Test d'adéquation à une loi uniforme, (un exemple de test du χ^2)

1 Enoncé

Un meunier a besoin, pour sa farine, d'un mélange de quatre variétés différentes de grains de blé, d'égales quantités chacunes et notées 1, 2, 3, 4.

1. Il veut savoir si, dans son silo, les différentes variétés sont bien mélangées. Pour cela, il prélève, à la sortie du silo, un échantillon de 100 grains de blé rendus radioactifs par des marqueurs différents selon les variétés? Il obtient les résultats suivant:

Variété	1	2	3	4
Nombre de grains radioactifs	18	27	35	20

   Le meunier veut savoir si ces données sont vraisemblables lorsque l'on fait l'hypothèse d'un mélange homogène des quatres variétés, ce qui correspond à un quart des marqueurs pour chaque variété.

   On appelle f_i la fréquence dans l'échantillon de la variété i et on pose

$$d^2 = 400 \sum_{i=1}^{4} \left(f_i - \frac{1}{4} \right)^2 .$$

   Calculer la valeur de d^2.

2. On suppose l'équiprobabilité de la présence d'un grain de blé de chaque variété et on simule 10 000 séries de 100 tirages de grains de blé.

   Pour chaque série de 100 tirages, on calcule d^2. Le tableau suivant donne le nombre de séries pour lesquelles la valeur de d^2 est strictement supérieur à l'entier j:

j	3	4	5	6	7	8	9	10
Nombre des séries telles que $d^2 > j$:	3915	2618	1715	1114	728	467	306	180

   Lire la valeur du 9-ème décile, arrondie à l'entier le plus proche, puis celle du 95-ème centile.

3. Si l'hypothèse d'équiprobabilité est vraie:

   (a) Peut-on affirmetr avec un risque d'erreur de 10% que le mélange étudié à la question 1. est homogène?

   (b) Même question avec un risque d'erreur de 5%.

   (c) Que peut-on dire si quelqu'un demande un risque d'erreur de 0%?

2 Travail demandé au candidat

1. Quelles sont les compétences néccessaires à l'élève pour faire cet exercice, et à l'enseignant pour être à même de l'expliquer devant une classe?

2. Quels sont les deux risques d'erreurs lorsque l'on effectue un tel test?

3. Pouvez expliquez l'intérêt de mettre les grands nombres : 100 et 10 000 dans l'énoncé?

   Pourquoi 10 000 simulations? 100 ne suffiraient-elles pas?

4. Pourriez vous donnez la valeur moyenne de d^2?
5. Que pensez vous du cas où le meunier aurait trouvé autant de grains de blé pour chaque variété?
6. Cet exercice introduit le test du χ^2 dans le cas d'une éventuelle adéquation à une loi équiprobable. Sauriez vous donnez la formule de d^2 pour cinq variétés de grain?
7. Proposez un test trés simple, et même un exercice, dans le cas ou l'on dispose de seulement 2 variétés de grain, sachant que l'échantillon du Meunier contient 55 grains de la variété 1?

source Documents d'accompagnement pour les séries S et ES, CNDP, 2004.