**Test khi deux**

*Principe :*

Le test du c2 (prononcez khi deux ou khi carré) fournit une méthode pour déterminer la nature d'une répartition, qui peut être continue ou discrète.

À la base d'un test de statistique classique, il y a la formulation d'une hypothèse appelée [hypothèse nulle](https://fr.wikipedia.org/wiki/Hypoth%C3%A8se_nulle) (ou hypothèse zéro)[N 1](https://fr.wikipedia.org/wiki/Test_du_%CF%87%C2%B2#cite_note-2), notée H0. Elle suppose que les données considérées proviennent de variables aléatoires suivant une loi de probabilité donnée, et l'on souhaite tester la validité de cette hypothèse.

1. On répartit les valeurs de l'échantillon (de taille n) dans k classes distinctes et on calcule les effectifs de ces classes. Il faut vérifier que pour les i de 1 à k, on a npi(1-pi) https://www.apprendre-en-ligne.net/crypto/icones/supegal.gif 5 (éventuellement répartir les valeurs autrement). Appelons oi (i=1,...,k) les effectifs observés et ei les effectifs théoriques.

|  |  |  |
| --- | --- | --- |
|  |  | |
|  |  | 0.05 | 0.01 |
| k-1 | 1 | 3.84 | 6.64 |
| 2 | 5.99 | 9.21 |
| 3 | 7.82 | 11.35 |
| 4 | 9.49 | 13.28 |
| 5 | 11.07 | 15.09 |
| 6 | 12.59 | 16.81 |
| 7 | 14.07 | 18.48 |
| 8 | 15.51 | 20.09 |
| 9 | 16.92 | 21.67 |
| 10 | 18.31 | 23.21 |
| 11 | 19.68 | 24.73 |
| 12 | 21.03 | 26.22 |

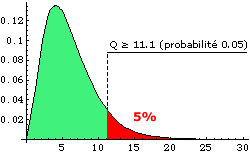
1. Calcule https://www.apprendre-en-ligne.net/random/Q.gif  
   La statistique Q donne une mesure de l'écart existant entre les effectifs théoriques attendus et ceux observés dans l'échantillon. En effet, plus Q sera grand, plus le désaccord sera important. La coïncidence sera parfaite si Q=0.
2. On compare ensuite cette valeur Q avec une valeur https://www.apprendre-en-ligne.net/random/khi.gif issue d'un [tableau](https://www.apprendre-en-ligne.net/random/tablekhi2.html) (voir extrait ci-contre) à la ligne k-1 et à la colonne a. k-1 est le nombre de degrés de liberté et a la tolérance.
3. Si Q > https://www.apprendre-en-ligne.net/random/khi.gif, et si n est suffisamment grand, alors l'hypothèse d'avoir effectivement affaire à la répartition théorique voulue est à rejeter avec une probabilité d'erreur d'au plus a.

*Exemple :*

On a lancé un dé 90 fois et on a obtenu les issues 1 à 6 (k=6) avec les effectifs suivants: 12, 16, 20, 11, 13, 18 (on a vérifié que 90 lancers sont suffisants: n(1/6)(5/6) https://www.apprendre-en-ligne.net/crypto/icones/supegal.gif 5 implique que n https://www.apprendre-en-ligne.net/crypto/icones/supegal.gif 36).   
Si le dé n'est pas pipé (notre hypothèse), on attend comme effectifs moyens théoriques 15 pour toutes les issues.

https://www.apprendre-en-ligne.net/random/Qex.gif

Pour k-1=5 degrés de liberté et un seuil de tolérance de 5%, la valeur https://www.apprendre-en-ligne.net/random/khi.gif du tableau est 11.1. Cela signifie que la probabilité que Q soit supérieur à 11.1 est de 5% (voir figure ci-dessous). Comme 4.266 < 11.1, on accepte l'hypothèse selon laquelle le dé est régulier.



Fonction de répartition de la loi du khi deux pour 5 degrés de libertés

*Test MATLAB*

1. Créez un objet de la distribution de probabilité normale standard. Générer un vecteur x de 100 va selon de la distribution.

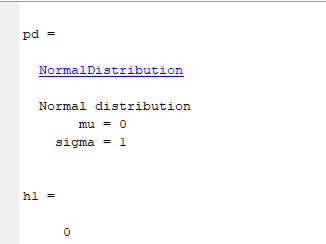
pd = makedist('Normal');

x = random(pd,100,1);

1. Tester l’hypothèse nulle que les données en x proviennent d’une population avec une distribution normale.

h = chi2gof(x)

[h](https://fr.mathworks.com/help/stats/chi2gof.html#btv1j1v-1_sep_shared-h) = chi2gof([x](https://fr.mathworks.com/help/stats/chi2gof.html#btv1j1v-1-x)) retourne une décision de test de l’hypothèse nulle que les données de vecteur x provient d’une distribution normale avec une moyenne et la variance estimée à partir de x, en utilisant le [test du Khi-deux d’ajustement](https://fr.mathworks.com/help/stats/chi2gof.html#btv1j1v-11). L’autre hypothèse est que les données ne provient pas d’une telle distribution. Le résultat de h est 1si le test rejette l’hypothèse nulle au seuil de signification de 5 % et 0 sinon.

 Exemple de programme Résultat

