

Chapitre 16

Échantillonnage

I Exercices

16.1 Échantillonnage et fluctuation d'échantillonnage

Lire d'abord le cours du paragraphe 16.1 page 180, jusqu'à l'exemple 16.1, et le paragraphe « Objectif de la première partie du cours ».

Exercice 16.1 (Simulation et échantillonnage)

Un grand supermarché a deux entrées A et B, la probabilité qu'un client entre par l'entrée A est 40 % soit 0,4.

1. Dans le tableau ci-dessous, simuler le choix d'un échantillon de dix clients qui arrivent au supermarché.
Voir l'exemple 16.2 et la méthode 16.1 page 181.
2. Calculer la fréquence de l'événement A.
3. Renouveler deux séries de 10 simulations sans compléter de tableau et calculer chaque fois la fréquence obtenue.
4. La fréquence obtenue n'est pas toujours égale à la probabilité 0,4. Comment appelle-t-on ce phénomène?
Voir le cours de l'exemple 16.3 page 182 jusqu'à la définition 16.1.

Nombre aléatoire entre 1 et 10										
L'événement s'est-il produit ?										

Avant de faire l'exercice suivant, lire dans le cours, l'exemple 16.4 page 183, et la remarque 16.2.

Exercice 16.2 (Échantillon de grande taille)

Dans cet exercice on a besoin d'un script Python nommé `echantillonnage2de` que l'on peut envoyer sur sa calculatrice Numworks ou que l'on peut copier sur son ordinateur, et qui est accessible en suivant le lien ci-dessous.

<https://my.numworks.com/python/jean-luc-poncin/echantillonnage2de>

Suite de l'énoncé page suivante

Dans la fonction Python `frsucces`, les variables `n`, `compteur`, `i` sont des entiers naturels.

Nous allons utiliser cette fonction, en la modifiant, pour revenir sur la situation de l'exercice 16.1 : simuler un échantillon de plusieurs clients arrivant au supermarché et qui décident ou non de passer par l'entrée A, puis augmenter la taille de l'échantillon.

```

1 from random import *
2 from math import *
3
4 def frsucces(n):
5     compteur=0
6     for i in range(n):
7         if(randint(1,10)<=...):
8             compteur=compteur+1
9     return(compteur/n)

```

1. Compléter dans les pointillés ci-dessus pour que la fonction `frsucces` simule la répétition de l'arrivée au supermarché de plusieurs clients et l'événement «entrer par l'entrée A».
2. Exécuter la fonction `frsucces` pour $n = 10$, puis $n = 100$, ainsi de suite et compléter le tableau ci-dessous.
3. Dans ce tableau, quand n devient grand, `frsucces(n)` se rapproche d'un nombre. Lequel? Expliquer ce que ça signifie.
4. Comment appelle-t-on cette propriété?

n	10	100	1 000	10 000	100 000
<code>frsucces(n)</code>					

Lire maintenant dans le cours tout le paragraphe 16.1.d page 184, de l'exemple 16.5 jusqu'à l'exemple 16.6 puis faire l'exercice ci-dessous.

Exercice 16.3 (Simuler un grand nombre d'échantillons)

On continue à utiliser la situation de l'exercice 16.1 (des clients arrivant au supermarché et qui décident ou non de passer par l'entrée A), mais nous allons maintenant simuler 100 échantillons de 50 clients.

Dans l'exemple 16.5 page 184, ce type de simulation est faite avec un tableur (LibreOffice ou Excel), mais ici, nous allons le faire avec la fonction Python `plusech` (*plusieursechantillons*) qui se trouve dans le script `echantillonnage2de`.

Ce script est entièrement reproduit ci-contre.

Dans les fonctions Python `frsucces`, et `plusech`, les variables `n`, `compteur`, `i`, `nbechant`, `j`, sont des entiers naturels.

Pour la ligne 14, on précise que `abs` veut dire *valeur absolue*, donc cette ligne signifie :

$$\text{si } |0,3 - f| \leq \frac{1}{\sqrt{30}}$$

Répondre aux questions page suivante.

```

1 from random import *
2 from math import *
3
4 def frsucces(n):
5     compteur=0
6     for i in range(n):
7         if(randint(1,10)<=3):
8             compteur=compteur+1
9     return(compteur/n)
10
11 def plusech(nbechant):
12     compteur=0
13     for j in range(nbechant):
14         if(abs(0.3-frsucces(30))<=1/sqrt(30)):
15             compteur=compteur+1
16     return(compteur/nbechant)

```

1. Les nombres encadrés sont adaptés pour l'exemple 16.5 page 184.
Modifier ces nombres pour que la fonction `plusech` simule 100 échantillons de 50 clients, sachant que la probabilité qu'un client passe par l'entrée A est $p = 0,4$.
2. Exécuter à la console la fonction `plusech` pour qu'elle simule 100 échantillons. On obtient normalement un nombre décimal entre 0,95 et 1.
3. Que représente ce nombre pour la simulation des 100 échantillons de 50 clients ?
4. Quelle propriété affirme que le résultat obtenu en 2. est normalement un nombre décimal entre 0,95 et 1 ?

Ces trois premiers exercices ont permis d'aborder l'idée de l'échantillonnage, et les capacités attendues du paragraphe 16.0 page 180. Les exercices suivants ont pour but de s'entraîner sur ces capacités attendues.

Exercice 16.4

Dans la population d'un pays la proportion de femmes et d'hommes est équilibrée : 50 % de femmes et 50 % d'hommes. On appelle p la proportion de femmes, et on a donc : $p = 0,5$.

On choisit au hasard 500 personnes pour constituer une assemblée que le gouvernement veut consulter sur certains sujets.

1. Cette situation est une situation d'échantillonnage.
 - a) Quelle est la population concernée ?
 - b) Quel est le caractère étudié et quel sa proportion p ?
 - c) Quel est l'échantillon et quel est sa taille ?
2. Un organisateur veut simuler le choix au hasard des 500 personnes, pour vérifier la fréquence de femmes dans l'échantillon, et il a saisi la fonction Python ci-dessous qui se trouve dans le script `echantillonnage2de`.

Dans la fonction `simul`, les variables `n`, `c`, `i` sont des entiers naturels.

- a) Le nombre retourné par la fonction `simul` n'est pas la fréquence simulée de femmes dans l'échantillon. Que représente ce nombre ?
- b) Exécuter ce script à la console pour simuler le choix au hasard des 500 personnes.
- c) Indiquer le résultat obtenu, puis calculer la fréquence f de femmes que cela représente. Arrondir au centième près.

```
from random import *
from math import *

def simul(n):
    c=0
    for i in range(n):
        if(randint(1,2)==1):
            c=c+1
    return(c)
```

3. Le jour où les 500 personnes sont choisies au hasard, on constate qu'il y a 265 femmes. Peut-on dire qu'il y avait 95 % de chances que cela se produise ? Avant de répondre, on calculera l'intervalle de fluctuation. Arrondir au centième près.

Exercice 16.5

La fonction Python ci-contre est dans le script `echantillonnage2de`.

Dans la fonction `jeu`, les variables `n`, `c`, `i` sont des entiers naturels.

```
from random import *
from math import *

def jeu(n):
    c=0
    for i in range(n):
        if(randint(1,6)==1):
            c=c+1
    return(c)
```

1. Cette fonction `jeu` simule un jeu bien connu. Lequel ?
2. Dans la fonction `jeu`, quel événement est simulé, et quelle est sa probabilité p ? Donner la probabilité p sous forme de fraction, puis arrondir au millième près.
3. Si on effectue 600 répétitions, on s'attend à quelle nombre de succès ?
4. Simuler 600 répétitions plusieurs fois en exécutant la fonction `jeu` à la console.
5. On n'obtient pas toujours le même nombre de succès. Comment nomme-t-on ce phénomène dans le cours ?
6. Quand on joue à ce jeu plusieurs fois, on considère que si la fréquence de succès n'est pas dans l'intervalle de fluctuation, il a 95 % pour que l'objet utilisé ne soit pas bien équilibré. Quelqu'un joue à ce jeu, le répète 60 fois et obtient 20 succès. Peut-on estimer, avec 95 % de chances d'avoir raison que l'objet utilisé ne soit pas bien équilibré ? Détailler d'abord les calculs avant de répondre. Arrondir au centième près.

16.2 Estimation d'une probabilité ou d'une proportion dans une population

Avant de faire les exercices suivants, lire le paragraphe 16.2 page 185, la propriété 16.4, et l'exemple 16.7.

Exercice 16.6

Dans un champ de céréale, on veut estimer quelle est la proportion p d'épis bien fécondés. Pour cela, on prélève au hasard un échantillon de 300 épis, et on constate que 171 sont bien fécondés.

Donner une estimation de la proportion p d'épis bien fécondés en donnant un intervalle. Détailler les calculs. Arrondir au centième près.

Exercice 16.7

1. Un candidat à une élection veut connaître la proportion p d'électeurs qui souhaitent voter pour lui, et fait appel à un institut de sondage. On interroge 1 000 personnes, on leur demande si elles voteront pour ce candidat, et 521 personnes répondent oui.
 - a) Donner une estimation de la proportion p d'électeurs qui souhaitent voter pour ce candidat en donnant un intervalle. Détailler les calculs. Arrondir au centième près.
 - b) Si toutes les valeurs de cet intervalle sont supérieures à 50 %, ce candidat a 95 % de chances d'être élu, mais est-ce bien le cas ici ?

2. L'élection a lieu, les premiers bureaux de vote ferment à 18 h, les derniers à 20 h. Entre 18 h et 20 h l'institut de sondage prend note de 10 000 résultats : il y a 5 190 votes favorables à ce candidat.
- a) Donner une estimation de la proportion p' d'électeurs qui ont votés pour ce candidat en donnant un intervalle. Détailler les calculs.
 - b) Ce candidat a-t-il 95 % de chances d'être élu ?
3. Pourquoi la 2^e estimation est-elle plus précise que la première ?

II Cours

16.0 Programme

Contenus

- Échantillon aléatoire de taille n pour une expérience à deux issues.
- Principe de l'estimation d'une probabilité, ou d'une proportion dans une population, par une fréquence observée sur un échantillon.
- Version vulgarisée de la loi des grands nombres : « Lorsque n est grand, sauf exception, la fréquence observée est proche de la probabilité. »

Capacités attendues

- Lire et comprendre une fonction Python renvoyant le nombre ou la fréquence de succès dans un échantillon de taille n pour une expérience aléatoire à deux issues.
- Observer la loi des grands nombres à l'aide d'une simulation sur Python ou tableur.
- Simuler N échantillons de taille n d'une expérience aléatoire à deux issues. Si p est la probabilité d'une issue et f sa fréquence observée dans un échantillon, calculer la proportion des cas où l'écart entre p et f est inférieur ou égal à $\frac{1}{\sqrt{n}}$.

16.1 Échantillonnage et fluctuation d'échantillonnage

Dans cette première partie du cours, on s'intéresse à un certain caractère dans une population donnée dont la proportion p est connue.

Cette proportion et ce caractère peuvent être par exemple :

- la proportion de tickets gagnants dans l'ensemble des billets qui ont été imprimés pour une tombola ;
- la proportion de pièces défectueuses dans l'ensemble des pièces produite par une usine en un mois ;
- la proportion de personnes qui utilisent les transports en commun dans la population d'une ville ;
- la proportion de clients achetant des fruits et légumes dans l'ensemble des clients d'une grande surface pendant un an.

La situation de l'échantillonnage est de considérer un groupe de n individus dans cette population, qu'on appelle un échantillon de taille n et de vérifier la fréquence f du caractère dans cet échantillon.

En reprenant les exemples précédents, cet échantillon et cette fréquence peuvent être :

- la fréquence de tickets gagnants parmi 100 billets de tombola pris au hasard ;
- la fréquence de pièces défectueuses parmi 1 000 pièces choisies au hasard dans la production d'une usine ;
- la fréquence de personnes qui utilisent les transports en commun parmi 500 personnes interrogées au hasard dans la population d'une ville ;
- la fréquence de clients achetant des fruits et légumes parmi 200 clients choisis au hasard parmi les clients d'une grande surface.

Revenons au caractère dans une population donnée dont la proportion p est connue.

Si l'on choisit un individu au hasard dans cette population, la probabilité que cet individu ait ce caractère est égale à la proportion p de ce caractère dans cette population. Voir l'exemple 16.1 page suivante.

Exemple 16.1

On sait que dans la population française la proportion de personnes atteintes par une certaine maladie est $p = 30 \% = 0,3$.

On choisit une personne au hasard dans cette population.

On appelle M l'événement : « la personne choisie au hasard est malade de cette maladie ».

La probabilité de M est donc : $p(M) = p = 0,3$.

Objectif de la première partie du cours

Le but de cette première partie du cours est d'expliquer comment simuler avec une calculatrice ou avec Python les situations ci-dessous.

1. Choisir une personne au hasard, et vérifier si elle est malade ou non, sachant que la probabilité qu'elle soit malade est $p = 0,3$.
2. Choisir un échantillon de n personnes au hasard, vérifier si elles sont malades ou non, et calculer la fréquence f de personnes malades.
3. Étudier ensuite comment évolue cette fréquence f par rapport à la probabilité p quand le nombre n de personnes de l'échantillon devient grand.
4. Étudier un grand nombre d'échantillons de n personnes.

16.1.a Simuler une expérience aléatoire et un événement de probabilité p **Exemple 16.2**

Avec une calculatrice, ou Python, ou dans un tableur on peut afficher des nombres au hasard, on appelle ça des nombres aléatoires.

On veut par exemple simuler l'expérience aléatoire de choisir une personne au hasard, et vérifier si l'événement M de probabilité $p = 0,3$ est réalisé ou pas. On fait afficher alors des nombres entiers aléatoires entre 1 et 10.

- Si le nombre affiché est inférieur ou égal à 3, on considère que l'événement s'est produit, en effet la probabilité que le nombre affiché soit 1 ou 2 ou 3 est $\frac{3}{10} = 0,3$;
- sinon, on considère qu'il ne s'est pas produit.

Quand l'événement se produit, on appelle ça *succès*, et quand il ne se produit pas, et on appelle ça *échec*.

Méthode 16.1 (Afficher un nombre entier aléatoire entre 1 et 10)**Numworks**

- module Calculs
- touche Boîte à outils
- descendre sur Aléatoire et approximation ▷
- descendre sur `randint(a,b)` et valider
- compléter ainsi : `randint(1,10)`

TI 82 : $\boxed{\text{math}}$ $\boxed{\leftarrow}$ $\boxed{5}$, et compléter ainsi : `entAléat(1,10)`

TI 83 :

- $\boxed{\text{math}}$ $\boxed{\rightarrow}$ $\boxed{\rightarrow}$ $\boxed{\rightarrow}$ $\boxed{5}$
- compléter comme ci-dessous, puis valider :
`bornin:1`
`bornsup:10`
`n:`

- on voit : `nbrAléatEnt(1,10)`
- valider

CASIO : OPTN F6 (\triangleright) F2 (PROB) F4 (RAND) F2 (Int)

et compléter ainsi : `RandInt#(1,10)`

En Python

Au début du script, saisir ceci : `from random import *`

L'instruction pour obtenir un nombre entier aléatoire entre 1 et 10 est : `randint(1,10)`

Dans le tableur de LibreOffice : `=ALEA.ENTRE.BORNES(1;10)`

16.1.b Simuler la répétition d'une expérience aléatoire – Fluctuation d'échantillonnage

Exemple 16.3

On veut maintenant simuler n fois l'expérience aléatoire de choisir une personne au hasard, et de vérifier si l'événement M de probabilité $p = 0,3$ est réalisé ou pas.

Cette simulation revient à choisir au hasard un échantillon de n personnes, à vérifier si elles sont malades ou non, et calculer la fréquence observée f de personnes malades.

Effectuons cette simulation pour $n = 10$ en répétant 10 fois ce qui est expliqué dans l'exemple 16.2.

On obtient alors le tableau ci-dessous.

Nombre aléatoire entre 1 et 10	7	7	5	5	6	6	1	7	3	8
Événement M , succès ou échec	échec	échec	échec	échec	échec	échec	succès	échec	succès	échec

Dans cette simulation, l'événement M s'est produit 2 fois (2 succès), la fréquence observée de malades dans cet échantillon est : $f = \frac{2}{10} = 0,2$.

Recommençons deux fois, on obtient alors les tableaux ci-dessous.

Nombre aléatoire entre 1 et 10	3	8	2	1	8	5	4	8	5	6
Événement M , succès ou échec	succès	échec	succès	succès	échec	échec	échec	échec	échec	échec

L'événement M s'est produit 3 fois (3 succès), la fréquence observée de malades dans cet échantillon est : $f = \frac{3}{10} = 0,3$

Nombre aléatoire entre 1 et 10	8	4	3	2	5	8	3	2	3	10
Événement M , succès ou échec	échec	échec	succès	succès	échec	échec	succès	succès	succès	échec

L'événement M s'est produit 5 fois (5 succès), la fréquence observée de malades dans cet échantillon est : $f = \frac{5}{10} = 0,5$.

Remarque 16.1

On remarque que les trois fréquences observées ne sont pas toujours égales à 0,3.

Propriété 16.1 (Fréquence observée et probabilité)

On considère une expérience aléatoire et un événement de probabilité p . On simule un échantillon de taille n en simulant n fois cette expérience aléatoire. On constate que la fréquence observée f de succès de cet événement n'est pas toujours égale à sa probabilité p .

Définition 16.1 (Fluctuation d'échantillonnage)

On considère une expérience aléatoire et un événement de probabilité p . On simule un échantillon de taille n en simulant n fois cette expérience aléatoire. Le fait que la fréquence observée f n'est pas toujours égale à la probabilité p est nommé fluctuation d'échantillonnage.

16.1.c Échantillon de grande taille – Loi des grands nombres

Exemple 16.4

Simulons maintenant des échantillons dont la taille n est de plus en plus grande.

Nous allons utiliser un script Python nommé `echantillonnage2de` que l'on peut envoyer sur sa calculatrice Numworks ou que l'on peut copier sur son ordinateur, et qui est accessible en suivant le lien ci-dessous.

<https://my.numworks.com/python/jean-luc-poncin/echantillonnage2de>

Ce script contient la fonction Python ci-dessous, nommée `frsucces`, qui permet de simuler n fois l'expérience aléatoire de choisir une personne au hasard et vérifier si elle est malade ou pas, de compter le nombre de personnes malades (nombre de succès), puis de calculer la fréquence de succès.

Les variables n , `compteur`, i sont des entiers naturels.

	Fonction Python	Explications
1	<code>def frsucces(n):</code>	On définit la fonction <code>frsucces</code> de variable n
2	<code>compteur=0</code>	Au début le compteur vaut zéro
3	<code>for i in range(n):</code>	On répète n fois les lignes 4 et 5.
4	<code>if(randint(1,10)<=3):</code>	<i>Si le nombre aléatoire entre 1 et 10 est inférieur ou égal à 3</i>
5	<code>compteur=compteur+1</code>	<i>la variable compteur augmente de 1</i>
6	<code>return(compteur/n)</code>	À la fin on retourne la valeur de <code>compteur/n</code>

On exécute la fonction `frsucces` pour des valeurs de n de plus en plus grande et on obtient par exemple le tableau ci-dessous.

n	10	100	1 000	10 000	100 000
<code>frsucces(n)</code>	0,2	0,26	0,294	0,3043	0,29995

Remarque 16.2

On constate que la fréquence de succès est de plus en plus proche de 0,3 c'est à dire que lorsque la taille de l'échantillon devient grande, la fréquence de malade f est proche de la probabilité $p = 0,3$.

Propriété 16.2 (Loi des grands nombres)

Lorsque la taille n de l'échantillon devient grande, sauf exception, la fréquence observée est proche de la probabilité.

Remarque 16.3

Dans la loi des grands nombres formulée ci-dessus, il convient de préciser deux choses.

- La loi des grands nombres indique que, lorsque la taille n de l'échantillon devient grande, la fréquence observée est proche de la probabilité. Oui, mais proche comment ? Avec quel écart ?
- La loi des grands nombres indique que, lorsque la taille n de l'échantillon devient grande, la fréquence observée est proche de la probabilité, en précisant *sauf exception*. Oui, mais combien y a-t-il d'exception ? ou plutôt, quelle est la proportion d'exception ?

Pour répondre à ces deux questions, il faut maintenant simuler un grand nombre d'échantillons.

16.1.d Simuler un grand nombre d'échantillons

Exemple 16.5

On continue à prendre comme exemple la situation de la maladie dont 30 % de la population est atteinte, en rappelant qu'ainsi la probabilité qu'une personne choisie au hasard soit malade est $p = 0,3$.

Dans le fichier annexe nommé `2de_echantillonnage_exemple_du_cours.pdf` on voit deux tableaux, créés avec un tableur, dans lesquels sont simulés 100 échantillons de 30 personnes choisies au hasard ($n = 30$).

Dans ces tableaux, on observe l'écart entre 0,3 et la fréquence f de malades, et on constate que l'échantillon où cet écart est le plus important est l'échantillon E42.

La propriété, la remarque et la définition ci-dessous vont maintenant donner des précisions sur l'écart entre la fréquence observée et la probabilité.

Propriété 16.3 (admise)

On répète n fois une expérience aléatoire avec un événement de probabilité p .

On appelle f la fréquence observée de succès de cet événement.

Alors, dans au moins 95 % des cas, l'écart entre f et p est inférieure ou égale à $\frac{1}{\sqrt{n}}$, autrement dit :

dans au moins 95 % des cas, $|f - p| \leq \frac{1}{\sqrt{n}}$.

Remarque 16.4

Dans la remarque 16.1 page 182, on indique que la fréquence observée fluctue autour de la probabilité p . La propriété ci-dessus précise en fait que dans 95 % des cas, la fréquence fluctue dans l'intervalle $\left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}} \right]$.

Définition 16.2 (Intervalle de fluctuation)

On répète n fois une expérience aléatoire avec un événement de probabilité p . On simule ainsi un échantillon de taille n . On appelle f la fréquence observée de succès de cet événement.

L'intervalle $\left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}} \right]$ est alors nommé *intervalle de fluctuation*.

Exemple 16.6 (Vérification de l'écart entre fréquence et probabilité)

Vérifions tout cela dans les tableaux du fichier annexe `2de_echantillonnage_exemple_du_cours.pdf`.

Ce sont des échantillons de taille $n = 30$, calculons donc $\frac{1}{\sqrt{n}} : \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{30}} \approx 0,18$

Les seuls échantillons où l'écart entre f et p est supérieur à 0,18 sont les échantillons E42, E77 et E87.

Échantillon E42 : $|f - p| = |0,07 - 0,3| = |-0,23| = 0,23 > 0,18$

Échantillon E42 et E87 : $|f - p| = |0,5 - 0,3| = 0,2 > 0,18$

La proportion d'échantillons où la distance entre f et p est supérieure à 0,18 est donc $\frac{3}{100} = 3\%$, donc la proportion d'échantillons où l'écart entre f et p est inférieure ou égale à 0,18 est 97 %, on a donc bien au moins 95 % des échantillons de taille n où l'écart entre la fréquence observée f et la probabilité p est inférieure ou égale à $\frac{1}{\sqrt{n}}$.

Pour obtenir d'autres exemples possibles de tableaux de simulations de 100 échantillons, on pourra utiliser un des deux fichiers annexes suivants :

- le fichier `2de_echantillonnage_fichier_de_test.ods` au format LibreOffice,
- le fichier `2de_echantillonnage_fichier_de_test.xlsx` au format Excel.

16.2 Estimation d'une probabilité ou d'une proportion dans une population

Dans cette deuxième partie du cours, on s'intéresse toujours à un certain caractère dans une population donnée, mais cette fois-ci lorsque la proportion p est inconnue.

Si l'on choisit une individu de la population au hasard, et qu'on appelle A l'événement « l'individu choisi au hasard a ce caractère », alors la proportion p est aussi la probabilité de A .

L'objectif est donc d'estimer une probabilité ou d'estimer une proportion dans une population.

D'après la première partie du cours, nous savons que lorsque la taille n d'un échantillon de la population devient grande, sauf exception, la fréquence observée est proche de la probabilité. Plus précisément nous savons que dans 95 % des cas l'écart entre la probabilité p et la fréquence f est inférieure ou égale à $\frac{1}{\sqrt{n}}$. On a donc la propriété ci-dessous.

Propriété 16.4

La proportion d'un caractère dans une population est p , et, dans un échantillon de taille n , la fréquence observée est f .

Alors, dans 95 % des cas, on peut estimer, que la proportion p appartient à l'intervalle

$$\left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}} \right].$$

Cela signifie aussi, pour un individu pris au hasard dans cette population, que, dans 95 % des cas, la probabilité p qu'il ait ce caractère appartient à cet intervalle.

Exemple 16.7

Une ville a imprimé 10 000 tickets de tombola. Une association en achète 80 et il y a 16 billets gagnants.

La population est l'ensemble des 10 000 tickets de tombola. Le caractère étudié est « billet gagnant ou non ».

Les 80 tickets de l'association constituent un échantillon de taille $n = 80$.

Pour estimer la proportion p de tickets gagnants dans les tickets imprimés par la ville, calculons f , puis $\frac{1}{\sqrt{n}}$, puis $f - \frac{1}{\sqrt{n}}$ et $f + \frac{1}{\sqrt{n}}$.

$$f = \frac{16}{80} = 0,2 \quad \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{80}} \approx 0,11$$

$$f - \frac{1}{\sqrt{n}} \approx 0,2 - 0,11 = 0,09 \quad f + \frac{1}{\sqrt{n}} \approx 0,2 + 0,11 = 0,31.$$

On peut donc estimer, avec 95 % de chances d'avoir raison, que la proportion p de tickets gagnants dans les tickets imprimés par la ville, est dans l'intervalle $[0,09 ; 0,31]$.

Cela signifie aussi qu'on peut estimer, avec 95 % de chances d'avoir raison, que la probabilité de gain à cette tombola est comprise entre 0,09 et 0,31.