L'ECHANTILLONNAGE.

La théorie de l'échantillonnage consiste, connaissant des propriétés d'une population, à déterminer des propriétés d'échantillons prélévés dans la population.
En réalité, on est plus souvent confronté au problème inverse, celui de l'estimation: déduire des informations sur la population totale à partir de renseignements donnés par des échantillons (voir chapitre Estimation). Mais les observations sur l'échantillonnage permettent d'obtenir des résultats utiles pour l'estimation.
Le tirage des éléments d'un échantillon aléatoire peut être exhaustif (sans remise); dans ce cas la composition de l'urne est modifiée à chaque tirage: les tirages ne sont donc pas indépendants.
Sinon le tirage est non exhaustif (avec remise); dans ce cas la composition de l'urne n'est pas modifiée à chaque tirage: les tirages sont donc indépendants.
Dans la plupart des cas où la population a un grand effectif N dont on tire un petit nombre d'éléments n (en pratique si $N \geq 10\,n$) on assimile un tirage sans remise à un tirage avec remise.

Théorème de la limite centrée

Soit X_1, X_2, …, X_n, n variables aléatoires indépendantes, suivant toutes la même loi, admettant une moyenne μ et une variance σ^2 ($\sigma > 0$).

Pour n suffisamment grand, la variable aléatoire $\overline{X_n} = \dfrac{X_1 + ... + X_n}{n}$ suit approximativement la loi normale $N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$

Distribution d'échantillonnage asymptotique de la moyenne.

Soit une population connue, d'effectif total N. On étudie un caractère pour lequel la moyenne de la population est m, l'écart-type σ.
Considérons tous les échantillons de taille n issus de la population.
E $_1$: effectif n, moyenne $\overline{x_1}$, écart-type σ'_1 ... E $_k$: effectif n, moyenne $\overline{x_k}$, écart-type σ'_k.

L'ensemble $\overline{X} = \{ \overline{x_1}, ..., \overline{x_k} \}$ est une série statistique appelée distribution des moyennes.

On montre que cette série statistique a pour moyenne m et pour écart-type $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}}$ si l'échantillonnage est non exhaustif, ($\dfrac{\sigma}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}}$ si l'échantillonnage est exhaustif.)

Pour n suffisamment grand ($n \geq 30$) on considère que la variable aléatoire $\overline{X}$ qui, dans le cas d'un tirage non exhaustif, à tout échantillon associe sa moyenne suit la loi normale $N(\mu, \sigma_{\overline{X}})$
avec $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}}$, σ étant l'écart-type de la population et n l'effectif des échantillons.
(et $N\left(m, \dfrac{\sigma}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}}\right)$ dans le cas d'un tirage exhaustif).

Distribution d'échantillonnage asymptotique de la fréquence.

Considérons une population d'effectif N dont un pourcentage p d'éléments possède une certaine propriété. De la même manière que précédemment on prélève des échantillons de taille n et on mesure pour chacun le pourcentage f des éléments possédant cette propriété.

E_1 : effectif n , pourcentage f_1 ... E_k : effectif n , pourcentage f_k

L'ensemble $F = \{ f_1, f_2, \dots f_k \}$ est une série statistique appelée distribution des pourcentages.

On montre que cette série statistique a pour moyenne p et pour écart-type $\sqrt{\frac{pq}{n}}$ (avec $q = 1 - p$) si l'échantillonnage est non exhaustif, ($\sqrt{\frac{pq}{n}} \sqrt{\frac{N-n}{N-1}}$ si l'échantillonnage est exhaustif).

> Pour n suffisamment grand ($n \geq 30$) on considère que la variable aléatoire F qui à tout échantillon associe le pourcentage des éléments possédant la propriété suit la loi $\mathcal{N}\left(p , \sqrt{\frac{pq}{n}} \right)$ dans le cas d'un tirage non exhaustif et $\mathcal{N}\left(p , \sqrt{\frac{pq}{n}} \sqrt{\frac{N-n}{N-1}} \right)$ dans le cas d'un tirage exhaustif.

EXERCICE 1

Les notes à l'épreuve de maths d'un examen ont pour moyenne 12 et pour écart-type 3.
Soit $\overline{X}$ la variable aléatoire qui, à tout échantillon aléatoire et avec remise de taille 100, associe la moyenne des notes de cet échantillon.
On suppose que $\overline{X}$ suit la loi normale $\mathcal{N}$ (12 ; 0,3)
On se propose de prélever un échantillon aléatoire avec remise de 100 notes.
Quelle est la probabilité d'avoir la moyenne d'un tel échantillon supérieure à 12,5 ?
Quelle est la probabilité d'avoir la moyenne d'un tel échantillon comprise entre 12,5 et 12,9 ?

EXERCICE 2

Dans une population, on constate qu'il naît 52% de garçons et 48% de filles.
On suppose que la variable aléatoire F qui, à tout échantillon de taille $n = 400$ prélevé au hasard et avec remise dans la population, associe le pourcentage de garçons dans cet échantillon, suit la loi normale $\mathcal{N}\left(p , \sqrt{\frac{p(1-p)}{n}} \right)$, avec $p = 0{,}52$.
On se propose de prélever un échantillon aléatoire avec remise de 400 nouveau-nés.
Déterminer la probabilité d'avoir, dans un tel échantillon, un pourcentage de garçons compris entre 50 % et 54 %.
Déterminer la probabilité d'avoir, dans un tel échantillon, un pourcentage de filles inférieur à 45 % .

EXERCICE 3

La taille moyenne des 180 élèves de 6^e d'un collège est de 135 cm, avec un écart-type de 7 cm. On note la taille moyenne d'un groupe de 32 élèves. Déterminer la probabilité :
pour que celle-ci soit comprise entre 134 et 136 cm
pour qu'elle soit supérieure à 137 cm
pour qu'elle soit inférieure à 135,8 cm.

EXERCICE 4

Un service d'étude sur des ventes par correspondance enregistre en moyenne 20 % de commandes. Quelle est la loi normale susceptible de rendre compte du pourcentage de commandes ?
Pour un nombre d'appels égal à 500, quelle est la probabilité d'observer plus de 115 commandes ? moins de 90 ?

EXERCICE 5

Une machine fabrique des pièces de forme circulaire en grande série.
A chaque pièce tirée au hasard on associe son diamètre en mm. On définit ainsi une variable aléatoire X.
On suppose que X suit la loi normale N (150 ; 0,21)
Soit M la variable aléatoire qui à chaque échantillon de 400 pièces prélevé au hasard avec remise associe la moyenne des diamètres des pièces de cet échantillon.
M est une variable aléatoire qui suit une loi normale.
Déterminer le nombre positif h tel que: $p(\mu - h \leq M \leq \mu + h) = 0{,}95$. (μ : diamètre moyen des pièces de la population)

EXERCICE 6

Une machine fabrique des pièces en grande série.
A chaque pièce tirée au hasard on associe sa longueur en mm. On définit ainsi une variable aléatoire X.
On suppose que X suit la loi normale N (28,20 ; 0,027)
On admet que la variable aléatoire M, qui à tout échantillon aléatoire non exhaustif de taille n associe la moyenne des longueurs des n pièces de cet échantillon, suit la loi normale $N\left(m, \frac{\sigma}{\sqrt{n}}\right)$ avec $m = 28{,}20$ et $\sigma = 0{,}027$.
Déterminer n pour que $p(28{,}195 \leq M \leq 28{,}205) = 0{,}95$.