Intervalle de fluctuation-intervalle de confiance
Extraits des programmes

Classe de seconde - statistiques et probabilités

Extrait du programme juillet 2009

Objectifs visés par l'enseignement des statistiques et probabilités à l'occasion de résolutions de problèmes dans le cadre de l'échantillonnage

- faire réfléchir les élèves à la conception et la mise en œuvre d'une simulation ;
- sensibiliser les élèves à la fluctuation d'échantillonnage, aux notions d'intervalle de fluctuation et d'intervalle de confiance et à l'utilisation qui peut en être faite.

CONTENUS	CAPACITES ATTENDUES	COMMENTAIRES
Échantillonnage Notion d'échantillon. Intervalle de fluctuation d'une fréquence au seuil de 95%*****. Réalisation d'une simulation.	- Concevoir, mettre en œuvre et exploiter des simulations de situations concrètes à l'aide du tableur ou d'une calculatrice. - Exploiter et faire une analyse critique d'un résultat d'échantillonnage.	Un échantillon de taille *n* est constitué des résultats de *n* répétitions indépendantes de la même expérience. À l'occasion de la mise en place d'une simulation, on peut : - utiliser les fonctions logiques d'un tableur ou d'une calculatrice, - mettre en place des instructions conditionnelles dans un algorithme. L'objectif est d'amener les élèves à un questionnement lors des activités suivantes : - l'estimation d'une proportion inconnue à partir d'un échantillon ; - la prise de décision à partir d'un échantillon

***** L'intervalle de fluctuation au seuil de 95%, relatif aux échantillons de taille *n*, est l'intervalle centré autour de *p*, proportion du caractère dans la population, où se situe, avec une probabilité égale à 0,95, la fréquence observée dans un échantillon de taille *n*. Cet intervalle peut être obtenu, de façon approchée, par simulation.

Le professeur peut indiquer aux élèves le résultat suivant, utilisable dans la pratique pour des échantillons de taille n > 25 et des proportions *p* du caractère comprises entre 0,2 et 0,8 : si *f* désigne la fréquence du caractère dans l'échantillon, *f* appartient à l'intervalle $\left[p-\frac{1}{\sqrt{n}}\,;\,p+\frac{1}{\sqrt{n}}\right]$ avec une probabilité d'au moins 0,95. Le professeur peut faire percevoir expérimentalement la validité de cette propriété mais **elle n'est pas exigible**.

<u>Classe de première</u> - Statistiques et probabilités

Dans le cas particulier d'expériences identiques et indépendantes à deux issues, on introduit la loi binomiale. L'acquisition de la loi binomiale permet de poursuivre la formation des élèves dans le domaine de l'échantillonnage et des procédures de prise de décision en contexte aléatoire.
L'objectif est d'amener les élèves à expérimenter la notion de « différence significative » par rapport à une valeur attendue et à remarquer que, pour une taille de l'échantillon importante, on conforte les résultats vus en classe de seconde.

CONTENUS	**CAPACITES ATTENDUES**	**COMMENTAIRES**
Échantillonnage <u>S-ES-L</u> Utilisation de la loi binomiale pour une prise de décision à partir d'une fréquence. <u>STL-STI2D</u> Utilisation de la loi binomiale pour une prise de décision à partir d'une fréquence observée sur un échantillon. <u>STMG</u> **Échantillonnage et prise de décision** Intervalle de fluctuation d'une fréquence. Prise de décision.	<u>S-ES-L</u> Exploiter l'intervalle de fluctuation à un seuil donné, déterminé à l'aide de la loi binomiale, pour rejeter ou non une hypothèse sur une proportion. <u>STL-STI2D-STMG</u> Déterminer à l'aide de la loi binomiale un intervalle de fluctuation, à environ 95 %, d'une fréquence. Exploiter un tel intervalle pour rejeter ou non une hypothèse sur une proportion	L'intervalle de fluctuation peut être déterminé à l'aide d'un tableur ou d'un algorithme. Le vocabulaire des tests (test d'hypothèse, hypothèse nulle, risque de première espèce) est hors programme.

Extraits des programmes de 2010 pour les séries L-ES-S , 2011 pour les séries STL-STI2D et 2012 pour la série STMG.

Classe de terminale- Statistiques et probabilités

STL-STI2D
Compléter la problématique de la prise de décision par celle de l'estimation par intervalle de confiance. On s'appuie sur la loi normale et, en mathématiques, on se limite au cadre d'une proportion. Toutefois, la pertinence des méthodes statistiques utilisées dans les disciplines scientifiques et technologiques, en particulier l'estimation d'une moyenne, peut s'observer par simulation.

STMG
Consolider les connaissances acquises dans le domaine de l'échantillonnage et aborder l'estimation par la détermination d'un intervalle de confiance pour une proportion.

S-ES-L
La loi normale permet d'initier les élèves à la statistique inférentielle par la détermination d'un intervalle de confiance pour une proportion à un niveau de confiance de 95 %.
Cette partie se prête particulièrement à l'étude de problèmes issus d'autres disciplines

Le recours aux représentations graphiques et aux simulations est indispensable.

CONTENUS **Intervalle de fluctuation**	**CAPACITES ATTENDUES**	**COMMENTAIRES**
STL-STI2D Intervalle de fluctuation d'une fréquence	S-ES-L-STL-STI2D Connaître l'intervalle de fluctuation asymptotique à 95 % d'une fréquence obtenue sur un échantillon de taille *n* : $\left[p-1,96\sqrt{\frac{p(1-p)}{n}}; p+1,96\sqrt{\frac{p(1-p)}{n}}\right]$ lorsque la proportion *p* dans la population est connue. STL-STI2D Exploiter un tel intervalle de fluctuation pour rejeter ou non une hypothèse sur une proportion.	ES-L La variable aléatoire F_n qui, à tout échantillon de taille *n*, associe la fréquence, prend ses valeurs dans l'intervalle de fluctuation asymptotique (*) au seuil de 95 % avec une probabilité qui s'approche de 0,95 quand *n* devient grand. On admet le résultat ci-contre, qui est conforté grâce à la simulation. S La démonstration du théorème suivant « si la variable aléatoire X_n suit la loi binomiale de paramètres n et p avec $0 < p < 1$, alors pour tout ~ dans $]0;1[$ et pour $I_n = \left[p-u_{\sim}\sqrt{\frac{p(1-p)}{n}}; p+u_{\sim}\sqrt{\frac{p(1-p)}{n}}\right]$; $\lim\limits_{n\to+\infty} P\left(\frac{X_n}{n}\in I_n\right)=1-\sim$ » donne l'expression de l'intervalle de fluctuation asymptotique (*) au seuil de $1-\sim$ de la variable aléatoire fréquence $F_n=\frac{X_n}{n}$ qui, à tout échantillon de taille *n*, associe la fréquence obtenue *f*. S-ES-L-STL-STI2D Avec les exigences usuelles de précision, on pratique cette approximation dès que $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$.

		S-ES-L En majorant $1,96\sqrt{p(1-p)}$ on retrouve l'intervalle de fluctuation présenté en classe de seconde. La problématique de prise de décision, déjà rencontrée, est travaillée à nouveau avec l'intervalle de fluctuation asymptotique.
STMG **Échantillonnage et prise de décision** Intervalle de fluctuation d'une fréquence. Prise de décision.	Connaître un intervalle de fluctuation à au moins 95 % d'une fréquence d'un échantillon de taille *n* : $\left[p-\frac{1}{\sqrt{n}}\,;\,p+\frac{1}{\sqrt{n}}\right]$ lorsque la proportion *p* dans la population est connue. Exploiter un tel intervalle pour rejeter ou non une hypothèse sur une proportion.	On peut faire observer qu'en approchant la loi binomiale par la loi normale de même espérance et d'écart type $\sqrt{p(1-p)n}$, on est conduit à l'intervalle $\left[p-1,96\sqrt{\frac{p(1-p)}{n}}\,;\,p+1,96\sqrt{\frac{p(1-p)}{n}}\right]$ qui est inclus dans $\left[p-\frac{1}{\sqrt{n}}\,;\,p+\frac{1}{\sqrt{n}}\right]$. On retrouve l'intervalle de fluctuation présenté en classe de seconde. Le vocabulaire des tests (test d'hypothèse, hypothèse nulle, risque de première espèce) est hors programme.

(*)Avec les notations précédentes : Un intervalle de fluctuation asymptotique de la variable aléatoire F_n au seuil 0,95 (de $1-\alpha$ pour la série S) est un intervalle déterminé à partir de *p* et de *n* et qui contient F_n avec une probabilité d'autant plus proche de 0,95 (de $1-\alpha$ pour la série S) que *n* est grand.

Pour une valeur de *p* fixée, l'intervalle aléatoire $\left[F_n-\frac{1}{\sqrt{n}}\,;\,F_n+\frac{1}{\sqrt{n}}\right]$ contient, pour *n* assez grand, la proportion *p* à estimer avec une probabilité au moins égale à 0,95 (de $1-\alpha$ pour la série S)

CONTENUS **Intervalle de confiance**	**CAPACITES ATTENDUES**	**COMMENTAIRES**
STL-STI2D Intervalle de confiance d'une proportion**.**	- Estimer une proportion inconnue avec un niveau de confiance de 95 % par l'intervalle : $\left[f-1,96\sqrt{\frac{f(1-f)}{n}}\,;\,f+1,96\sqrt{\frac{f(1-f)}{n}}\right]$ calculé à partir d'une fréquence *f* obtenue sur un échantillon de taille *n*. - Juger de l'égalité de deux proportions à l'aide des intervalles de confiance à 95 % correspondant aux fréquences de deux échantillons de taille *n*.	Cette expression de l'intervalle de confiance, pour *n* assez grand, est admise. On constate par simulation que, pour *n* ≥ 30, sur un grand nombre d'intervalles de confiance, environ 95 % contiennent la proportion à estimer. La différence entre les deux fréquences observées est considérée comme significative quand les intervalles de confiance à 95 % sont disjoints. C'est l'occasion d'étudier des méthodes statistiques pratiquées dans les disciplines scientifiques ou technologiques. En liaison avec les enseignements technologiques et scientifiques, on peut observer par simulation la pertinence d'un intervalle de confiance de la moyenne d'une population, pour un caractère suivant une loi normale.

S-ES-L Intervalle de confiance au niveau de confiance 0,95(**). Niveau de confiance.	− Estimer une proportion inconnue à partir d'un échantillon. − Déterminer une taille d'échantillon suffisante pour obtenir, avec une précision donnée, une estimation d'une proportion au niveau de confiance 0,95. On énonce que *p* est élément de l'intervalle $\left[f-\frac{1}{\sqrt{n}}\,;\,f+\frac{1}{\sqrt{n}}\right]$ avec un niveau de confiance de plus de 95 %, où *f* désigne la fréquence observée sur un échantillon de taille *n*.	Les attendus de ce paragraphe sont modestes et sont à exploiter en lien avec les autres disciplines. Avec les exigences usuelles de précision, on utilise cet intervalle dès que $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$. La simulation de sondages sur tableur permet de sensibiliser aux fourchettes de sondage. Il est important de noter que, dans d'autres champs, on utilise l'intervalle $\left[f-1{,}96\sqrt{\frac{f(1-f)}{n}}\,;\,f+1{,}96\sqrt{\frac{f(1-f)}{n}}\right]$ qu'il n'est pas possible de justifier dans ce programme. S Il est intéressant de démontrer que, pour une valeur de *p* fixée, l'intervalle $\left[F_n-\frac{1}{\sqrt{n}}\,;\,F_n+\frac{1}{\sqrt{n}}\right]$ contient, pour *n* assez grand, la proportion *p* avec une probabilité au moins égale à 0,95. On énonce alors que *p* est élément de l'intervalle $\left[f-\frac{1}{\sqrt{n}}\,;\,f+\frac{1}{\sqrt{n}}\right]$ avec un niveau de confiance de plus de 95 %, où *f* désigne la fréquence observée sur un échantillon de taille *n*.
STMG **Estimation** Intervalle de confiance d'une proportion	Estimer une proportion inconnue par l'intervalle $\left[f-\frac{1}{\sqrt{n}}\,;\,f+\frac{1}{\sqrt{n}}\right]$ où *f* est la fréquence obtenue sur un échantillon de taille *n*.	Cet intervalle contient la proportion dans au moins 95 % des cas pour *n* grand, ce qui peut être illustré par simulation. La notion de niveau de confiance ne fait pas l'objet de développements

(**) Un intervalle de confiance pour une proportion *p* au niveau de confiance 0,95 (de $1-\alpha$ pour la série S) est la réalisation, à partir d'un échantillon, d'un intervalle aléatoire contenant la proportion *p* avec une probabilité supérieure ou égale à 0,95 (de $1-\alpha$ pour la série S), intervalle aléatoire déterminé à partir de la variable aléatoire F_n qui, à tout échantillon de taille *n*, associe la fréquence.
Les intervalles de confiance considérés ici sont centrés en la fréquence observée *f*.

Extraits des programmes de 2011 pour les séries L-ES-S , 2011 pour les séries STL-STI2D et 2012 pour la série STMG.