

FLUCTUATIONS D'ÉCHANTILLONNAGE

Simulations avec XCAS...

Le problème

Le but de notre séance de vérifier un résultat admis en 2^{nde} concernant les sondages

Partie 1: SONDAGES ET URNES

1. Expérience

Une urne contient un très grand nombre de boules dont 40% sont rouges. On en tire un échantillon de taille 100 au hasard.

On mesure la fréquence de boules rouges obtenues.

Un résultat classique en statistique affirme que la probabilité que la fréquence observée soit dans l'intervalle $\left[0,40 - \frac{1}{\sqrt{100}}, 0,40 + \frac{1}{\sqrt{100}}\right]$ est au moins supérieure à une valeur que nous allons essayer de déterminer expérimentalement.

2. Que voulons-nous que l'ordinateur fasse ?

Nous allons généraliser le problème.

Nous allons demander à l'ordinateur d'effectuer un grand nombre N de tirages d'échantillons de taille n, sachant que la proportion de boules rouges est p.

Nous allons représenter le nuage des points d'ordonnées les fréquences de boules rouges obtenues par échantillon.

On trace également les droites d'équations :

$$y = p \pm \frac{1}{\sqrt{n}}$$

3. Procédure

Essayez d'analyser la procédure suivante :

```
echantillon(n,p,N):={
Freq:=NULL;
pour k de 1 jusque N faire
  S:=0;
  pour j de 1 jusque n faire
    si evalf(hazard(0,1))<p alors S+=1; fsi
  fpour
  Freq:=Freq,[k,evalf(S/n)];
fpour
P:=affichage(nuage_points([Freq]),epaisseur_point_3);
D1:=affichage(droite(y=p-1./sqrt(n)),bleu+epaisseur_ligne_3);
D2:=affichage(droite(y=p+1./sqrt(n)),rouge+epaisseur_ligne_3);
retourne(P,D1,D2)
};;
```

puis analysez le résultat.

4. À vous de jouer

Il faudrait légèrement modifier la procédure précédente pour obtenir le pourcentage de fréquences observées de boules rouges.

5. Le théorème

Le résultat suivant sera admis :

Théorème

Lorsqu'on prélève un échantillon de taille n dans une population où la fréquence d'un caractère est p , alors, sous certaines conditions, la probabilité que cet échantillon fournisse une fréquence appartenant à

$$I = \left[p - \frac{1}{\sqrt{n}}, p + \frac{1}{\sqrt{n}} \right]$$

est **au moins égale** à %

Partie 2: SONDAGES ET URNES : LES ÉLECTIONS

Vu à la radio : « En perte de vitesse ces derniers mois, la cote du président syldave est remontée de 49% à 52% après son mariage avec la princesse Carlotta de Bordurie qui a su séduire par ses sourires les électeurs de Syldavie. »

1. La véritable cote de popularité du président auprès des 60 millions de Syldave est inconnue. Désignons-la par p .
On connaît deux fréquences observées sur deux échantillons de 1 000 personnes : $f_1 = 49\%$ et $f_2 = 52\%$.
Montrez que $f_i \in I_i = \left[p - \frac{1}{\sqrt{n}}, p + \frac{1}{\sqrt{n}} \right]$ équivaut à $p \in J_i = \left[f_i - \frac{1}{\sqrt{n}}, f_i + \frac{1}{\sqrt{n}} \right]$.
2. Écrivez les intervalles J_1 et J_2 correspondant à f_1 et f_2 . Représentez-les sur une droite graduée. Commentez alors l'annonce faite à la radio.

Partie 3: SONDAGE PERSONNEL

Lu à la syldavision :

Le président syldave a affirmé hier lors de sa conférence de presse que sa cote de popularité est de 52%.

1. Un groupe dissident soutenu par la Bordurie effectue un sondage auprès de 625 personnes et obtient une cote de 47% d'opinions favorables : le président syldave aurait-il menti ?
2. Un autre sondage effectué auprès de 625 autres personnes donne une cote de 49%. Cela permet-il de confirmer la cote de 52% annoncée par le président syldave ?

Partie 4: Arnaque ?

La princesse Carlotta anime un jeu à la télévision. Elle annonce que dans une urne se trouvent 50 boules noires et 50 boules blanches. Elle offrira une photo dédicacée de son mari à toute personne obtenant au moins 11 boules blanches en tirant 30 boules au hasard de l'urne.

L'ambassadeur bordure participe au jeu. Il tire 30 boules et obtient 10 blanches seulement. Furieux, il décide de déclarer la guerre à la Syldavie affirmant que le jeu est truqué. Qu'en pensez-vous ? Parviendrez-vous à éviter la guerre ?