

Master 1 ingénierie TISD DM 3

Emeline Schmisser, `emeline.schmisser@math.univ-lille1.fr`, bureau 314 (bâtiment M3).

Exercice 1 (sondage et χ^2 d'adéquation)

1. Télécharger les tables `sondage_primaire_1` et `sondage_primaire_2` et les ouvrir dans R. Attention : les colonnes et les lignes ont des noms.

Ces fichiers donnent les sondages réalisés pour le premier et le deuxième tour de la primaire socialiste (ainsi que les résultats pour le premier tour). Nombre correspond au nombre de personnes interrogées (ou au nombre de votants). Les noms des colonnes correspondent à l'institut de sondage et à la date (OP pour Opinion Way, I pour Ipsos).

On veut vérifier que les sondages correspondent bien à ce qui a été observé. Pour cela, on va faire un test du χ^2 d'adéquation.

2. Rappeler le test du χ^2 d'adéquation.
3. Créer un vecteur `resultat` qui contient les résultats de la primaire. Créer une table `P1` qui contient, pour chaque sondage, le nombre de gens qui veulent voter pour chaque candidat. `P1` ne contient pas les résultats de l'élection. `P1` doit correspondre aux valeurs n_i .
Indication : les tableaux donnent des pourcentages. Pensez à la fonction `as.numeric` pour créer le vecteur `resultat`.
4. Créer la table `P2`, de même taille que `P1`, qui correspond aux valeurs $n * p_i$.

5. Calculer les 10 statistiques du χ^2 .

Indication : on peut ne pas utiliser de boucles grâce aux commandes `rowSums` et `colSums`.

6. Quelle loi suivent ces statistiques (préciser le degré de liberté). À votre avis, qu'est-ce qui peut expliquer ces résultats ?
7. Pourquoi est-ce que cela aurait été faux de comparer directement les sondages aux résultats (sans reconstituer d'abord les effectifs) ?
8. La fonction `chisq.test` permet de faire directement un test du χ^2 . Entrer la commande `chisq.test(P1[1,],as.numeric(resultats))`. Est-ce qu'on trouve la même chose que précédemment ? Si non, modifier la commande pour obtenir le résultat souhaité.
9. Utiliser la commande `chisq.test` pour tester si le dernier sondage sur le second tour des primaires correspond aux résultats.

Exercice 2 (Sondage et indépendance)

On va travailler avec le tableau contenant les sondages du second tour de la primaire. On souhaite savoir si les différents sondages suivent bien la même loi.

1. Rappeler le test du χ^2 d'indépendance.
2. Créer une table `P3` qui contient, pour chaque sondage, le nombre de personnes qui veulent voter pour chaque candidat. Créer deux vecteurs qui contiennent les totaux de chaque ligne et de chaque colonne.
3. Créer une table `P4` qui contienne les valeurs attendues.

4. Calculer la statistique de test.
5. Quelle loi suivent cette statistique ?
6. Refaire le test avec R. Conclure.
On s'intéresse maintenant aux trois derniers sondages.
7. Utiliser R pour faire le test du χ^2 pour les trois sondages les plus récents. Quelle est la probabilité d'avoir une statistique de test plus petite ? Qu'est-ce que cela pourrait indiquer ?
8. Refaire le test avec les deux sondages les plus anciens. Conclure

Exercice 3 (Gini et courbe de Lorentz)

Les salaires d'une entreprise sont répartis comme suit :

Proportion	Salaires
10%	1000
30%	1200
20%	1300
15%	1500
10%	1700
10%	1900
5%	2300

1. Donner la formule théorique pour calculer la courbe de Lorentz en fonction des vecteur Salaires s et Proportion p .
2. Donner la formule théorique du coefficient de Gini (la calculer).
3. Tracer sur un même graphique la courbe de Lorentz et la première bissectrice pour cette entreprise.
4. Calculer le coefficient de Gini.