Exercice 110 page 407

110. Un fabricant de diodes électroluminescentes (LED) garantit que la probabilité p qu'une diode ne fonctionne pas vaut au plus 0,03. Pascal s'est fait livrer 5 000 diodes. On note X le nombre de diodes défectueuses parmi 5 000 diodes et $F = \frac{X}{5\,000}$ leur proportion.

1. Quelle est la loi suivie par X si $p = 0{,}03$?
2. D'après l'approximation de Moivre-Laplace, dans quel intervalle fluctue F avec une probabilité de 0,95 ?
3. On en déduit que X fluctue à plus de 95 % dans l'intervalle $[127\,;\,174]$. Vérifier cette affirmation en calculant $P(127 \leqslant X \leqslant 174)$.

Pascal a constaté que 172 diodes ne fonctionnaient pas dans le lot de 5 000 diodes qu'il a commandé. Il trouve ce nombre trop élevé. Le fabricant lui explique que si ce nombre appartient à l'intervalle de fluctuations à 95 % de X, il n'y a pas de raison de considérer ce lot comme non conforme.

4. Suivant cette règle de décision, quelle est la probabilité qu'un lot ne soit pas conforme ?
5. Suivant cette règle de décision, le lot commandé par Pascal est-il non-conforme ?
6. Si le lot de Pascal n'avait contenu aucune diode défectueuse, aurait-il été considéré comme conforme ?

Pascal trouve la règle de décision absurde. Il propose une autre règle : si le lot contient moins de 170 diodes défectueuses, alors il est jugé conforme, sinon il n'est pas jugé conforme.

7. Suivant la règle de décision de Pascal, quelle est la probabilité qu'un lot ne soit pas conforme ?
8. Un lot ne contenant aucune diode défectueuse serait-il jugé conforme ?
9. Quelle règle de décision vous paraît la plus adaptée au problème : celle du fabricant ou celle de Pascal ?
10. Selon sa règle de décision, le lot reçu par Pascal est-il conforme ?

1) $X \sim \mathcal{B}(5000\,;\,0{,}03)$ car répétition 5000 fois de façon identique et indépdte ...

2) $n = 5000$ donc $n \geq 30$

$np = 5000 \times 0{,}03 = 150$ donc $np \geq 5$

$n(1-p) = 4850$ donc $n(1-p) \geq 5$

$p - 1{,}96\,\frac{\sqrt{p(1-p)}}{\sqrt{n}} = 0{,}03 - 1{,}96\,\frac{\sqrt{0{,}03 \times 0{,}97}}{\sqrt{5000}} \simeq 0{,}025$ par défaut

$u_{0{,}05} = 1{,}96$

$p + 1{,}96\,\frac{\sqrt{p(1-p)}}{\sqrt{n}} = 0{,}03 + 1{,}96\,\frac{\sqrt{0{,}03 \times 0{,}97}}{\sqrt{5000}} \simeq 0{,}035$ par excès

$IF = [0{,}025\,;\,0{,}035]$

3) $P(127 \leq X \leq 174) = P(X \leq 174) - P(X \leq 126)$ (*) Bcd $P(X \leq k)$

$\approx 0{,}9769 - 0{,}0234$

$\approx 0{,}9535$ à 10^{-4} près

4) La probabilité qu'un lot ne soit pas conforme est inférieure ou égale à 0,05.

5) Non le lot est conforme car $172 \in [127\,;\,174]$

6) Le lot aurait été considéré comme non conforme car $0 \notin [127, 174]$

7) $P(X \geq 170) = 1 - P(X \leq 169) \simeq 1 - 0{,}94487 \simeq 0{,}0551$

8) Avec cette nouvelle règle, si aucune diode n'est défectueuse alors le lot est conforme

9) La règle de Pascal semble la plus adaptée.

10) Avec 172 leds défectueuses, le lot aurait été jugé non conforme.

Intervalles de fluctuation et de confiance

mercredi 17 juin 2020 12:16

Exercice 107 page 407

107. Selon l'INSEE, environ 51 % de la population française active ayant un emploi appartient à la catégorie socio-professionnelle (CSP) « employés et ouvriers ». Au concours d'entrée à l'ENA en 2009, 17 admis sur 81 ont des parents appartenant à cette CSP.

1. Quel est l'intervalle de fluctuations asymptotique de la fréquence d'apparition de la CSP « employés et ouvriers » dans un échantillon de 81 français tirés aléatoirement avec remise ?

2. Peut-on considérer que les admis d'origine socio-professionnelle « employés et ouvriers » sont sous-représentés à l'ENA ? On comparera leur proportion à l'intervalle de fluctuations précédent.

$n = 81$ $p = 51\% = 0,51$

* $n = 81$ donc $n \geq 30$
* $np = 81 \times 0,51 = 41,31$ donc $np \geq 5$
* $n(1-p) = 39,69$ donc $n(1-p) \geq 5$.

1) $p - 1,96\dfrac{\sqrt{p(1-p)}}{\sqrt{n}} = 0,51 - 1,96\dfrac{\sqrt{0,51 \times 0,49}}{\sqrt{81}} \approx 0,401$ par défaut

$p + 1,96\dfrac{\sqrt{p(1-p)}}{\sqrt{n}} \approx 0,619$ par excès

$IF = [0,401 ; 0,619]$ au seuil $0,95$.

2) $f = \dfrac{17}{81} \approx 0,21$ donc $f \notin [0,401 ; 0,619]$ donc les admis d'origine socio professionnelle « employés et ouvriers » sont sous représentés à l'ENA.

Intervalle de confiance

vendredi 19 juin 2020 09:08

Théorème 3.

Soit $X_n \sim \mathcal{B}(n;p)$ et p la proportion inconnue d'apparition d'un caractère. Pour n assez grand, $P\left(\frac{X_n}{n} - \frac{1}{\sqrt{n}} \leqslant p \leqslant \frac{X_n}{n} + \frac{1}{\sqrt{n}}\right) \geqslant 0{,}95$.

On sait que : Pour n suffisamment grand, $P\left(\frac{X_n}{n} \in \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]\right) \geq 0{,}95$ si $X_n \sim \mathcal{B}(n;p)$

ST

$$\frac{X_n}{n} \in \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right] \iff p - \frac{1}{\sqrt{n}} \leq \frac{X_n}{n} \leq p + \frac{1}{\sqrt{n}}$$

$-p$

$$\iff -\frac{1}{\sqrt{n}} \leq \frac{X_n}{n} - p \leq \frac{1}{\sqrt{n}}$$

$\times(-1)$

$$\iff \frac{1}{\sqrt{n}} \geq p - \frac{X_n}{n} \geq -\frac{1}{\sqrt{n}}$$

$$\iff -\frac{1}{\sqrt{n}} \leq p - \frac{X_n}{n} \leq \frac{1}{\sqrt{n}}$$

$+\frac{X_n}{n}$

$$\iff \frac{X_n}{n} - \frac{1}{\sqrt{n}} \leq p \leq \frac{X_n}{n} + \frac{1}{\sqrt{n}}$$

$$\iff p \in \left[\frac{X_n}{n} - \frac{1}{\sqrt{n}} ; \frac{X_n}{n} + \frac{1}{\sqrt{n}}\right]$$

Des évènements équivalents ont m̂ proba

donc $\underbrace{P\left(\frac{X_n}{n} \in \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]\right)}_{\text{supérieure ou égale à } 0{,}95} = P\left(p \in \left[\frac{X_n}{n} - \frac{1}{\sqrt{n}} ; \frac{X_n}{n} + \frac{1}{\sqrt{n}}\right]\right)$

Intervalles de confiance

Exercice 117 page 409

117. Selon le sondage « les Français et la musique », réalisé par Opinion Way pour la Sacem en 2011, le classement des styles musicaux qu'écoutent les 15-24 ans est le suivant : 1) RnB pour 45 % d'entre eux ; 2) Pop/Rock pour 42 % ; 3) rap pour 38 % ; 4) musiques électroniques pour 28 % ; 5) chanson française pour 26 %.

Ces pourcentages concernent les 15-24 ans de l'échantillon, ce qui représente 222 personnes. On considère que la population française est assez nombreuse pour qu'on puisse assimiler le tirage de l'échantillon à un tirage avec remise.

1. Donner un intervalle de confiance au niveau asymptotique 95 % pour estimer la proportion des 15-24 ans dans la population totale écoutant du RnB.

2. Même question pour le Pop/Rock ; le rap ; les musiques électroniques ; la chanson française.

3. On convient que si les intervalles de confiance de deux proportions se chevauchent, alors ces deux proportions ne peuvent pas être considérées comme étant différentes. Parmi les 15-24 ans de la population totale, peut-on considérer qu'il y en a significativement plus qui écoutent du RnB que du Pop/Rock ?

4. Selon le même critère, quels sont les styles musicaux significativement moins écoutés que le RnB parmi les 15-24 ans de la population totale ?

1) Pour le RnB : $IC_1 = \left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}}\right] = \left[0{,}45 - \frac{1}{\sqrt{222}} ; 0{,}45 + \frac{1}{\sqrt{222}}\right]$

$IC1 : [0{,}38 ; 0{,}52]$

2) Pour le Pop/Rock, $IC2 = \left[0{,}42 - \frac{1}{\sqrt{222}} ; 0{,}42 + \frac{1}{\sqrt{222}}\right] \approx [0{,}35 , 0{,}49]$

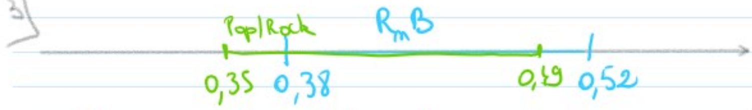
Pour le Rap : $IC_3 = \left[0{,}38 - \frac{1}{\sqrt{222}} ; 0{,}38 + \frac{1}{\sqrt{222}}\right] \approx [0{,}31 ; 0{,}45]$

Pour les musiques électroniques

$IC_4 = \left[0{,}28 - \frac{1}{\sqrt{222}} ; 0{,}28 + \frac{1}{\sqrt{222}}\right] \approx [0{,}21 ; 0{,}35]$

Pour la chanson française $IC_5 = \left[0{,}26 - \frac{1}{\sqrt{222}} ; 0{,}26 + \frac{1}{\sqrt{222}}\right] \approx [0{,}19 , 0{,}33]$

3)



Si les intervalles se chevauchent, l'intersection n'est pas vide : il y a des él^{ts} communs aux 2 intervalles donc on peut trouver une valeur p dans IC1 et dans IC2 les proportions ne sont pas différentes.

On ne peut pas dire qu'il y a plus de personnes qui écoutent du RnB que du Pop/Rock

4) Les styles moins écoutés que le RnB sont : les musiques électroniques, la musique française

car les intervalles ne se chevauchent

Intervalles de confiance

Exercice 122 page 411

122. Deux candidats s'affrontent lors du second tour de l'élection présidentielle. Un institut de sondage souhaite effectuer une enquête auprès d'un échantillon de n personnes tirées aléatoirement afin de connaître les intentions de vote. Étant donné la taille de la population par rapport à n, ce tirage peut être considéré comme un tirage avec remise. L'institut de sondage pense que le scrutin sera serré et souhaite effectuer un sondage suffisamment précis pour pouvoir prédire le vainqueur. Combien de personnes n l'institut doit-il interroger pour que les intervalles de confiance au niveau ~~asymptotique~~ 95 % ait une longueur inférieure à 0,02 ?

$IC = \left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}}\right]$ au seuil de 0,95.

Amplitude de l'intervalle $IC = f + \frac{1}{\sqrt{n}} - \left(f - \frac{1}{\sqrt{n}}\right) = f + \frac{1}{\sqrt{n}} - f + \frac{1}{\sqrt{n}}$

" $= \frac{2}{\sqrt{n}}$.

L'énoncé se traduit par : $\frac{2}{\sqrt{n}} < 0,02$

$\frac{2}{\sqrt{n}} < 0,02 \iff \frac{1}{\sqrt{n}} < 0,01 \iff \sqrt{n} > \frac{1}{0,01} \underset{(n>0)}{\iff} n > 100^2$

$n > 10\,000$

Il faudra interroger au min 10 001 personnes pour avoir un intervalle de confiance de longueur inférieur à 0,02.