Chapitre 3
Détermination de la taille de l'échantillon

Lorsqu'on prélève un échantillon pour estimer un paramètre, on court toujours le risque de découvrir un peu trop tard que l'échantillon prélevé est trop petit ou, au contraire, trop grand. Il est *trop petit* si l'intervalle de confiance qu'il fournit est trop large, ou ne permet pas de confirmer une hypothèse; il est *trop grand* si l'intervalle de confiance est inutilement court, c'est-à-dire, plus précis et donc plus coûteux que nécessaire. Nous discutons ici quelques façons d'estimer *avant l'échantillonnage* la taille minimale de l'échantillon qui fournit la précision voulue. Nous traiterons séparément le cas d'une moyenne ou d'un total; et celui d'une proportion ou d'un effectif.

3.1 Le cas d'une moyenne ou d'un total

Nous traitons en détail le cas où le paramètre à estimer est la moyenne μ; la solution pour un total τ en découle immédiatement. Tout d'abord, nous devons clarifier la question, ce qui n'est pas évident puisque la tendance naturelle est de la poser de façon simpliste, comme « est-ce qu'un échantillon de taille 500 suffit pour une population de 10 000? ». Il est impossible, hélas! de répondre par oui ou par non. Un échantillon doit fournir une estimation aussi précise que possible de la moyenne, et nous savons que la précision s'améliore indéfiniment lorsque la taille de l'échantillon augmente. Par conséquent, on ne peut pas exiger « la meilleure précision possible »: la meilleure précision possible n'est atteinte que lorsqu'on prélève la population entière. Voici comment on doit poser la question : quelle est la taille de l'échantillon qui assure tel degré précision? Avant de répondre, précisons la notion de « précision ».

Nous exprimerons la précision par sa marge d'erreur: un estimateur est précis si la marge d'erreur d'un intervalle de confiance est étroite. Il faut donc, avant de commencer, décider de la marge d'erreur qu'on est disposé à tolérer. Par exemple, si on tire un échantillon de ménages afin d'estimer le revenu moyen des ménages, on construira un intervalle de confiance qui prendra la forme

Estimateur ± marge d'erreur

On pourrait aboutir à quelque chose comme

40 000 $ ± 25 000 $

En d'autres termes, on affirmerait que le revenu moyen se situe entre 15 000 $ et 65 000 $. Manifestement, un tel intervalle de confiance est presque inutile, la marge d'erreur est trop grande. Le jour où l'on reprendra cet échantillonnage, on voudra s'assurer que la marge d'erreur soit plus étroite, disons 5 000 $ plutôt que 25 000 $. Il faudra alors déterminer la taille de l'échantillon *n* qui donnera une telle marge d’erreur. Rappelons que la marge d’erreur est donnée par $2\sigma_{\bar{y}}$ [1]. Nous devons donc déterminer la taille de l'échantillon nécessaire pour que la marge d’erreur soit égale à un nombre donné *E :*

$$\boxed{\textit{Exigé}:\ \ 2\,\sigma_{\bar{y}} = E}$$

Puisque $2\,\sigma_{\bar{y}} = 2\sqrt{1-f}\,\dfrac{S}{\sqrt{n}}$, où $f = \dfrac{n}{N}$, nous devons déterminer la valeur de n qui satisfait l'équation

$$2\sqrt{1-f}\,\frac{S}{\sqrt{n}} = E.$$

Oublions pour l'instant le facteur de correction $\sqrt{1-f}$, qui de toute façon est négligeable dans la plupart des situations réelles. L’équation à résoudre devient

$$\frac{2S}{\sqrt{n}} = E\,,$$

ce qui donne la solution provisoire $n_o = (2S/E)^2$.

C’est donc la règle que nous allons suivre en général :

> *Si N est assez grand, la valeur de n pour laquelle* $2\sigma_{\bar{y}} = E$ *est approximativement*
>
> $$n_o = \left(\frac{2S}{E}\right)^2.$$

Cette solution correspond au cas où $f = 0$, c’est-à-dire, lorsque la population est infinie, et dans la plupart des applications, la solution provisoire n_o produit une bonne approximation de la solution réelle n. Mais si, comme dans l’exemple traité dans cette section, le facteur de correction est important (si N n’est pas très grand), il suffit d'un simple ajustement pour obtenir la solution exacte. L’ajustement est :

> *Lorsque N n’est pas très grand, il convient d’ajuster la taille estimée par l’équation*
>
> $$n = \frac{n_o}{1 + \dfrac{n_o}{N}}$$

[1] En pratique c’est $\hat{\sigma}_{\bar{y}}$ qu’on calcule. Mais idéalement, on préférerait avoir $\sigma_{\bar{y}}$

***Exemple* 3.1.1** *Taille d'échantillon pour une marge d'erreur donnée*

Considérons l'échantillon présenté au tableau A.02 de l'annexe et supposons qu'on veuille estimer le montant total des salaires de 2001. La population elle-même est constituée de $N = 200$ professeurs et l'échantillon est de taille $n = 50$. La moyenne et l'écart-type de la variable Y = *salaire de 2001* sont:

$$\bar{y} = 48\,447{,}32 \text{ \$ et } s = 10\,025{,}95$$

Supposons qu'on ait l'intention de prélever l'an prochain un deuxième échantillon de la même population afin d'estimer le salaire moyen des professeurs. Quelle doit être la taille de l'échantillon pour que la marge d'erreur soit de 1 000 $?

Nous supposerons que le nombre de professeurs d'universités l'an prochain ne changera pas, donc N = 200; et que l'écart-type S restera à peu près constant. Nous estimerons donc S par s = 10 025,95 $. Comme première approximation, donc, nous avons $n_o = (2s/1000)^2 = 402{,}0789$. Faisons l'ajustement:

$$n = \frac{n_o}{1+\frac{n_o}{N}} = \frac{402{,}0789}{1+\frac{402{,}0789}{200}} = 134$$

On doit donc prélever un échantillon d'environ 134 professeurs pour que la marge d'erreur de l'estimateur soit de 1 000 $ □

Remarque *Dans l'exemple que nous venons de traiter, il a fallu faire deux suppositions: l'une est que N ne changera pas; l'autre est que l'écart-type ne changera pas. En général, la première supposition ne pose pas de problème, car N ne joue pas un grand rôle: la population est normalement si grande qu'une petite erreur dans l'estimation de N est sans effet.* (*L'exemple qu'on vient de traiter n'est pas typique en ce sens. La petitesse de la population là faisait qu'il a été utile de bien estimer sa taille.*)

La deuxième supposition est plutôt douteuse, et plus lourde de conséquences. En général, si on s'attend à une augmentation des salaires, on peut aussi s'attendre à une augmentation de l'écart-type S. Donc il est sage de majorer un peu l'estimation de S.

En pratique il sera presque toujours nécessaire de faire des suppositions à propos de la population afin d'estimer la taille de l'échantillon. Heureusement, on est souvent moins démuni dans un contexte réel que dans une situation hypothétique comme celle-ci. Dans notre exemple, on pourrait avoir de bonnes raisons de croire que les salaires augmenterons, disons, de 5 %. Dans ce cas, on s'attend à ce que l'écart-type aussi augmente de 5%, et alors on estimera S par 10025,95 (1,05) = 10527,25. Avec cette nouvelle estimation, on trouvera n = 94 au lieu de n = 90 . □

Remarque *La marge d'erreur est la demi-largeur d'un intervalle de confiance. Dans ce cours, le niveau de confiance est toujours de 95 %, et donc la marge d'erreur dont il est question ici est une « marge d'erreur à 95 % ». Cette qualification est sous-entendue. Il faut se rappeler, cependant, qu'il est possible de déterminer des intervalles de confiance à d'autres niveaux que 95 %.*

Quel est le sens de ce 95 %? On a une marge d'erreur, disons, de 5 000 $ dans l'estimation de la moyenne des salaires si la probabilité est de 95 % que l'erreur d'estimation ne dépassera pas 5 000 $; ou si la probabilité est de 5 % que l'erreur d'estimation dépasse 5 000 $. □

La marge d'erreur dans le dernier exemple est exprimée en dollars. Souvent, la marge d'erreur est exprimée en pourcentage : une marge d'erreur de 1 000 $ est peut-être acceptable dans une population où la moyenne est de 50 000 $, mais elle serait sûrement jugée excessive dans une population où la moyenne est de 20 000 $. En d'autres termes, c'est plutôt la marge d'erreur *relative* qui compte, c'est-à-dire, la marge d'erreur divisée par la moyenne. On pourrait dire, par exemple, qu'on souhaite estimer le revenu familial d'une population avec une marge d'erreur de 10 %—on veut dire par là 10 % de la moyenne. En d'autres termes, on exige que

$$2\sigma_{\bar{y}} = 0{,}10\mu$$

ou, plus généralement,

$$2\sigma_{\bar{y}} = R\mu$$

Algébriquement, le problème n'a pas changé : on a juste remplacé E par $R\mu$. Par conséquent, nous avons déjà la formule :

Si N est assez grand, la valeur de n pour laquelle $2\sigma_{\bar{y}} = R\mu$ *est approximativement* $$n_o = \left(\frac{2S}{R\mu}\right)^2.$$

Remarque *La condition* $2\sigma_{\bar{y}} = R\mu$ *est équivalente à* $2N\sigma_{\bar{y}} = RN\mu = R\tau$. Donc la valeur de n_o trouvée donnera une marge d'erreur relative de R dans l'estimation du total aussi.

***Exemple* 3.1.2** *Taille d'échantillon pour une marge d'erreur relative donnée*

Dans l'exemple précédent, supposons qu'on veuille estimer la moyenne de telle sorte que la marge d'erreur soit de 5 %.

Donc

$$n_o = \left(\frac{2S}{R\mu}\right)^2 = \left(\frac{2S}{0{,}05\mu}\right)^2 = \left(\frac{2S/\mu}{0{,}05}\right)^2$$

On estime S par s et μ par $\bar{y}$ et donc S/μ par 10025,95/48447,32 = 0,2069455. Donc

$$n_o = \left(\frac{2(0{,}2069455)}{0{,}05}\right)^2 = 68{,}52228$$

L'ajustement porte ce chiffre à

$$n = \frac{68{,}52228}{1+\frac{68{,}52228}{200}} = 51.$$ □

Remarque *Dans le dernier exemple, il a fallu estimer S et* μ, *deux paramètres distincts. On constate, cependant, que c'est le quotient des deux qu'il faut estimer puisque la formule de* n_o *ne dépend que du quotient* S/μ (*c'est-à-dire,* du coefficient de variation *de la population*)

$$n_o = \left(\frac{2(S/\mu)}{0{,}05}\right)^2$$

On a estimé le coefficient de variation de la population par celui de l'échantillon $s/\bar{y}$ = 0,2069455.

En pratique, il est souvent plus aisé d'estimer le coefficient de variation que l'écart-type. Dans l'exemple ci-dessus, l'écart-type aura peut-être changé en un an sous l'effet d'une augmentation de salaire; mais si les salaires augmentent, disons, de 5%, la moyenne et l'écart-type augmenteront tous deux de 5%. Par conséquent, le coefficient de variation restera fixe □

Remarque *Contrairement à ce qu'on a tendance à croire, la taille de la population joue un rôle relativement mineur dans la détermination de la taille de l'échantillon, à moins qu'elle ne soit très petite. Le tableau ci-dessous reprend le problème présenté à l'exemple 3.1.2: quelle est la taille nécessaire pour que la marge* *d'erreur dans l'estimation de la moyenne soit de 5 %. On présente la réponse en fonction de la taille de la population.*

Taille de la population	100	200	300	400	500	800	1000	5000	10000	15000	20000	∞
Taille d'échantillon nécessaire	41	51	56	59	60	63	64	68	68	68	68	69

On constate que dès que la population atteint 5 000, il n'est plus vraiment nécessaire d'augmenter la taille de l'échantillon. □

Estimation des* paramètres *inconnus
Les problèmes présentés dans cette section n'ont de solution que si l'on connaît, ou si l'on peut estimer, les valeurs de S ou de S/μ, selon le cas. Dans les exemples traités ici, on disposait d'un échantillon qui permettait d'estimer ces paramètres, du moins approximativement. Que fait-on si on n'a pas d'échantillon? On peut envisager plusieurs approches. Nous supposons dans ce qui suit que la population est grande et donc qu'il n'est pas nécessaire de faire de distinction entre σ et S .

Échantillonnage préliminaire On peut estimer S par s, μ *par* $\bar{y}$ et donc le coefficient de *variation* $C = S/\mu$ par $s/\bar{y}$ à l'aide d'un petit échantillon préliminaire tiré à cette fin. La taille de cet échantillon préliminaire sera généralement assez petite, puisque l'estimation ne doit pas nécessairement être très précise; elle dépendra surtout des ressources qu'on voudra bien lui accorder.

Estimation du coefficient de variation C
Lorsqu'on n'a pas d'échantillon pilote, l'estimation de $C = S/\mu$ est plus délicate que celle de S, puisqu'il y a deux paramètres à estimer.

Un avantage, par contre, du coefficient de variation est qu'il est plus stable d'une population à l'autre que l'écart-type S. Deux populations peuvent être sensiblement différentes quant à leur écart-type, mais comparables quant à leur coefficient de variation. Car S et μ évoluent souvent dans le même sens: lorsque les valeurs de la variable sont grandes, S et μ ont toutes deux tendance à être grands; lorsqu'ils sont *petits*, S et μ sont tous deux petits. Par conséquent le quotient S/μ ne sera pas très différent dans les deux populations. Toute expérience avec des populations relativement semblables peut donc servir à estimer C : la même population l'année dernière; une population analogue dans une autre compagnie, une autre division, une autre succursale; une autre variable dans la même population, etc.

3.2 Le cas d'une proportion ou d'un effectif

La situation avec l'estimation d'une proportion p est semblable à celle d'une moyenne, sauf que le caractère un peu particulier de l'écart-type de $\hat{p}$ permet certaines estimations qu'on ne peut pas faire dans le cas de variables quantitatives. Le problème posé est le même: on exige que la marge d'erreur de l'estimateur soit égal à un certain nombre E :

$$2\sigma_{\hat{p}} = E\,,$$

Donc il faut déterminer la valeur de n pour laquelle

$$2\sqrt{\frac{p(1-p)}{n}}\ \sqrt{\frac{N-n}{N-1}} = E.$$

Si l'on néglige provisoirement le facteur de correction $\sqrt{\frac{N-n}{N-1}}$, on a l'équation suivante à résoudre:

$$2\sqrt{\frac{p(1-p)}{n}} = E.$$

On a donc la solution provisoire

$$n_o = \frac{4p(1-p)}{E^2}.$$

Cette solution provisoire est généralement acceptable, à moins que la population ne soit plutôt petite par comparaison à l'échantillon. Si la population n'est pas très grande, alors on effectue l'ajustement suivant:

$$n = \frac{n_o}{1+\frac{n_o-1}{N}} \approx \frac{n_o}{1+\frac{n_o}{N}}$$

Pour ne pas multiplier les formules inutilement, nous utiliserons toujours l'expression approximative $\frac{n_o}{1+\frac{n_o}{N}}$, puisqu'elle est identique à celle que nous avons employée pour la moyenne.

Le fait que nous ne connaissions pas la valeur de p est moins grave qu'on pourrait le croire, car il est possible de trouver une borne au produit $p(1\text{-}p)$. La borne la plus conservatrice (c'est-à-dire, la plus grande) est celle-ci:

$$p(1\text{-}p) \leq 1/4.$$

C'est un fait que le produit $p(1\text{-}p)$ n'est jamais supérieur à 1/4, une valeur atteinte seulement lorsque p = ½. Donc l'estimation la plus pessimiste qu'on puisse obtenir de n_o est celle où l'on remplace p par ½. On a alors:

$$n_o = \frac{1}{E^2}$$

***Exemple* 3.2.1** *Taille d'échantillon pour l'estimation d'une proportion avec marge d'erreur donnée*

Le responsable de marketing dans une compagnie de boissons gazeuses veut estimer la proportion p de consommateurs qui préfère un nouveau produit destiné à remplacer l'ancien. Quelle est la taille de l'échantillon qu'il doit prélever s'il veut que la marge d'erreur de son estimateur ne soit pas supérieur à 0,04?

La valeur de n_o est $1/E^2 = 1/(0{,}04)^2 = 625$. Nous ne ferons pas d'ajustement, puisque la population est à toute fin pratique infinie. Donc si on prélève un échantillon de taille 625, l'écart-type de l'estimateur $\hat{p}$ ne sera pas supérieur à 0,02. □

L'estimation que nous avons faite dans le dernier exemple est correcte si de fait $p = ½$: la marge d'erreur serait alors égale à 0,04. Lorsque $p \neq 1/2$, l'estimation est pessimiste: elle donne une valeur de n_o plus grande que nécessaire. Le tableau suivant donne la taille de l'échantillon nécessaire sous plusieurs hypothèses concernant p. On constate qu'on ne commet pas une erreur très importante lorsque p n'est pas trop éloignée de 0,5. Mais lorsque p est petit, on surestime n.

p	0,01	0,05	0,10	0,20	0,30	0,40	0,50
n_o	25	119	225	400	525	600	625

Il est rare en pratique qu'on soit à un tel point démuni qu'on ne peut estimer p qu'en le remplaçant par 1/2. On a souvent une idée de la valeur de p, ce qui permet de fixer une borne. Par exemple, dans un sondage politique destiné à estimer le pourcentage d'électeurs qui voterons NPD, on sait d'avance que ce pourcentage n'atteindra pas 50 %. On peut même dire, avec un peu d'expérience, que ce pourcentage est inférieur à, disons, 30 %. Dans ce cas, on sait que $p(1\text{-}p) < (0{,}30)(0{,}70) = 0{,}21$. La raison pour cela est claire lorsqu'on examine la fonction $p(1\text{-}p)$. Le graphique de cette fonction est présenté à la figure ci-dessous.

On voit dans le graphique que si p est inférieure à un certain nombre $p_o < 1/2$, alors $p(1\text{-}p) < p_o(1\text{-}p_o)$. De même, si p est supérieure à un nombre $p_o > 1/2$, alors $p(1\text{-}p) > p_o(1\text{-}p_o)$.

***Exemple* 3.2.2** *Taille d'échantillon pour l'estimation d'une proportion*

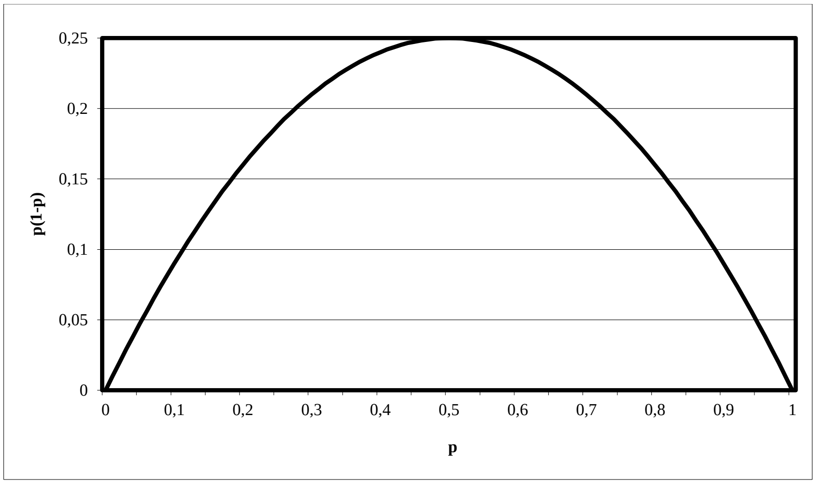
Vous voulez estimer le pourcentage de pièces défectueuses dans un lot de 5000 pièces. Vous savez que normalement, le taux de défectuosité est de l'ordre de 5 %, et vous pensez qu'il est encore de cet ordre là. Il est possible qu'il ait augmenté quelque peu, mais vous êtes sûr qu'il n'est pas supérieur à 10%.

a) Quelle est la taille de l'échantillon que vous devriez tirer si vous voulez que votre estimateur ait une marge d'erreur de 0,04?

b) Quelle est la taille de l'échantillon que vous devriez tirer si vous voulez avoir une marge d'erreur *relative* de 25 %? Vous savez d'avance que p ne peux avoir baissé au-dessous de 3 %.

Solution

a) Nous avons $n_o = 4p(1\text{-}p)/(0{,}04)^2$, et puisque le numérateur prend au maximum la valeur (0,1)(0,9) = 0,09, on a $n_o = 4(0{,}09)/0{,}0016 = 225$. L'ajustement donne $n = 225/(1+225/5000) \approx 216$. Remarquez que si on avait pris la valeur 1/2 pour p, nous aurions obtenu $n_o = 625$ et $n = 556$, beaucoup plus élevé que nécessaire.



b) Ce que vous voulez, c'est que $2\sigma_{\hat{p}}$ = 0,25p ⇒ n_o = 4p(1-p)/(0,25p)^2 = $\frac{4(1-p)/p}{0{,}25^2}$. La plus grande valeur de $\frac{4(1-p)/p}{0{,}25^2}$ est atteinte lorsque p est à son minimum; ce minimum est, selon la description de la situation, 0,03.

On prendra donc $n_o = \frac{4(1-0{,}03)/0{,}03}{0{,}25^2}$ = 2069. □

3.3 Résumé

1 *Marge d'erreur absolue* : Si N est assez grand, la valeur de n pour laquelle $2\sigma_{\bar{y}} = E$ est approximativement

$$n_o = \left(\frac{2S}{E}\right)^2.$$

2. On ajuste cette estimation provisoire à l'aide de la formule $n = \dfrac{n_o}{1+\dfrac{n_o}{N}}$

3 La marge d'erreur est égale à E si la probabilité est de 95% (à peu près) que l'erreur ne sera pas supérieure à E.

4 Si N est assez grand, la valeur de n pour laquelle $2\sigma_{\bar{y}} = R\mu$ est approximativement

$$n_o = \left(\frac{2S}{R\mu}\right)^2.$$

5 Pour estimer une proportion p de telle sorte que $2\sigma_{\hat{p}} = E$, la taille approximative de l'échantillon qu'il faut tirer est donnée par $n_o = \dfrac{4p(1-p)}{E^2}$. Pour une population de taille N, on peut ensuite

faire l'ajustement habituel $n \approx \dfrac{n_o}{1+\dfrac{n_o}{N}}$. Si p est inférieure à un certain nombre $p_o < 1/2$, alors $p(1-p) < p_o(1-p_o)$. De même, si p est supérieure à un nombre $p_o > 1/2$, alors $p(1-p) > p_o(1-p_o)$. Lorsque la valeur de p est totalement inconnue, on peut remplacer p dans la formule par ½.

3.4 Exercices

3.1 D'une population de 1 640 étudiants, on veut prélever un échantillon pour estimer certains paramètres concernant la variable y : montant dépensé quotidiennement en transport. Un échantillonnage préliminaire a donné une estimation de μ_y et de S_y: $\mu_y \approx 140$, $S_y \approx 25$. Déterminer la taille de l'échantillon qu'on doit prélever s'il faut que

a) la marge d'erreur de l'estimateur de la moyenne soit de 2 $.

b) la marge d'erreur de l'estimateur de la moyenne soit de 4 %.

c) la marge d'erreur de l'estimateur du total soit de 2 000 $.

d) la marge d'erreur de l'estimateur du total soit de 4 %.

3.2 D'une population de taille 678 vous devez prélever un échantillon dans le but d'estimer la moyenne de la population avec une marge d'erreur de 10 %. Vous prélevez un échantillon préliminaire de taille 12 et obtenez les résultats suivants:

5,17	42,21	81,26	52,34	64,15	3,06
30,36	26,96	24,24	4,54	120,24	48,68

Estimez la taille de l'échantillon que vous devriez prélever.

3.3 [*Tableau* A.06]Vous êtes en 2002, et vous disposez des données du tableau A.06 obtenues en 2000 sur un échantillon de 58 villes du monde. Vous allez maintenant prélever un échantillon de villes afin d'estimer le nombre moyen d'heures de travail des ingénieurs en 2001. Quelle est la taille de l'échantillon qu'il faudrait prélever si vous voulez estimer le nombre moyen d'heures de travail des ingénieurs avec une marge d'erreur de 1 % si vous comptez estimer votre moyenne par la moyenne échantillonnale? Vous supposerez que la population des villes est très grande.

3.4 [*Tableau* A.07] Le tableau A.07 présente des données sur un échantillon de 102 maisons vendues en 2001. Vous avez l'intention de mener une étude comparable en 2002 afin d'estimer le prix moyen des maisons. Vous supposerez que la population visée est très grande, et que le prix des maisons en 2002 est probablement supérieur à celui de 2001 de quelque 15%. Déterminez la taille de l'échantillon qu'il faudrait prélever si vous voulez que votre marge d'erreur à 95% soit d'environ 5 000 $.

3.5 Vous voulez estimer le nombre d'étudiants dans une certaine université qui ont accès à l'Internet. Le nombre d'étudiants dans l'université est de 45 635. Sachant par expérience que la proportion d'étudiants ayant accès à l'Internet se situe quelque part entre 15 % et 25 %, déterminez la taille de l'échantillon que vous devriez prélever

a) si vous tenez à ce que votre marge d'erreur soit d'environ 500 étudiants;

b) si vous tenez à ce que votre marge d'erreur relative soit de 10 %.

3.6 Vous voulez estimer le nombre d'étudiants dans une certaine université qui possèdent un ordinateur. Le nombre d'étudiants dans l'université est de 45 635. Sachant par expérience que la proportion d'étudiants qui possèdent un ordinateur se situe quelque part entre 40 % et 65 %, déterminer la taille de l'échantillon que vous devriez prélever

a) si vous tenez à ce que votre marge d'erreur soit d'environ 1 000 étudiants;

b) si vous tenez à ce que votre marge d'erreur relative soit de 2 %.

3.7 Vous voulez estimer la proportion d'étudiants dans une certaine université qui ont accès à l'Internet. Le nombre d'étudiants dans l'université est de 45 635. Sachant par expérience que la proportion d'étudiants ayant accès à l'Internet se situe quelque part entre 15 % et 25 %, déterminez la taille de l'échantillon que vous devriez prélever

a) si vous tenez à ce que votre marge d'erreur soit d'environ 2 % (c'est-à-dire, 2 points de pourcentage);

b) si vous tenez à ce que votre marge d'erreur *relative* soit de 10 % (par rapport à la vraie proportion p) .

3.8 Vous voulez estimer la proportion d'étudiants dans une certaine université qui possèdent un ordinateur. Le nombre d'étudiants dans l'université est de 45 635. Sachant par expérience que la proportion d'étudiants qui possèdent un ordinateur se situe quelque part entre 40 % et 65 %, déterminer la taille de l'échantillon que vous devriez prélever

a) si vous tenez à ce que votre marge d'erreur soit d'environ 1 %;

b) si vous tenez à ce que votre marge d'erreur relative soit de 5 %.