

ASI (L2) : TP3 Calculs probabilistes avec *Excel* et *Rstat*

Objectifs du TP :

Savoir utiliser Excel et Rstat pour calculer des moyennes pondérées, des variances pondérées et savoir faire des approximations par la loi binomiale et la loi de Poisson y compris le calcul du χ^2 d'adéquation.

1. Calculs probabilistes avec Excel

1.1 Premiers calculs

Soit X une variable aléatoire qui prend ses valeurs x_i avec les probabilités respectives p_i définies par :

x_i	0	1	2	5	9
p_i	0.30	0.20	0.2	0.1	(illisible)

Trouver la probabilité manquante puis calculer la moyenne pondérée de X , sa variance pondérée et son écart-type avec Excel. On commencera par créer une colonne correspondant à X^2 et on calculera la moyenne des carrés.

On lira l'aide sur la fonction `SOMMEPROD` pour trouver une solution plus courte.

On comparera avec Excel $m(X^2) - m(X)^2$ et $m((X - m(X))^2)$. A quelles phrases correspondent ces formules ?

Réponse : "la différence entre la moyenne des carrés et le carré de la moyenne", "la moyenne du carré des écarts à la moyenne".

Calculer avec Excel la variable centrée issue de X , la variable réduite issue de X puis la centrée réduite issue de X .

1.2 Excel comme outil probabiliste exploratoire

Peut-on trouver deux valeurs y_1 et y_2 équiprobables dont la moyenne est $m_y = 3,123456$ et la variance est $V_y = 1,654321$? Pourquoi ?

Idem avec z_1 et z_2 pour $V_z = 3,123456$ et $m_z = 1,654321$.

1.3 Les moyennes

Calculer les moyennes arithmétique, géométrique, harmonique et quadratique avec Excel des deux valeurs $x = 2$ et $y = 8$.

Rappel :

$$m_a(x, y) = (8 + 2)/2 = 10/2 = 5.$$

$$m_g(x, y) = \sqrt{8 + 2} = \sqrt{16} = 4$$

$$m_h(x, y) = 1/((1/2) + (1/8)) = 8/5 = 1.6$$

$$m_q(x, y) = \sqrt{(2^2 + 8^2)/2} = \sqrt{34} = 5.83$$

Essayer avec les n valeurs 1,2, 3... 10. Au passage, qui connaît les formules ? Le "formulaire des lois statistiques" du cours peut-il nous aider ?

<http://www.info.univ-angers.fr/pub/gh/wstat/tablois.ps>

Réponse :

Oui, car c'est la loi uniforme.

Comment retrouver un de ces résultats avec le panneau des lois statistiques discrètes usuelles à l'adresse ?

<http://www.info.univ-angers.fr/pub/gh/vitrine/Democgi/loisStatp.htm>

Réponse :

Taper U 1 10.

1.4 Valeurs de la loi Binomiale

Comment se sert-on de la fonction `LOI.BINOMIALE` pour afficher les valeurs de la loi binomiale $\mathcal{B}(30; 0,04)$ avec Excel ? Idem avec la page web `loisStatp`.

1.5 Approximation par la loi Binomiale (début)

Des essais récents sur 145 échantillons de 30 cartes-mères de P8 fournissent le tableau suivant de nombre de cartes-mères défectueuses :

x_i	0	1	2	3	4	5	6	7	...	30
n_i	44	55	29	10	5	1	1	0	...	0

Quelle est la somme totale des effectifs ?

Quelle est la loi binomiale candidate ?

Quelles sont les probabilités associées fournies par Excel ?

Quels sont les effectifs entiers correspondants ?

Réponse :

On vérifie que la somme totale des effectifs $Nt = \sum n_i$ vaut bien 145. La loi binomiale à considérer est ici $B(n, p)$ avec $n = 30$. Pour p , puisque la moyenne pondérée vaut $m = \sum x_i.n_i/Nt = 1.2 = n.p$, on déduit que $p = 0.04$. Les 12 premières valeurs de la loi binomiale $\mathcal{B}(30; 0, 04)$ ainsi que leur produit par 145 sont

x_i	p_i	$p_i.145$	<i>arrondi</i>
0	0.2938576	42.6093520	43
1	0.3673221	53.2617045	53
2	0.2219237	32.1789365	32
3	0.0863037	12.5140365	13
4	0.0242729	3.5195705	4
5	0.0052591	0.7625695	1
6	0.0009130	0.1323850	0
7	0.0001304	0.0189080	0
8	0.0000156	0.0022620	0
9	0.0000016	0.0002320	0
10	0.0000001	0.0000145	0
11	0.0000000	0.0000000	0

Les termes d'ordre supérieur à 12 sont négligeables. La somme des 9 premières probabilités est 0.9999825, soit après multiplication par 145, un total de 144.9974625 ce qui ressemble beaucoup à 145.

1.6 Calcul du χ^2

Rappeler la formule de comparaison du χ^2 .

Quelles sont les conditions d'application ? On pourra utiliser la page

<http://www.info.univ-angers.fr/pub/gh/wstat/formules.ps>

Donner la valeur du χ^2 pour

O_i	5	12	11	6
T_i	7	10	10	7

A quelle valeur de la table `k:\stat_ad\chi2.dbf` faut-il comparer au risque de première espèce $\alpha = 5\%$? Fonction Excel correspondante ?

Que peut-on en conclure ?

Réponse :

Valeurs observées :

5 12 11 6

Valeurs théoriques :

7 10 10 7

Somme des théoriques : 34 = somme des observées.

CHI-DEUX avec 4 valeurs = 1.21428571 pour 3 ddl (degrés de liberté)

Détail :

Différences	2.00	-2.00	-1.00	1.00
Cumul	0.57	0.97	1.07	1.21

Valeur de chi2 lue dans la table a 5 %

en arrondi : 7.815 pour 3 ddl au seuil de 5 %

Conclusion : au risque... on accepte...

La fonction Excel correspondante (comme indiqué dans `formules.ps`) est nommée `KHIDEUX.INVERSE`. Vérifier que `KHIDEUX.INVERSE(0.05;3)` fait bien 7,81.

1.7 Approximation par la loi Binomiale (fin)

Terminer l'approximation par la loi Binomiale des lots de cartes-mères défectueuses.

Réponse :

Pour appliquer le calcul du χ^2 , il faut regrouper des classes. On obtient finalement les valeurs théoriques 43, 53, 32 et 17 mises en regard des valeurs observées 44, 55, 29 et 17.

Le χ^2 est alors de 0,38 et il indique que ce sont des valeurs proches (le χ^2 maximal est 7.81 pour 3 ddl). On écrit donc la phrase "légale" : *au risque de 5 % on accepte l'hypothèse que les valeurs observées suivent la loi théorique $\mathcal{B}(30; 0,04)$.*

2. Calculs probabilistes avec Rstat

Reprendre le calcul de m_X et σ_X puis celui de la variable centrée issue de X , de la variable réduite issue de X et enfin de la variable centrée réduite issue de X .

Vérifier les calculs de moyennes.

Montrer la fonction `dbinom` (passer sous silence pour l'instant `pbinom`, `qbinom` et `rbinom`) pour afficher les valeurs de $\mathcal{B}(30; 0,04)$.

Utiliser le fichier `k:\stat_ad\tp3_cm.r` à l'aide de l'instruction *source* pour éviter de saisir les valeurs pour les lots de carte-mère défectueuses.

Essayer de faire l'approximation binomiale pour ces lots, y compris le calcul du χ^2 .

Montrer le fichier `rsexos.ps` cité sur la page

<http://www.info.univ-angers.fr/pub/gh/wstat/asi.htm>