

A.R.I.R.S.

Actions de test par envoi de CV

Rapport général remis à la HALDE

Mai 2008

Table des matières

<i>A.R.I.R.S.</i>	1
1 Présentation générale des tests	3
1.1 Note Méthodologique	3
1.1.1 Méthode	3
1.1.2 Choix des offres d'emploi pour lesquelles des candidatures sont émises.....	3
1.1.3 Construction des candidatures en réponse à ces offres d'emploi.....	4
1.1.4 Envoi des candidatures par voie électronique	5
1.1.5 Relevé des réponses	5
1.2 L'Analyse statistique des données	6
1.2.1 Introduction.....	6
1.2.2 Le test exact de Fisher.....	7
1.2.3 Le test d'indépendance du χ^2 de Pearson	9
1.2.4 Le test de Student.....	12
1.2.5 Mode de calcul opérationnel	13
1.2.6 Remarques sur l'exigence de significativité	13
1.2.7 Nos résultats livrent une information dans le cadre retenu	14
1.3 Procédure d'examen contradictoire des résultats	14
1.3.1 Qualité et équivalence des CV	15
1.3.2 Envois	15
1.3.3 Réponses	15
1.3.4 Traitements statistiques.....	15
1.3.5 Principe du test.....	16
2 Résultats généraux	17
2.1 Contexte du test	17
2.1.1 Périmètre de l'étude.....	17
2.1.2 Délais de réponse	17
2.1.3 Répartition des réponses positives	17
2.2 Synthèse des résultats	18
2.2.1 Discrimination liée à l'Origine	18
2.2.2 Discrimination liée à l'âge.....	19

1 Présentation générale des tests

1.1 Note Méthodologique

1.1.1 Méthode

La méthode de « **l'audit par couple** » ou « **test de discrimination** » qui a été utilisée consiste à envoyer, en réponse à une offre d'emploi, deux CV accompagnés parfois d'une lettre de candidature et d'un formulaire de recrutement qui ne diffèrent que par une caractéristique : la variable à tester. La méthode, classique dans les pays anglo-saxons et légale en France est recommandée par le BIT. Elle peut aussi consister en l'envoi de plus de 2 CV en parallèle pour une même offre.

Nos tests ne portent que sur le tri de CV opéré par les entreprises, pourquoi ?

C'est la mesure la plus fiable. En effet, poursuivre le test au-delà du tri de CV nécessiterait que des candidats factices se présentent aux entretiens pour lesquels nous avons reçu une convocation. Cela poserait de redoutables problèmes de fiabilité. Il faudrait en effet une grande technicité dans la sélection et la préparation des candidats factices qui jouent un rôle, ainsi que dans la traçabilité des entretiens (enregistrements son et vidéo...).

Pour chaque entreprise, un échantillon de cv est envoyé. Nous comparons le taux de succès de candidats dits de « référence » (non susceptible d'être discriminé) à des candidats ayant une caractéristique telle qu'ils sont, par hypothèse, susceptibles d'être discriminés. Les deux variables de discriminations utilisées dans les tests sont les suivantes :

- l'origine (du Maghreb et d'Afrique subsaharienne) qui est signalée par le patronyme ;
- l'âge (autour de 45 ans).

Pour chacune des entreprises testées, le processus de travail se fait en 4 étapes :

- choix des offres d'emploi¹ pour lesquelles des candidatures seront émises ;
- construction des candidatures en réponse à ces offres d'emploi ;
- envoi des candidatures ;
- relevé des réponses.

Pour mener à bien notre mission, nous avons développé un outil informatique qui assure la traçabilité de nos activités. Les principales fonctionnalités de cet outil sont décrites dans cette note.

1.1.2 Choix des offres d'emploi pour lesquelles des candidatures sont émises

Le choix des offres d'emploi est fait par notre équipe qui, par sa connaissance du marché de l'emploi, cherche à sélectionner des offres d'emploi pour lesquelles :

- Le taux de retour sera satisfaisant**, cela permettra d'augmenter le nombre d'observations valides (nombre de réponses positives reçues pour une entreprise), ce qui est essentiel pour la fiabilité statistique des écarts constatés.

¹ Nous appelons « offre d'emploi » toutes les offres, quel que soit le type de contrat proposé : CDD, CDI ou stage.

- Le risque que l'entreprise découvre qu'elle fait l'objet d'un test est le plus faible**, ce qui est un point essentiel de notre mission. Pour cela, il ne faut pas que le profil cherché soit trop rare. En effet, il nous faut construire des CV de qualité mais dont les caractéristiques soient suffisamment variées pour que nos CV ne se ressemblent pas trop. Il faut donc que l'offre d'emploi accepte un minimum de variété dans les profils recherchés.
- La date de publication de l'offre est récente**. Il nous faut en effet réagir suffisamment vite pour maximiser le taux de retour. Au bout d'un certain temps, une entreprise a reçu suffisamment de CV, elle ne consulte plus les nouveaux CV.

Une fois que l'offre a été sélectionnée, nous la saisissons dans l'outil informatique, en renseignant les informations suivantes :

- intitulé de l'offre (poste à pourvoir) ;
- référence de l'offre ;
- date de publication ;
- texte de l'offre ;
- méthode d'envoi requise par l'entreprise (saisie d'un formulaire, e-mail)

Cas particuliers

Pour respecter la période de test, nous avons dans certains cas pu sélectionner plusieurs fois une même offre, généralement à plusieurs semaines d'intervalle. Ces offres sont considérées par la suite comme une nouvelle offre à part entière.

Pour 2 entreprises, du fait du très faible nombre d'offres d'emploi disponibles pendant la période du test, nous avons également envoyé des candidatures spontanées en plus des candidatures en réponses à une offre. Nous l'avons précisé lorsque cela a été le cas.

1.1.3 Construction des candidatures en réponse à ces offres d'emploi

Pour chacune des offres d'emploi sélectionnées, nous construisons une ou plusieurs séries de candidatures. Chaque série comporte une candidature dite de référence et une candidature dite testée. Une série peut comporter 3 candidats aux caractéristiques différentes (Réfèrent, Origine, Age). Les tests statistiques mis en œuvre tiennent compte de cette particularité.

Une candidature se compose de :

- Une identité** : nom, prénom, adresse e-mail spécialement construite, adresse postale, un numéro de téléphone portable.

Un CV, reprenant ces éléments d'identité ainsi qu'un parcours professionnel, de formation, de compétences... Ces CV sont construits par nos équipes de manière à assurer l'équivalence entre les 2 ou 3 CV d'une même série, mais également de manière à maximiser le taux de retour afin d'avoir un nombre suffisant d'observations valides. Toutefois, nous prenons garde à la crédibilité de notre opération : une entreprise se poserait vraisemblablement des questions si elle recevait 30 CV de diplômés de HEC dans la même semaine ! Pour assurer la non visibilité du test nous faisons donc varier les diplômes et compétences des candidats selon les séries. Lorsque nous construisons un CV, nous ne savons pas à quelle variable il sera attribuée (candidat de référence, âgé, porteur de patronyme originaire du Maghreb ou d'Afrique Sub-saharienne...)

- Si demandée, une lettre de motivation. Cette lettre est également parfois jointe à notre initiative pour rendre moins visible notre action de test.

1.1.4 Envoi des candidatures par voie électronique

La candidature est adressée à l'entreprise selon le mode retenu par l'offre d'emploi : par e-mail ou par saisie d'un formulaire de candidature.

Cas d'une candidature à adresser par e-mail

CV et lettre de motivation sont adressés à l'entreprise par courrier électronique. Généralement, les entreprises répondent également par e-mail à ce type de candidature. Ces e-mails reçus sur cette boîte e-mail sont relevés en fin de programme et serviront en cas de contestation.

Cas d'une candidature à saisir sur un formulaire de candidature

CV et lettre de motivation sont saisis sur un formulaire de candidature.

2 méthodes ont été utilisées pour assurer la traçabilité de nos saisies, selon le système utilisé par l'entreprise :

- Certains outils de recrutement permettent au candidat de se créer un compte, grâce à un login et un mot de passe. Dans ce cas, les candidatures déposées peuvent être consultées par la suite. Nous saisissons dans notre système informatique les login et mot de passe utilisés, ce qui permettra par la suite de contrôler les saisies réalisées, s'il est nécessaire. La saisie sera par ailleurs notifiée dans notre système informatique.
- Parfois cette fonctionnalité n'existe pas et le compte candidat ne permet pas toujours de revenir sur tous les champs du formulaire. Nous utilisons donc un logiciel permettant de filmer ce qui se passe sur l'écran d'un ordinateur. Cette méthode permet de garder systématiquement la trace de tous nos envois de CV par voie électronique.

1.1.5 Relevé des réponses

Pour les réponses reçues par courrier électronique, que ce soit les accusés de réception, les réponses positives ou négatives, un transfert automatique des e-mails reçus par les candidats fictifs nous a permis de conserver une trace informatique systématique de tous les e-mails reçus.

Pour les réponses reçues sur la boîte vocale d'un téléphone portable, nous avons écouté les messages téléphoniques depuis un ordinateur, dont les conversations téléphoniques ont été enregistrées.

1.2 L'Analyse statistique des données

1.2.1 Introduction

Il y a discrimination dès lors que de manière répétée, après envoi de candidatures globalement identiques sur un même poste (qui ne diffèrent que par l'âge ou le patronyme), une candidature est préférée à une autre. Pour mesurer l'existence d'une discrimination, lors d'une opération de test, il nous faut donc comparer le taux de succès des candidats en fonction de leur âge et de leur patronyme.

Lorsqu'un type de candidat obtient plus de réponses qu'un autre qui avait pourtant postulé aux mêmes emplois, l'écart entre les deux mesure une « **discrimination selon une variable** ». Une discrimination peut opérer, par exemple, à l'avantage du candidat potentiellement discriminé ou à celui du candidat référent. L'addition de ces deux discriminations de sens inverse donne une mesure de la « **discrimination brute** ». Mais on s'accorde évidemment à penser que le plus important est de connaître la « **discrimination nette** », c'est-à-dire le nombre de fois où le candidat potentiellement discriminé est défavorisé (en déduisant les fois où il est préféré).

En d'autres termes, lorsque plusieurs séries de CV sont envoyées, la mesure d'une discrimination consiste donc à comparer le nombre de réponses positives obtenues par les candidats en lice. Puisque les candidats avaient postulé aux mêmes offres, il suffit de constater que l'un des deux a obtenu au total plus de réponses positives que l'autre pour mesurer la discrimination dite nette, c'est ce que nous faisons. Présenter de la sorte les résultats des tests de discrimination est plus clair et compréhensible. Cette présentation est plus proche du « vécu quotidien » des demandeurs d'emploi, elle en diffère à l'évidence puisque les demandeurs d'emplois ignorent combien les candidats qui leur étaient opposés obtiennent de réponses positives. Notre mesure des discriminations donne précisément cette information.

Mais l'écart de réussite entre les types de candidats (la discrimination nette) qui est une information intéressante est-il significatif ou fortuit ?

Pour répondre à cette question, nous utiliserons 3 méthodes mathématiquement fiables et reconnues par la communauté scientifique :

- Le Test exact de Fisher
- Le test d'indépendance du χ^2 de Pearson
- Le T-Test de Student apparié

Nous allons présenter les 3 méthodes.

Dans les calculs, nous utiliserons les 3 méthodes.

1.2.2 Le test exact de Fisher

1.2.2.1 Principe général

Considérons deux variables X et Y ayant chacune 2 modalités (X1, X2 et Y1, Y2).

Ces variables se présentent sous un tableau de contingence 2x2 dont les effectifs seront notés (a,b,c,d), comme suit :

	<u>Y1</u>	<u>Y2</u>	
<u>X1</u>	<u>a</u>	<u>b</u>	<u>a+b</u>
<u>X2</u>	<u>c</u>	<u>d</u>	<u>c+d</u>
	<u>a+c</u>	<u>b+d</u>	<u>n</u>

où $n = a+b+c+d$

Si les deux variables X et Y sont indépendantes, la probabilité d'observer le tableau (a,b,c,d), est donnée par la loi hypergéométrique suivante :

$$P(a,b,c,d) := \frac{(a+c)! (b+d)! (a+b)! (c+d)!}{n! a! b! c! d!}$$

Plus le tableau (a,b,c,d) s'éloigne de l'indépendance, c'est-à-dire plus il est déséquilibré, et plus la probabilité P(a,b,c,d) est petite.

Calculer la valeur du test exact de Fisher consiste à calculer la somme suivante : probabilité de la table observée (a,b,c,d) + probabilité de chaque table (a', b', c', d') présentant un plus grand écart à l'indépendance (l'équilibre) que la table observée (a,b,c,d), tout en maintenant les totaux de chaque ligne et chaque colonne.

Voir par exemple TUFFERY (2006)² et Laurencelle (2007)³

² « Étude statistique et préparation des données ». Document non publié disponible sur le site de l'auteur. Stéphane TUFFERY est l'auteur de « *Data Mining et statistique décisionnelle* », Éditions Technip, nouvelle édition revue et enrichie, juin 2007.

³ Laurencelle, L. (2007). Les tableaux de fréquences 2 × 2 et leur traitement statistique. *Tutorials in Quantitative Methods for Psychology Vol. 3(1)*, 14-25.

1.2.2.2 Mise en œuvre

Poser les bonnes hypothèses

Pour notre propos nous voulons vérifier l'hypothèse H0 selon laquelle les variables sont indépendantes.

Si les 2 variables ne sont pas indépendantes, cela veut dire qu'une des deux variables influence l'autre. Dans ce cas de figure cette influence se fait nécessairement dans une direction ou dans une autre (l'augmentation de l'une fait diminuer l'autre, ou, l'augmentation de l'une fait augmenter l'autre).

Lorsque la direction dans laquelle une variable peut influencer l'autre est a priori connue du chercheur, alors il est préférable de ne pas nier cette information et de la prendre en compte dans le test statistique. En faisant ainsi le chercheur réalise un test dit « directionnel » ou unilatéral.

Pour une compréhension intuitive, on pourrait dire que si les 2 variables ne sont pas indépendantes, alors les répartitions des effectifs dans le tableau ne vont pas être équilibrées - schématiquement, ce déséquilibre fait que le tableau va pencher d'un côté ou de l'autre. Si nous savons de quel côté le tableau risque de pencher alors, nous devons surveiller précisément ce côté (unilatéral), plutôt que de surveiller des deux côtés (bilatéral).

Introduire cette information dans le test revient à dire « quelle est l'hypothèse que nous formulons si H0 n'est pas vérifiée », autrement dit d'émettre une hypothèse alternative H1.

Pour la mise en œuvre de notre test, la première étape, consistant à formuler des hypothèses, se présente donc ainsi :

Nous posons l'hypothèse H0 selon laquelle les variables sont indépendantes contre l'hypothèse alternative H1, selon laquelle le groupe de candidats potentiellement discriminés (en raison du patronyme ou de l'âge) a reçu moins de réponses positives que le groupe de référence.

Au vu des résultats généraux exposés à la section 2 : Résultats généraux, il nous semble légitime de poser cette hypothèse. En effet, nous constatons bien une discrimination globale, fréquente et répétée concernant l'origine et l'âge.

Calculer la valeur du Test exact de Fisher.

Prenons l'exemple du tableau suivant :

Variable Y = Type de candidat ; 2 modalités : Référent / Candidat potentiellement discriminé

Variable Z = Résultat de la candidature ; 2 modalités : Pas de réponse positive / Réponse positive

Z	Y	Référent	Potentiellement discriminé	Somme
Pas de réponse positive		10	18	28
Réponse positive		10	2	12
	Somme	20	20	40

Notre hypothèse H0 est que les deux variables (Type de candidat et Résultat de la candidature) sont indépendantes, contre l'hypothèse H1 selon laquelle le groupe de candidats potentiellement discriminés a reçu moins de réponses positives que le groupe de référence.

La probabilité correspondant à notre tableau observé est :

$$P(10,10,18,2) = (20! \times 20! \times 28! \times 12!) / (40! \times 10! \times 10! \times 18! \times 2!) = 0,00628326$$

Les situations dans lesquelles les candidats ayant un prénom d'origine maghrébine ont reçus encore moins de réponses positives, tout en conservant identiques les totaux des lignes et des colonnes de notre tableau, sont les suivantes :

$$P(10,11,19,1) = 0,0008931$$

$$P(10,12,20,0) = 0,000058477$$

La probabilité exacte du test de Fisher est donc de :

$$0,00628326 + 0,0008931 + 0,000058477 = 0,00723484$$

Dans notre exemple, il y a donc 0,7% de probabilité que les variables soient indépendantes ou encore 99,3% de chances qu'il faille rejeter H0 au profit de H1. Il est donc très fortement probable (plus de 99 chances sur 100) que les candidats potentiellement discriminés aient effectivement subi une discrimination..

1.2.2.3 Décider d'accepter H0 ou de la rejeter au profit de H1

Dans la pratique, nous ne rejeterons l'hypothèse H0 au profit de H1 que si le test statistique nous indique que, sous une hypothèse d'indépendance, l'observation d'une telle répartition ou d'une répartition plus déséquilibrée encore n'est que « peu probable » ($p < 10\%$) ou « extrêmement peu probable » ($p < 5\%$). Si cette probabilité est inférieure à 1%, nous la considérerons comme « quasiment nulle », puisqu'inférieure à l'unité et donc proche de zéro.

1.2.3 Le test d'indépendance du χ^2 de Pearson

La méthode

Afin d'étudier la discrimination, on cherche à savoir si le type de candidature (Candidat référent / Candidat potentiellement discriminé) a une influence sur son recrutement. Le principe est de tester l'indépendance entre ces deux variables. Les données sont en effet décrites par deux variables :

Considérons les deux mêmes variables que précédemment :

Variable Y = Type de candidat ; 2 modalités : Référent / Candidat potentiellement discriminé

Variable Z = Résultat de la candidature ; 2 modalités : Pas de réponse positive / Réponse positive

On appelle $X_{Y;Z}$ le nombre de candidatures correspondant au couple (Y;Z). Par exemple $X_{0;0}$ est égal au nombre de candidatures rejetées des individus potentiellement discriminés.

C'est sur ces effectifs X que se fonde le test statistique du χ^2 de Pearson.

On peut résumer l'information dans le tableau de contingence suivant :

	$Y = 0$	$Y = 1$
$Z = 0$	$X_{0,0}$	$X_{1,0}$
$Z = 1$	$X_{0,1}$	$X_{1,1}$

Par la suite, nous prendrons en exemple les données suivantes, adaptées à notre contexte :

Z	Y	Candidat potentiellement discriminé	Candidat de référence
	Refus	110	88
	Acceptation	11	33

Calcul de la significativité des écarts

Comment savoir si les écarts observés entre les candidats sont le fruit du hasard ou d'une discrimination ? En effet, des candidats de « références » peuvent aussi recevoir des réponses différentes bien que leurs CV puissent être similaires. S'il n'y a pas discrimination, c'est parce qu'il n'y a pas de lien entre le fait d'avoir telle origine et de recevoir un nombre donné de réponses. Les variables sont indépendantes.

On fait l'hypothèse H_0 suivante : les variables Y et Z sont indépendantes. Le test statistique mis en œuvre a pour but de valider ou de rejeter cette hypothèse à un niveau de confiance donné, par exemple 5%.

Si cette hypothèse est vraie, alors la probabilité d'être maghrébin et d'avoir sa candidature rejetée est égale à la probabilité d'être maghrébin multipliée par la probabilité d'avoir sa candidature rejetée. En effet pour deux événements A et B indépendants :

$$\mathbb{P}(A \text{ et } B) = \mathbb{P}(A) * \mathbb{P}(B)$$

On peut donc fournir des effectifs dits théoriques. Calculons-la pour la première case du tableau. On a 121 candidatures du candidat maghrébin et 121 candidatures de référence, autrement dit la probabilité d'être maghrébin est d'une chance sur deux. De même on a 198 refus et 44 admissions, donc la probabilité de refus est de 198 sur 242. Au total la probabilité d'être maghrébin et d'avoir une candidature refusée est donc de 198 sur $2 * 242$. C'est la probabilité théorique d'être dans ce cas-là, sous l'hypothèse d'indépendance des variables. L'effectif théorique est égal à l'effectif total (242) multiplié par la probabilité théorique. On effectue ce calcul pour les quatre cases du tableau et on obtient le tableau suivant :

Z	Y	Candidat potentiellement discriminé	Candidat de référence
	Refus	99	99
	Acceptation	22	22

On note $E_{Y;Z}$ les effectifs théoriques correspondant aux couples $(Y;Z)$.

Comme on le voit, dans les effectifs théoriques on obtient le même nombre de refus pour chaque candidat, puisqu'on a supposé que la variable qui différencie les candidats (l'origine Y) était indépendante du succès Z de la candidature. Le test consiste à déterminer si la différence qu'on observe entre nos données et ces effectifs théoriques est significative.

La statistique de test du χ^2 est définie ainsi :

$$U = \sum_{Y=0}^1 \sum_{Z=0}^1 \frac{(X_{Y,Z} - E_{Y,Z})^2}{E_{Y,Z}}$$

Avec nos données, on obtient $U = 13,44$.

Sous l'hypothèse H_0 d'indépendance, U suit une loi du χ^2 à 1 degré de liberté. C'est une loi statistique tabulée, dont voici un extrait :

Quantile	0.05
Valeur	3.841

Cela signifie que la valeur critique à 5% vaut 3,841.

Ainsi on obtient une valeur de la statistique U ici nettement supérieure à la valeur critique. En effet, U vaut 13.44 quand il suffit que U soit supérieur à 3.481 pour que la probabilité que les variables soit indépendantes soit inférieure à 5%. Le test rejette donc l'hypothèse d'indépendance avec un niveau de confiance de 5%.

On peut également dire que la probabilité que les variables soient dépendantes est supérieure à 95%, ou encore dans notre contexte que la probabilité qu'il y ait une discrimination est supérieure à 95%.

Après avoir calculé U , on peut avoir un calcul plus fin de la probabilité qu'il y ait une discrimination en se rapportant au tableau suivant :

Si le résultat est supérieur à :		La probabilité que les variables soient liées est de :
2,71	⇒	90%
3,84	⇒	95%
5,41	⇒	98%
6,64	⇒	99%
10,83	⇒	99,9%

Tableau 1 - Probabilité de dépendance des variables

Dans notre exemple, U est égal à 13,44, la probabilité qu'il y ait une discrimination est donc supérieure à 99,9%.

Robustesse du résultat

Le problème principal compte tenu des données est le faible effectif. Différents critères sont utilisés par les statisticiens pour valider le résultat d'un test du χ^2 . Par exemple Cochran propose comme critère minimal que toutes les cases du tableau aient un effectif non nul et que 80% des cases (dans le cas de variables discrètes) ait un effectif au moins égal à 5.

Nos données sont donc acceptables dans l'exemple qui précède selon les critères de Cochran et Cramer (couramment utilisés), mais pas selon le critère de Kend. Il faut garder à l'esprit que ces critères sont de toute façon arbitraires.

Il semble donc, dans l'exemple qui précède, que le résultat ait une significativité statistique.

1.2.4 Le test de Student

Une analyse par séries

Dans cette approche, nous n'analysons pas la répartition des effectifs dans les cases d'un tableau de contingence (approche du χ^2 ou du test exact de Fisher) mais comparons les réponses aux candidats pour chaque série à l'intérieur de chaque mission.

Nous procédons comme suit : pour chaque poste candidaté et chaque série, nous établissons la moyenne du nombre de réponses positives reçues par chacun des groupes.

Pour que les réponses apportées aux candidats soient comparables entre elles, nous ne prenons pas en compte le nombre brut de réponses positives mais le ratio des réponses positives sur les réponses envoyées (la moyenne).

Nous réalisons toujours une comparaison entre deux groupes : les candidats potentiellement discriminés d'une part (candidats ayant un prénom d'origine maghrébine ou candidats âgés) et les candidats référents d'autre part. Dans le cas où dans une série, 3 groupes de candidatures différentes ont été envoyées (candidats ayant un prénom d'origine maghrébine, candidats âgés, candidats référents), le score des candidats potentiellement discriminés est égal à la moyenne des scores des deux groupes potentiellement discriminés.

Principe général

La question qui se pose est de savoir si au sein d'une entreprise, les candidatures des candidats référents ont obtenu les mêmes résultats que celles des autres candidats.

Ainsi, si nous considérons les résultats suivants :

		Groupe 1	Groupe 2		Groupe 1et2	Groupe 3	
Offre d'emploi	Série	Score des candidats ayant un prénom d'origine maghrébine	Score des candidats âgés		Score des candidats potentiellement discriminés	Score des candidats référents	Différence de score (D)
N°1	1	1	1		1	0	1
	2	0	np		0	1	-1
	3	1	0		0,5	1	-0,5
N°2	1	0	np		0	1	-1
	2	0	np		0	0	0
N°3	1	np	1		1	0	1

np : non pertinent, aucune candidature envoyée pour ce groupe dans cette série

Moyenne des différences	- 0,08
--------------------------------	--------

Nous constatons que la moyenne des différences est égale à -0,08. Cette moyenne observée est-elle, pour autant, significativement différente de zéro ?

Répondre à cette question revient à réaliser un test de Student pour observations paires⁴. Le risque d'erreur associé à ce test est fixé à 5%.

Mise en œuvre du test

1. Nous posons l'hypothèse que la moyenne des différences est égale à zéro

$H_0 : \mu_D = 0$

Contre l'hypothèse alternative H_1 selon laquelle la moyenne est différente de zéro

$H_1 : \mu_D \neq 0$

2. Nous calculons la valeur t observée, définie par la formule

$$t_{obs} = \frac{D_{moy} - 0}{SD / \text{racine}(N)}$$

Où,

- D_{moy} est la moyenne des différences
- SD est l'écart-type des scores de différence
- N est le nombre de lignes dans le tableau

3. Nous comparons la valeur du t observé à la valeur présente dans la table de Student pour le nombre de degrés de liberté concerné ($ddl = N-1$) et pour le risque choisi

4. Nous décidons de rejeter H_0 au profit de H_1 si la valeur absolue de t observée est supérieure à la valeur lue dans la table. Sinon, nous ne pouvons pas rejeter l'hypothèse H_0

Dans notre exemple, la valeur de t_{obs} est de : -0,2225

La valeur critique lue dans la table de Student pour 5 degrés de liberté avec un risque de 5% dans un test bilatéral est de : 2,571.

La valeur absolue du t observé étant inférieure à la valeur critique, nous ne rejetons pas l'hypothèse H_0 selon laquelle la moyenne des différences est égale à zéro.

Sur cet exemple nous ne concluons donc pas que les candidats des deux groupes ont obtenu des scores significativement différents.

1.2.5 Mode de calcul opérationnel

Opérationnellement, nous avons calculé nos trois tests par le logiciel de calcul statistique S-plus.

1.2.6 Remarques sur l'exigence de significativité

Nos tests établissent une discrimination parfois **significative** ou **très significative**. Dans certains cas, les échantillons et/ou les écarts de taux de succès entre les candidats ne sont pas assez importants pour établir aussi sûrement la discrimination au moment du test de discrimination. Un détail du CV peut avoir provoqué la décision non discriminante d'un recruteur (entreprise où a eu lieu un stage ou une expérience professionnelle par exemple) et c'est pourquoi plusieurs tests de discrimination doivent être menés. Si un seul test ne livre qu'un simple indice, plusieurs observations même non statistiquement significatives permettent de dégager **une tendance**.

⁴ voir par exemple HOWEL, D.C. (1998). *Méthodes statistiques en sciences humaines*. DeBoeck Université.

1.2.7 Nos résultats livrent une information dans le cadre retenu

Il est possible que nos observations soient contingentes. La pratique de certains recruteurs lors de la période choisie ne peut être généralisée à tous les recruteurs, à toute l'entreprise. Nos tests constatent donc une discrimination parfois très significative au moment du test (sur les mois du test), pour les offres auxquelles nous avons répondu et dans le cadre des candidatures factices qui sont les nôtres.

1.3 Procédure d'examen contradictoire des résultats

Nous avons demandé aux entreprises si elles avaient des remarques concernant la méthode et les résultats du test les concernant. Nous avons transmis toutes les informations qui pouvaient permettre de porter une appréciation la plus précise possible de la situation de chaque entreprise objet du test. Les entreprises peuvent estimer que des limites inhérentes au processus de fabrication et d'envoi de CV ou de relevé des réponses lors d'un test de situation peuvent avoir affecté certaines observations et par conséquent le résultat constaté.

Par exemple, lorsque nous construisons nos paires de CV équivalents entre eux, nous ne pouvons évidemment pas construire des CV strictement identiques, sinon l'entreprise comprendrait qu'un test de situation est en cours ce qui fausserait évidemment les résultats. Nous introduisons donc de légères différences dans le fond et la forme des CV. Il est tout à fait envisageable que ces différences qui nous paraissent légères ne le soient en fait pas aux yeux du recruteur chargé de convoquer les candidats en entretien. Il est possible d'autre part que des candidatures ne soient pas parvenues au destinataire ou que nos relevés de réponse ne convergent pas avec ceux de l'entreprise.

Nous avons tenu compte des observations des entreprises

Pour garantir l'équivalence des CV présentés il a été procédé à une nouvelle expertise indépendante de nos séries de CV qui a permis de montrer que les CV des candidats susceptibles d'être discriminés sont supérieurs en moyenne aux CV des candidats de référence alors même que nous avons expertisé des CV pour lesquels une discrimination était constatée.

Nous avons suivi les observations des entreprises en retenant leur bonne foi sauf lorsque nous avons la certitude que l'information fournie était inexacte.

Nous avons examiné si la présence d'une lettre de motivation influençait le jugement des recruteurs et pouvait avoir faussé nos résultats ce qui n'est pas le cas.

Nous avons écarté les candidatures pour lesquelles un doute existait sur le processus de transmission et retenu ici encore la bonne foi des entreprises.

Ce processus a conduit à retenir 4691 candidatures valides. 929 CV ont été supprimés.

Les CV supprimés proviennent de :

- Séries déséquilibrées :
 - 1 seul CV gardé pour une série car par exemple l'autre CV a été invalidé
 - Séries avec un candidat référent et deux candidats d'origine : Dans ce cas, nous avons supprimé une des 2 candidatures d'origine. La suppression d'un des deux candidats est réalisée aléatoirement (loi uniforme sans remise, chaque occurrence a la même probabilité d'être choisie)
- Autres raisons de suppressions suite aux vérifications :
 - Pas de trace vidéo
 - CV non associés à une offre en particulier
 - Incertitude sur la fiabilité de l'adresse courriel ou du téléphone du candidat
 - 2 CV d'une même série pouvant être jugés non équivalents

1.3.1 Qualité et équivalence des CV

- Contestation de l'équivalence des CV
 - o Nous acceptons directement ou faisons expertiser en cas de divergence.
- Rejet d'une candidature pour un motif ignoré de nous (contrat liant un intermédiaire de l'emploi à ses clients, entreprise non connue de nous lors de la postulation, etc.)
 - o Nous acceptons ces explications sans examen
- Le niveau de nos candidatures serait pauvre et expliquerait le rejet des candidats
 - o Le taux élevé de succès de nos candidats et l'évaluation indépendante d'un large échantillon de CV atteste de la qualité des CV présentés. Qui qu'il en soit cette observation quant à la qualité globale de nos CV ne change rien au résultat du test.
- La présence non systématique d'une lettre de motivation favoriserait certains candidats.
 - o En moyenne, la présence d'une lettre n'a pas l'effet attendu sur le taux de succès des candidatures.

1.3.2 Envois

- Contestation de l'envoi d'une candidature
 - o Nous ne suivons pas l'entreprise si nous avons en notre possession toutes les preuves de l'envoi.

1.3.3 Réponses

- Déclaration d'une réponse positive par une entreprise dont nous n'avons pas la trace.
 - o Nous acceptons faute de preuve du contraire.
- Déclaration de réponse négative ou de non réponse par l'entreprise.
 - o Nous avons la preuve du contraire et nous ne suivons pas l'entreprise (message audio par exemple).
- Réponse pour une autre offre que l'offre initialement testée.
 - o Nous ne pouvons distinguer les réponses pour une offre ou pour une autre et les gardons toutes.
 - o Dans le cas d'une entreprise qui fournit une liste des réponses pour les offres ou d'autres offres nous suivons la règle souhaitée par elle, acceptons l'information fournie et ne gardons que les réponses pour l'offre

1.3.4 Traitements statistiques

- Puisqu'un test statistique n'autorise pas à conclure ceci met en cause tout le traitement statistique
 - o Nous ne tenons pas compte de cette objection car les tests sont effectués pour chaque type de discrimination
- Nos séries comportent plus de 2 candidats
 - o Nous avons supprimé dans toutes les entreprises testées toutes les séries « déséquilibrées » (deux candidats potentiellement discriminés et un candidat de référence par exemple)

- Nos candidatures sont équivalentes au sein de séries mais pas comparables au sein d'une offre (il y a plusieurs séries par offre et chaque série a des caractéristiques différentes)
 - o Nous faisons tous nos calculs par série

1.3.5 Principe du test

- Candidature concernant une filiale, une « franchise », un « managé »
 - o Nous utilisons le site de recrutement de l'entreprise, ne pouvons sélectionner les offres en fonction de la situation juridique d'ailleurs changeante des entités et maintenons toutes les candidatures
- L'entreprise a recruté un autre candidat susceptible d'être discriminé que celui que nous avons proposé
 - o Nous ne retenons pas cette explication car elle ne remet pas en cause la validité du test
- L'entreprise aurait recruté un pourcentage élevé de candidats susceptible d'être discriminés en fournissant des évaluations statistiques de ces proportions
 - o Nous ne retenons pas cette explication car elle ne remet pas en cause la validité du test

2 Résultats généraux

2.1 Contexte du test

2.1.1 Périmètre de l'étude

Le test s'est déroulé du 03/07/2007 au 21/12/2007.

Nous avons postulé sur 1327 offres, par le biais de 2288 séries de 2 ou 3 candidats (couples ou triplets). Une ou plusieurs série(s) de CV peuvent être envoyées pour chaque offre.

Cas n°1 : Série où 1 candidat de référence et 1 candidat âgé ont postulé.

Cas n°2 : Série où 1 candidat de référence et 1 candidat au nom et prénom maghrébin ont postulé.

Cas n°3 : Série où 1 candidat de référence, 1 candidat au nom et prénom maghrébin et 1 candidat âgé ont postulé.

Tableau de répartition des séries :

	Effectifs	%
Cas n°1	316	13,81
Cas n°2	1857	81,16
Cas n°3	115	5,03

Au total, 4691 candidatures ont été envoyées.

- 2288 CV de candidats de référence soit 48,77 % des candidats
- 1972 CV de candidats susceptibles d'être discriminés selon leur origine soit 42,04 % des candidats
- 431 CV de candidats âgés soit 9,19 % des candidats

2.1.2 Délais de réponse

Le délai moyen de réponse est de 21 jours. Le dernier envoi a été réalisé le 21/12/2007.

2.1.3 Répartition des réponses positives

Nous avons reçu 648 réponses positives, ainsi réparties entre les différents candidats :

- 366 réponses positives pour les candidats de référence
- 234 réponses positives pour les candidats susceptibles d'être discriminés selon leur origine
- 48 réponses positives pour les candidats âgés

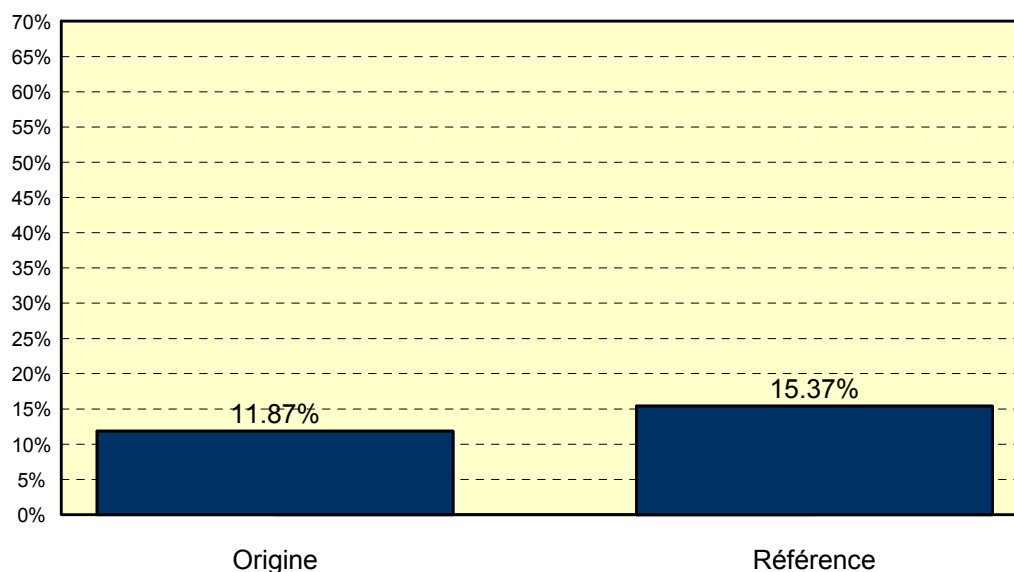
Les réponses positives reçues conduisent à un taux de réponse positive moyen de 13,81 %.

2.2 Synthèse des résultats

2.2.1 Discrimination liée à l'Origine

Afin de savoir si une discrimination existe, nous comparons le taux de réponses positives des candidats de référence à celui des candidats aux noms et prénoms originaires du Maghreb ou d'Afrique Sub-saharienne.

Taux de réponses positives



Le tableau suivant présente le nombre de réponses positives reçues par type de candidat lorsqu'ils postulaient sur les mêmes offres.

	Origine	Référence	Total
Nombre de CV envoyés	1972	1972	3944
Nombre de réponses positives	234	303	537
Taux de réponses positives	11,87	15,37	13,62

Le taux de convocation pour les candidats de référence est de 15,37%. Le taux de convocation pour les candidats susceptibles d'être discriminés est de 11,87%.

Soit 22,77% de chances en moins pour ces derniers d'être convoqués en entretien.

Reste à savoir si ces écarts sont statistiquement significatifs. Nous utiliserons pour cela le test exact de Fisher, le test du Chi-Deux et le T-Test apparié par série.

D'après le test de Fisher, la probabilité qu'il n'y ait pas de discrimination est de 0,1575%. Le risque d'erreur associé à ce test est fixé à 5%.

Le test exact de Fisher permet de conclure à l'existence d'une discrimination.

D'après le test du Chi-Deux, la probabilité qu'il n'y ait pas de discrimination est de 0,1593%. Le risque d'erreur associé à ce test est fixé à 5%.

Le test du Chi-deux permet de conclure à l'existence d'une discrimination.

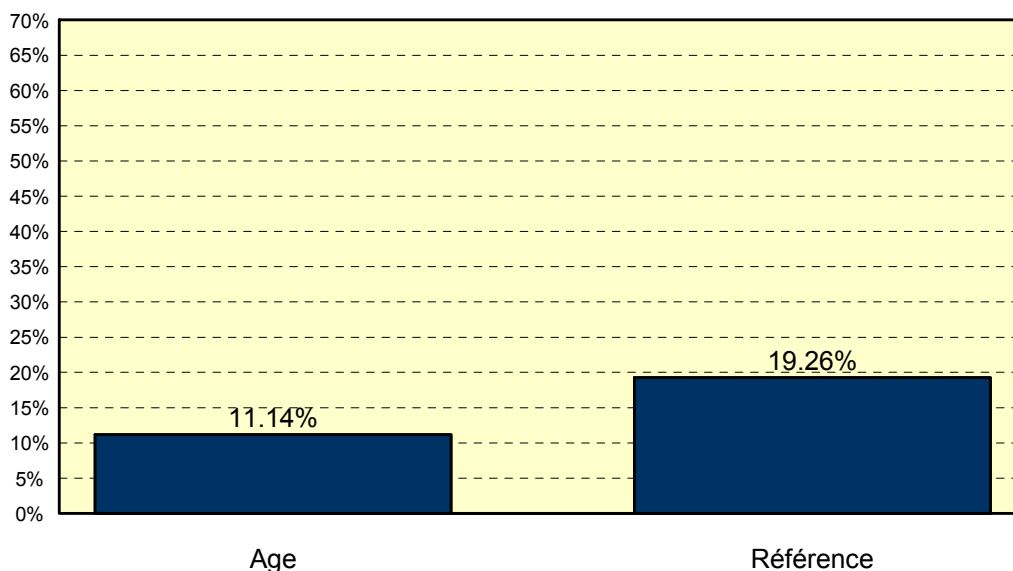
D'après le test de Student apparié, la probabilité qu'il n'y ait pas de discrimination est de 0,0018%. Le risque d'erreur associé à ce test est fixé à 5%.

Le test de Student apparié permet de conclure à l'existence d'une discrimination.

2.2.2 Discrimination liée à l'âge

Afin de savoir si une discrimination existe, nous comparons le taux de réponses positives des candidats de référence à celui des candidats âgés.

Taux de réponses positives



Le tableau suivant présente le nombre de réponses positives reçues par type de candidat lorsqu'ils postulaient sur les mêmes offres.

	Age	Référence	Total
Nombre de CV envoyés	431	431	862
Nombre de réponses positives	48	83	131
Taux de réponses positives	11,14	19,26	15,2

Le taux de convocation pour les candidats de référence est de 19,26%. Le taux de convocation pour les candidats susceptibles d'être discriminés est de 11,14%.

Soit 42,17% de chances en moins pour ces derniers d'être convoqués en entretien.

Reste à savoir si ces écarts sont statistiquement significatifs. Nous utiliserons pour cela le test exact de Fisher, le test du Chi-Deux et le T-Test apparié par série.

D'après le test de Fisher, la probabilité qu'il n'y ait pas de discrimination est de 0,1195%. Le risque d'erreur associé à ce test est fixé à 5%.

Le test exact de Fisher permet de conclure à l'existence d'une discrimination.

D'après le test du Chi-Deux, la probabilité qu'il n'y ait pas de discrimination est de 0,1256%. Le risque d'erreur associé à ce test est fixé à 5%.

Le test du Chi-deux permet de conclure à l'existence d'une discrimination.

D'après le test de Student apparié, la probabilité qu'il n'y ait pas de discrimination est de 0,0027%. Le risque d'erreur associé à ce test est fixé à 5%.

Le test de Student apparié permet de conclure à l'existence d'une discrimination.