychotomous item response models and standardized indices. *British Journal of Mathematical and Statistical Psychology*, **38**(1), 67–86.

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
DIFFICULTÉS D'APPLICATION

Difficultés d'application dans un contexte de sessions à la demande

1. Estimation de l'indice liée au patron de réponse du groupe
 - a. Il faut disposer du patron de réponse d'un groupe
Quid des premières sessions à la demande ?
 - b. Si on accumule de nouvelles données, les valeurs des coefficients vont changer
2. Comment fixer la valeur seuil pour la prise de décision ?
 - a. En fonction du groupe ? (percentile)
 - b. En préfixant une valeur seuil ? (expérience)

25 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
EXPLOITATION DE LA BANQUE D'ITEMS

Contexte du TEF

- **Banque d'items calibrés** → on connaît la difficulté a priori des items
- **Versions équivalentes** → on connaît la distribution habituelle des compétences

donc on peut, à partir des fonctions de probabilité de réussite des items :

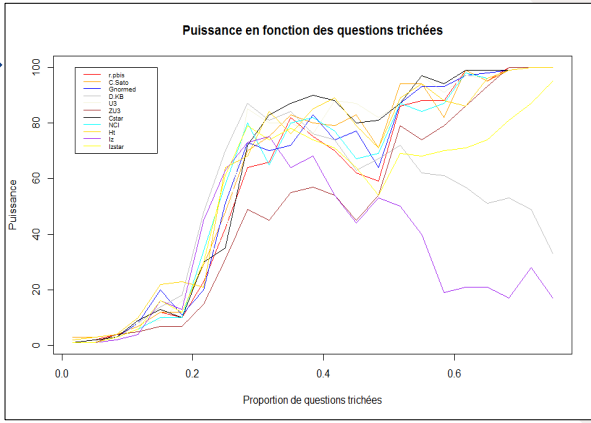
- **Générer des patrons de réponse plausibles pour une population similaire**
- **Générer les indices statistiques de vraisemblance**
- **Fixer des seuils pour ces indices en fonction du % de faux positifs qu'on accepte de traiter**

Pour tout nouveau résultat

- On peut calculer les indices relativement au groupe simulé
- Et donc déterminer un taux de suspicion par rapport à la distribution statistique des indices pour le groupe de référence

26 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
MÉTHODES DE DÉTECTION DES PATRONS DE RÉPONSE ABERRANTS



Jorge N. Tendeiro, Rob R. Meijer, A. Susan M. Niessen (2016).
PerFit: An R Package for Person-Fit Analysis in IRT.
Journal of Statistical Software, 74(5), 1-27

Approche non paramétrique : **Ht**
Approche paramétrique : **Iz***

Aw et al. (2019). **Mise à l'épreuve de méthodes de détection de patrons de réponses suspects**. Actes du 31^e colloque de l'Admee-Europe. Luxembourg, 8-10 janvier 2019.

27 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

Comment détecter
des « délits d'initié » ?

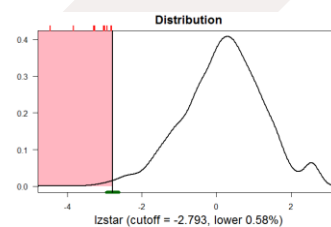
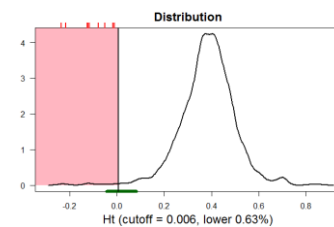
Comment repérer les items
ayant perdu leur confidentialité ?

28 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
APPLICATION EN SITUATION RÉELLE

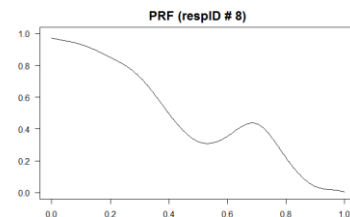
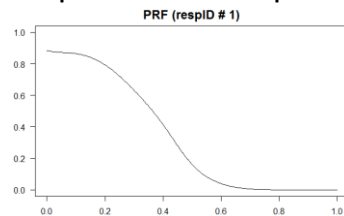
Simulation de données et définition de seuils a priori

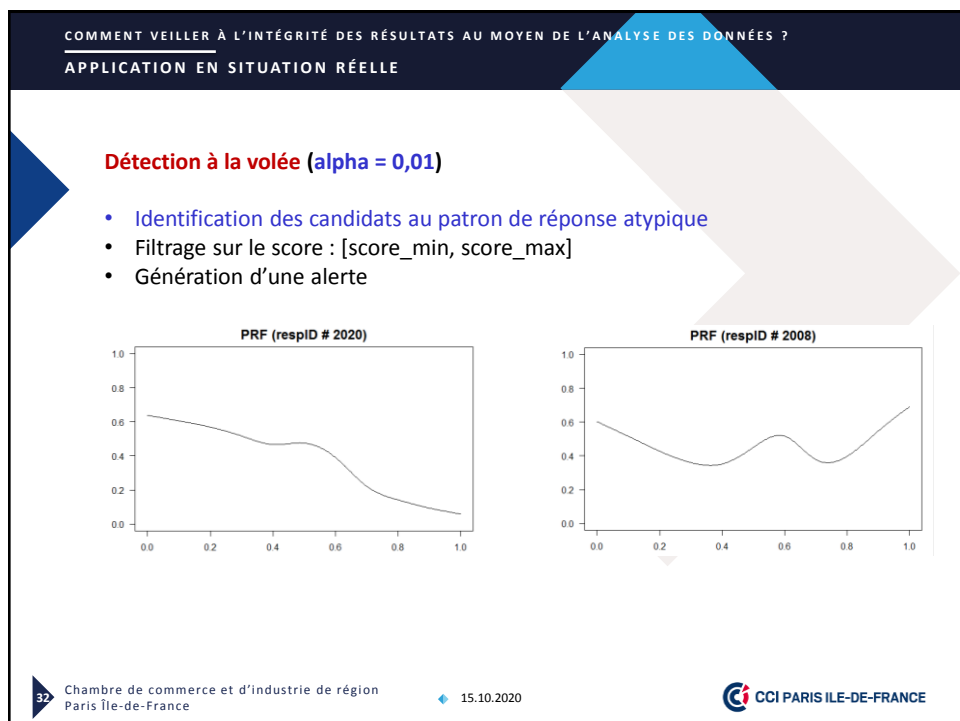
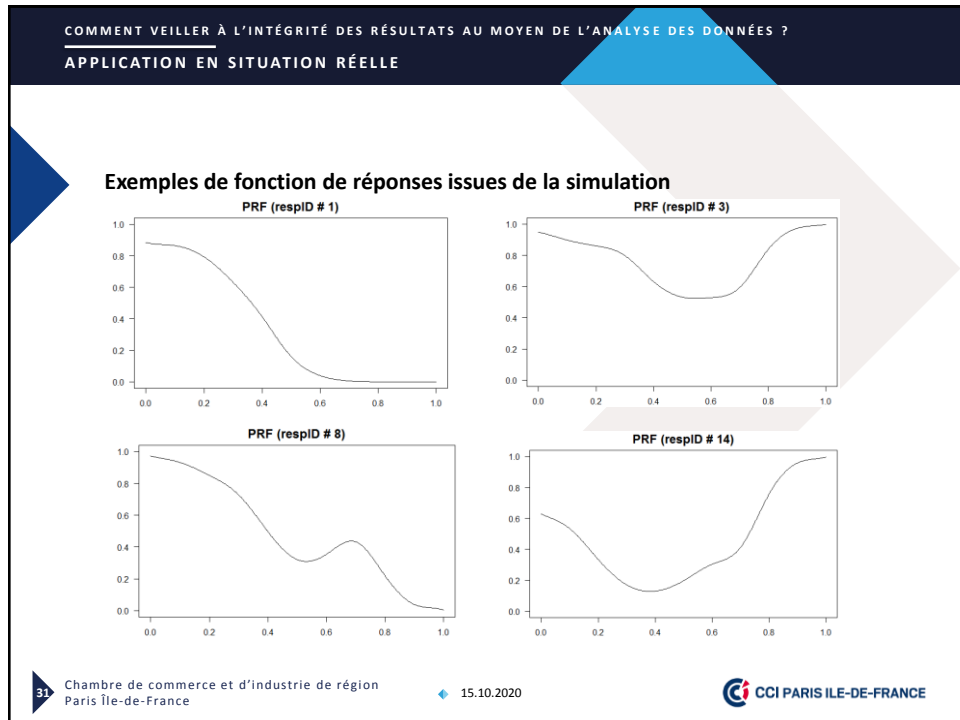
- À partir des paramètres d'items et d'une distribution type de la population on génère des patrons de réponse en s'appuyant sur la probabilité de réussite
 - Si $P_i(\theta) > \text{alea}()$ réponse = 1, sinon réponse = 0
- À partir de cet ensemble (probable) de données on détermine les seuils de détection (1% et 5%)

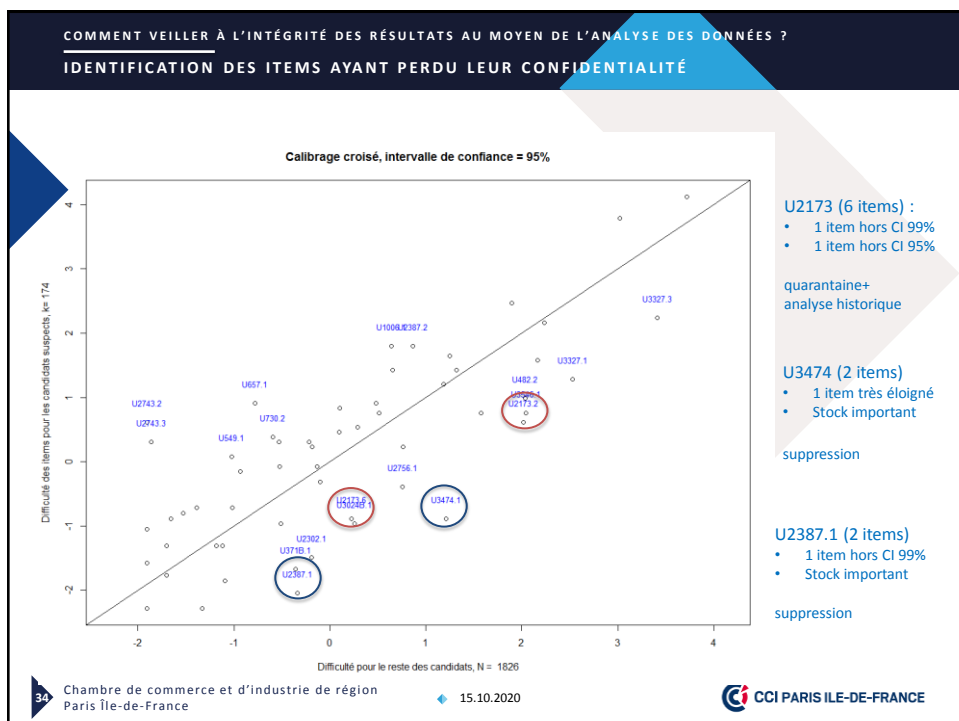
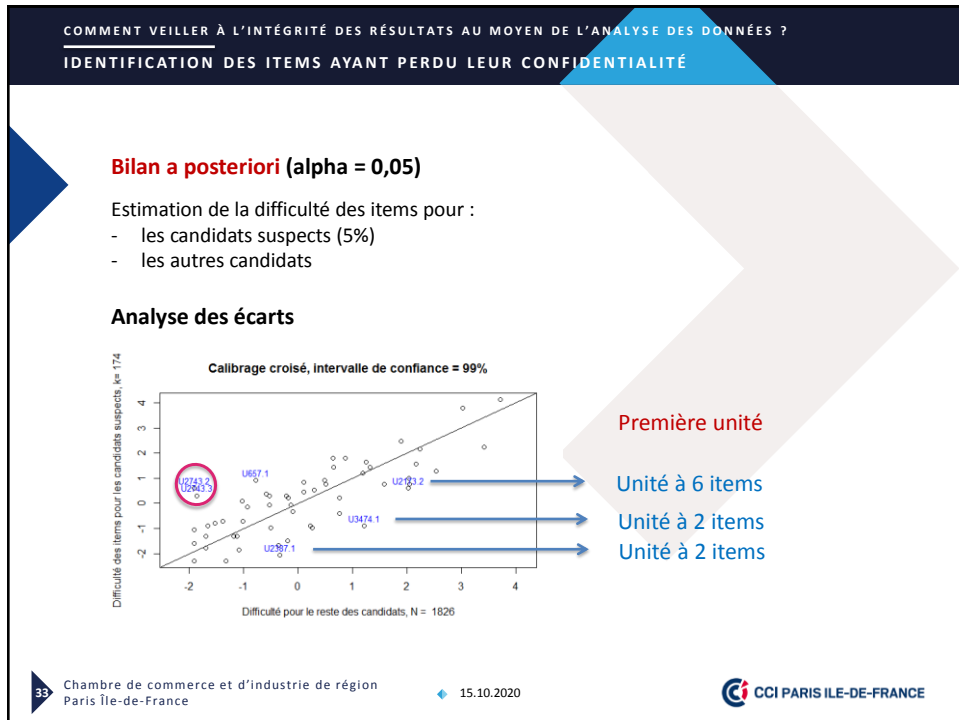


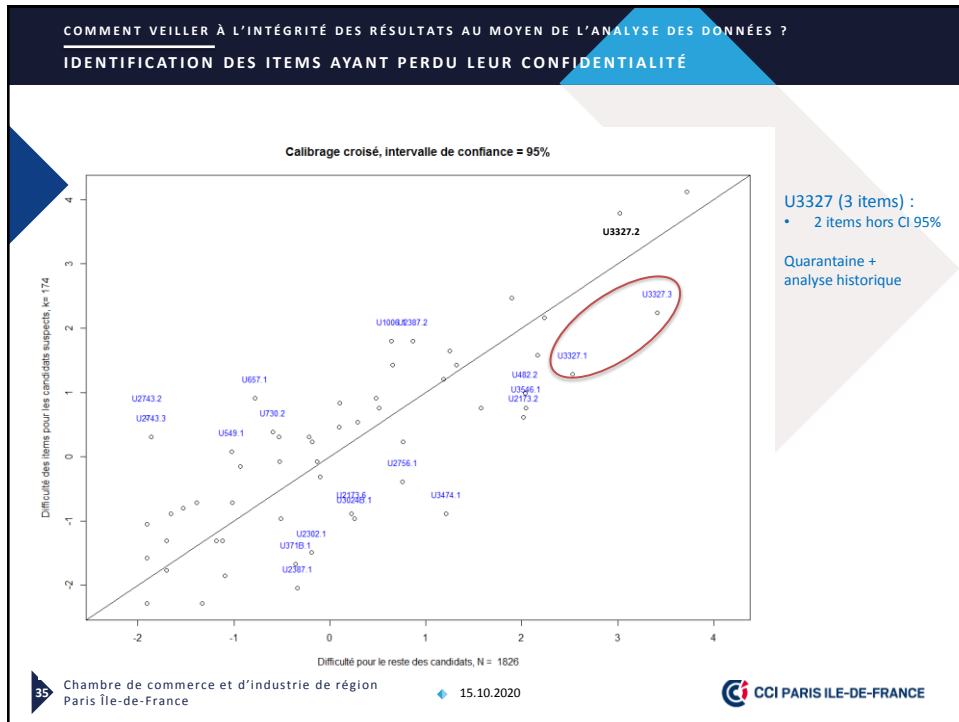
COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?
APPLICATION EN SITUATION RÉELLE

Exemples de fonction de réponses issues de la simulation









COMMENT VEILLER À L'INTÉGRITÉ DES RÉSULTATS AU MOYEN DE L'ANALYSE DES DONNÉES ?

PERSPECTIVES

Épreuves au format QCM : prise en compte du temps de réponse par unité

- Comparaison par rapport une distribution du temps de réponse pour l'unité
- En tenant compte du succès ou de l'échec aux items de l'unité
- Et du score total du candidat

Épreuves d'expression écrite : prise en compte de la dynamique de saisie

- Fréquence et longueur des pauses (rassemblement des idées)
- Rapidité de saisie du texte (restitution d'un texte mémorisé ?)
- Fréquence des transformation du texte
 - o correction
 - o Réorganisation

37 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE

dcasanova@cci-paris-idf.fr

LE FRANÇAIS DES AFFAIRES
RENCH FOR BUSINESS

38 Chambre de commerce et d'industrie de région Paris Île-de-France 15.10.2020 CCI PARIS ÎLE-DE-FRANCE