T.D-2 :Chapitre-2

Mohamed Dakkon
F.S.J.E.S de Tétouan ,
Année universitaire 2020-2021.

Module:Echantillonnage et Estimation

Exercice1 :Solution

D'après l'énoncé de l'exercice, on constate que la population des étudiants de cette Faculté est composée de plusieurs strats. Donc afin d'améliorer la représentativité de l'échantillon, on doit adopter la méthode de stratification proportionnelle, qui consiste à répartir les 1000 étudiants proportionnellement au nombre total des étudiants dans chaques strate. Autrement dit, la construction de l'échantillon se fera conformément au poids de chaque strate comme suit :

$$n_h = \left(\frac{n}{N}\right) \times N_h$$

Avec :

h : désigne le rang du h^{eme} strate ;

n_h : est la taille de l'échantillon dans la strate h ;

N_h : est l'effectif de la population dans la strate h ;

n : est la taille de l'échantillon ;

N : est la taille de la population.

Exercice1 :Solution

Niveau / Sexe	S_1	S_2	S_3	S_4	S_5	S_6
Masculin	200	160	100	90	40	25
Féminin	150	90	60	50	20	15

- Exemple : $n_{11} = (\frac{1000}{10000}) \times 2000 = 200$
- La stratification consiste à découper la population étudiée en groupes homogènes, appelés strates, et à tirer indépondamment un échantillon aléatoire dans chaque strate.

Exercice 2 :Solution

1. Puisqu'on connait approximativement le nombre total de salariés dans chaque strate et les variables étudiées étant en bonne corrélation avec le nombre de salariés, l'échantillon doit etre réparti proportionnellement au nombre total de salariés dans chaque strate. Autrement dit, la méthode qu'on va adopter est celle de stratification proportionnelle.
La construction de l'échantillon dans le cadre de cette méthode se fera conformément au poids de chaque strate comme suit :

Exercice 2 :Solution

$$n_h = (\frac{n}{N}) \times N_h$$

Avec :
h : désigne le rang du h^{eme} strate ;
n_h : est la taille de l'échantillon dans la strate h ;
N_h : est l'effectif de la population dans la strate h ;
n : est la taille de l'échantillon ;
N : est la taille de la population ;
D'ou la structure de l'échantillon donnée dans le tableau ci-aprés :

Exercice 2 :Solution

Strates	Effectif de l'échantillon n_h
Strate 1	188
Strate 2	112
Totaux	300

Exemple :

$n_1 = (\frac{300}{800000}) \times 500000 = 188$

$n_2 = (\frac{300}{800000}) \times 300000 = 112$

Exercice 2 :Solution

2. Afin de procéder à la répartition optimale de l'échantillon, il faut prendre en considération les facteurs suivants :

- Le budget alloué à l'enquete C_0 ;
- Prendre l'avantage d'observations dans les strates ayant un poids $W_h = \frac{N_h}{N}$ plus important ;
- Prendre plus d'observations dans les strates caractérisées par une grande dispersion (σ plus grand) par rapport à celles ou la dispersion est faible ;
- Prendre plus d'observation dans les strates ou le cout d'obtention des informations (cout de l'observation C_h) est faible.

Exercice 2 :Solution

Donc la taille de l'échantillon dans la strate h est optimale pour :

$$n_h = \frac{k \times W_h \times \sigma_h}{\sqrt{C_h}}$$

Avec la constante k est déterminée d'après le budget globale C_0 par la formule suivante :

$$k = \frac{C_0}{\sum_{h=1}^{p}(W_h \times \sigma_h \times \sqrt{c_h})}$$

$$\Rightarrow k = \frac{100000}{(0,625 \times 1,5 \times \sqrt{400}) + (0,375 \times 4 \times \sqrt{300})} = 2235,6$$

Exercice 2 :Solution

D'ou la structure optimale de l'échantillon pour les deux strates est déterminée comme suit :

$$n_1 = \frac{k \times W_1 \times \sigma_1}{\sqrt{C_1}} = \frac{2235,6 \times 0,625 \times 1,5}{\sqrt{400}} = 105$$

$$n_2 = \frac{k \times W_2 \times \sigma_2}{\sqrt{C_2}} = \frac{2235,6 \times 0,375 \times 4}{\sqrt{300}} = 193$$

Exercice 2 :Solution

On vérifie bien que si on retient 105 individus dans la première strate et 193 individus dans la deuxième strate, on aura besoin d'un budget de :$(105 \times \times 400) + (193 \times 300) = 99900$ DH (ne dépasse pas le budget alloué à cette enquete fixe à 100000 DH)

Exercice 3 : Solution

1- Puisque la base de sondage est connue, la méthode à adopter afin de déterminer la taille de l'échantillon est la méthode probabiliste.

Exercice 3 : Solution

1- Puisque la base de sondage est connue, la méthode à adopter afin de déterminer la taille de l'échantillon est la méthode probabiliste.

2- Puisque le taux de sondage est égal à 5% et la taille de la population est égale à 300, donc la taille de l'échantillon est déterminée comme suit :

$$t = \frac{n}{N} \;\Rightarrow\; n = N \times t \;\Rightarrow\; n = 300 \times 0,05 = 15 \;\; clients$$

Exercice 3 : Solution

1- Puisque la base de sondage est connue, la méthode à adopter afin de déterminer la taille de l'échantillon est la méthode probabiliste.

2- Puisque le taux de sondage est égal à 5% et la taille de la population est égale à 300, donc la taille de l'échantillon est déterminée comme suit :

$$t = \frac{n}{N} \quad \Rightarrow \quad n = N \times t \quad \Rightarrow \quad n = 300 \times 0,05 = 15 \;\; clients$$

3- Pour la désignation des 15 individus de l'échantillon, on va adopter deux méthodes les plus couramment utilisées, à savoir, la méthode du sondage élémentaire (utilisation des tables au hasard) et la méthode systématique comme nous les avons présentées auparavant.

Exercice 3 : Solution

Application de la méthode du sondage systématique : Il convient de rappeler que le cadre de la méthode des tirages systématiques les unités échantillons sont prélevées dans la population suivant une progression arithmétique, la base de celle-ci étant choisie au hasard et la raison est calculée de façon à couvrir entièrement la population de référence. Pour cela :

Exercice 3 : Solution

- Il faut déterminer la raison de la progression arithmétique à partir du taux de sondage comme suit :
  Pour cet exercice le taux de sondage = 5% = $\frac{5}{100} = \frac{1}{20}$, ce qui implique que la raison est égale à 20.
- On retiendra comme base de sondage un nombre tiré au hasard entre 1 et 20, 10 par exemple.Puisque le pas de sondage ou la raison est de 20, l'échantillon comprendra les unités de rang :
  10 30 50 70 90 110 130 150 170 190 210 230 250 270 290

Exercice 4 : Solution

La méthode à utiliser est la méthode des quotas car on connait la distribution de la population (15000 entreprises) selon les variables de controle (le chiffre d'affaires : X et le nombre d'employés : Y).
Soit $t = \frac{n}{N} = 10\%$ le taux de sondage choisi par l'institut de sondage.
Avec :
n : désigne la taille de l'échantillon à constituer,
N : désigne la taille de la population.
Donc, la taille de l'échantillon est de $n = 10\% \times 15000 = 1500$ entreprises.
Commençons par les quotas relatifs à la variable de controle chiffre d'affaires, puis, nous calculerons les quotas relatifs à la variable de controle nombre d'employés.

Exercice 4 : Solution

- Les quotas relatifs à la varaible de controle chiffre d'affaires (X) :

X en milliers de DHs	Effectifs population : N_s	Quotas population	Effectifs échantillon : n_i	Répartition des enqueteurs
0-50	5500	36,67%	550	11
50-100	4500	30%	450	9
100-200	1750	11,66%	175	3
200-500	2000	13,34%	200	4
500-1000	500	3,33%	50	1
> 1000	750	5%	75	2
Total	15000 = N	100%	1500 = n	30 = Nombre des enqueteurs

Avec :

Exercice 4 : Solution

- Quotas population = $\frac{N_s}{N}$;
- L'effectif échantillon peut etre calculé de deux façons :
  - Soit : $n_i = N_s \times 10\%$
  - Soit : n_i = n(taille de l'échantillon)× quotas population ;
- La répartition des enqueteurs est obtenue en multipliant le nombre total des enqueteurs par chaque quota de la population.
- Totale :En aucune manière :
  - l'effectif de la population ne doit dépasser N ;
  - la somme des quota ne doit dépasse 100% ;
  - l'effectif de l'échantillon ne doit dépasser n ;
  - Et la répartition des enqueteurs, ne doit dépasser le total des enqueteurs.

Exercice 4 : Solution

- De la meme façon on peut calculer Les quotas relatifs à la deuxième varaible de controle : le nombre d'employés (Y) :

Y : Le nombre d'employés	Effectifs population : N_s	Quotas population	Effectifs échantillon : n_i	Répartition de enqueteurs
0-10	7050	47%	705	14
10-50	3700	24,66%	370	7
50-100	2500	16,66%	250	5
100-500	1500	10%	150	3
> 500	250	1,68%	25	1
Total	15000 = N	100%	1500 = n	30

Exercice 5 :Solution :

a) La variable de controle est une varaible qui permet d'assurer la conformité de la structure de l'échantillon par rapport à la structure de la population avec X : Age et Y : Nombre d'enfants.
b) Le taux de sondage désigne le rapport entre la taille de l'échantillon et la taille de la population mère, donc le taux de sondage $t = \frac{n}{N}$, avec $N = 1000$, ce qui implique que $n = 200$ $(n = N \times t = 1000 \times 20\%)$
c) Le principe de la méthode des quotas : en fonction des variables de controle, l'échantillon est choisi de façon à constituer une image aussi fidèle que possible de la population. En d'autres termes, la structure de l'échantillon doit etre identique à la structure de la population.

Exercice 5 :Solution :

L'application de la méthode des quotas à cet exemple - en respectant la meme démarche de l'exercice précédent - nous donne la composition de l'échantillon pour les deux varaible de controle X et Y :

X	n_i	Y	n_i
< 20	22	0	10
20-25	30	1	110
25-30	40	2	20
30-35	50	3	30
35-40	38	4	15
> 40	20	≥ 5	15
Total	200	Total	200

Exercice 5 :Solution :

e) La structure des interviews :

X	n_i	Y	n_i
< 20	2	0	1
20-25	3	1	11
25-30	4	2	2
30-35	5	3	3
35-40	4	4	1,5 $\rightarrow$ 2 ou 1
> 40	2	≥ 5	1,5 $\rightarrow$ 1 ou 2
Total	20	Total	20

Exercice 5 :Solution :

f) Les conditions de la mise en oeuvre da la méthode des quotas sont respectivement :

- Choix des varaibles de controle ;
- Organisation pratique des enqueteurs ;
- Le controle des enqueteures.