Échantillonnage

Chapitre 12 : L'estimateur par le ratio

Exercices corrigés

Jaouad Madkour
Faculté d'Économie et de Gestion de Tanger
Échantillonnage

Table des matières

Partie 1 — Énoncés 1

Partie 2 — Corrections 4

1 Exercices du chapitre

Exercice 1 — Biais et variance approchés

Une commune compte $N = 3\,000$ ménages et, selon le recensement, $X = 13\,500$ personnes. On veut estimer la dépense alimentaire mensuelle moyenne $\overline{Y}$ des ménages par un sondage aléatoire simple de $n = 150$ ménages, avec l'estimateur par le ratio sur la taille du ménage, $\widehat{\overline{Y}}_R = \overline{X}\,\overline{y}/\overline{x}$. Une enquête antérieure fournit des ordres de grandeur que l'on prendra pour exacts : $\overline{Y} = 3\,600$ dh, $S_y = 1\,800$ dh, $S_x = 2$ personnes, et une corrélation $\rho = 0{,}7$ entre la dépense et la taille.

(a) Calculer $\overline{X}$, $R = \overline{Y}/\overline{X}$, CV_x et CV_y. La condition de gain du ratio est-elle remplie ?

(b) Calculer $S_e^2 = S_y^2 - 2R\,\rho\,S_y\,S_x + R^2 S_x^2$, l'effet de plan du ratio, puis les erreurs-types de $\overline{y}$ et de $\widehat{\overline{Y}}_R$.

(c) Calculer le biais approché $\overline{Y}\,(1-f)\left(CV_x^2 - \rho\,CV_x\,CV_y\right)/n$ et le comparer à l'erreur-type. Quelle part de l'erreur quadratique moyenne représente-t-il ?

(d) Combien de ménages faudrait-il interroger par sondage aléatoire simple, en estimant par $\overline{y}$, pour atteindre la même précision ? Comparer à la taille effective $n/deff$ et expliquer la différence.

(e) L'estimateur par la régression (chapitre 13) a un effet de plan $1-\rho^2$. Comparer. Calculer la pente $B = \rho\,S_y/S_x$ et l'ordonnée à l'origine $\overline{Y} - B\,\overline{X}$ de la droite des moindres carrés, et expliquer l'écart.

Exercice 2 — La mini-population avec trois ménages

On reprend les cinq ménages du cours :

Ménage	A	B	C	D	E
Dépense y_k (milliers de dh)	4	6	7	9	14
Taille x_k (personnes)	2	3	3	4	6

On tire cette fois $n = 3$ ménages par sondage aléatoire simple, et l'on estime $\overline{Y} = 8$ par $3{,}6\,\overline{y}/\overline{x}$.

(a) Énumérer les dix échantillons et calculer, pour chacun, $\overline{y}$, $\overline{x}$ et l'estimation par le ratio.

(b) Calculer l'espérance, le biais, la variance et l'erreur quadratique moyenne de l'estimateur par le ratio. Les comparer à la variance de $\overline{y}$, $(1-f)\,S^2/n$ avec $S^2 = 14{,}5$.

(c) Pour $n = 2$, le cours a obtenu un biais de $-0{,}0557$ et une variance de $0{,}1078$. Par combien le biais et la variance sont-ils divisés quand n passe de 2 à 3 ? Comparer à l'évolution du facteur $(1-f)/n$.

(d) Sachant que $S_x^2 = 2{,}3$, $S_{xy} = 5{,}75$ et $S_e^2 = 0{,}3025$, calculer le biais et la variance approchés :

$$B \approx \frac{1-f}{n}\cdot\frac{R\,S_x^2 - S_{xy}}{\overline{X}}, \qquad V \approx \frac{1-f}{n}\,S_e^2$$

Les approximations sont-elles meilleures que pour $n = 2$ (où elles donnaient $-0{,}0532$ et $0{,}0907$) ?

Exercice 3 — La dépense alimentaire par personne

Dans une commune périurbaine de $N = 400$ ménages, on tire $n = 10$ ménages par sondage aléatoire simple. On relève leur taille x_k, leur dépense alimentaire mensuelle y_k (en centaines de dh) et leur milieu de résidence :

Ménage	1	2	3	4	5	6	7	8	9	10
Taille x_k	2	3	4	4	5	6	3	5	7	1
Dépense y_k	12	15	20	18	22	28	17	24	30	8
Rural	non	non	oui	non	oui	oui	non	non	oui	non

(a) Estimer la dépense alimentaire par personne $R = Y/X$ de la commune. Faut-il connaître le

nombre total de personnes X ?

(b) Calculer les résidus $\hat{e}_k = y_k - \widehat{R}\,x_k$ et vérifier que leur somme est nulle. Estimer l'erreur-type de $\widehat{R}$ et donner un intervalle de confiance à 95 %.

(c) Calculer la moyenne des dix rapports y_k/x_k. Pourquoi diffère-t-elle de $\widehat{R}$? Laquelle estime R ?

(d) Estimer la dépense alimentaire moyenne des ménages ruraux. Calculer son erreur-type de deux façons : par la linéarisation du cours, avec $e_k = \delta_k\,(y_k - \overline{y}_d)$; puis par la formule du chapitre 4, $(1-f)\,s_d^2/n_d$. Comparer.

2 Exercice cumulatif (chapitres 12, 11, 10 et 9)

Exercice 4 — Les commerces d'une ville moyenne

Une ville compte $N_I = 30$ quartiers et, selon le registre du commerce, $N = 1\,500$ commerces. On tire 4 quartiers par sondage aléatoire simple, puis 10 commerces par sondage aléatoire simple dans chaque quartier tiré, et l'on relève leur chiffre d'affaires mensuel (chiffres fictifs) :

Quartier tiré	1	2	3	4
Nombre de commerces M_i	80	45	60	35
Chiffre d'affaires moyen des 10 commerces $\overline{y}_i$ (milliers de dh)	42	30	36	25

(a) Calculer la probabilité d'inclusion et le poids d'un commerce de chaque quartier tiré. Calculer la somme des poids des 40 commerces interrogés et la comparer à N.

(b) Estimer le chiffre d'affaires total des commerces de la ville par l'estimateur sans biais $\widehat{Y} = (N_I/n_I)\sum_i M_i\,\overline{y}_i$, puis le chiffre d'affaires moyen par $\widehat{Y}/N$.

(c) Estimer le nombre de commerces par $\widehat{N} = (N_I/n_I)\sum_i M_i$, puis la moyenne par le ratio $\widehat{Y}/\widehat{N}$ et le total par $N\,\widehat{Y}/\widehat{N}$. Pourquoi les deux estimations diffèrent-elles ?

(d) En négligeant la variance du second degré, estimer l'erreur-type de l'estimateur par le ratio avec la formule du chapitre 9 : $\dfrac{1-f_I}{n_I\,\overline{M}^2}\,s_e^2$, où $\overline{M} = N/N_I$ et s_e^2 est la variance des résidus $M_i\,\overline{y}_i - \widehat{R}\,M_i$.

(e) On aurait pu tirer les quartiers avec des probabilités proportionnelles à leur nombre de commerces (chapitre 11), puis 10 commerces par quartier. Calculer alors la probabilité d'inclusion d'un commerce. Si les mêmes quartiers avaient été tirés, quelle estimation de la moyenne obtiendrait-on ? Quel point commun avec le ratio ?

3 Applications en sciences économiques

Exercice 5 — La production d'olives d'une province

Une province compte $N = 800$ exploitations oléicoles ; le recensement agricole donne leur superficie totale en oliviers, $X = 4\,000$ ha. Un sondage aléatoire simple de $n = 40$ exploitations donne, pour la production d'olives y (en tonnes) et la superficie en oliviers x (en ha), les résultats suivants (chiffres fictifs) :

$$\overline{y} = 9{,}2, \qquad \overline{x} = 4{,}6, \qquad s_y = 5{,}5, \qquad s_x = 3, \qquad r = 0{,}88$$

(a) Estimer la production totale de la province par $N\,\overline{y}$, avec son erreur-type et un intervalle de confiance à 95 %.

(b) Estimer le rendement moyen R (tonnes par hectare), puis la production totale par l'estimateur par le ratio, avec son erreur-type et un intervalle de confiance à 95 %.

(c) Calculer l'effet de plan estimé du ratio, et vérifier la condition de gain sur l'échantillon.

(d) Donner un intervalle de confiance à 95 % du rendement.

(e) Pourquoi les deux estimations de la production diffèrent-elles ? Laquelle retenir, et pourquoi faut-il l'avoir décidé avant de voir les résultats ?

Exercice 6 — L'évolution du chiffre d'affaires

Une chambre de commerce suit $N = 600$ entreprises industrielles. Le fichier fiscal lui donne leur chiffre d'affaires total de l'année 2024 : $X = 720$ millions de dh. Pour estimer le chiffre d'affaires de 2025 sans attendre ce fichier, elle interroge un sondage aléatoire simple de $n = 60$ entreprises sur leur chiffre d'affaires de 2025 (y) et de 2024 (x), en millions de dh (chiffres fictifs) :

$$\bar{y} = 1{,}32, \qquad \bar{x} = 1{,}25, \qquad s_y = 0{,}9, \qquad s_x = 0{,}8, \qquad r = 0{,}95$$

(a) Estimer le taux d'évolution du chiffre d'affaires entre 2024 et 2025 par $\widehat{R} = \bar{y}/\bar{x}$, avec son erreur-type et un intervalle de confiance à 95 %. Peut-on conclure à une hausse ?

(b) Estimer le chiffre d'affaires total de 2025 par $N\bar{y}$, puis par $X\widehat{R}$, avec leurs erreurs-types et leurs intervalles à 95 %.

(c) Comparer $N\bar{x}$ à X. Que révèle cette comparaison sur l'échantillon ? Comment le ratio en tient-il compte ?

(d) Calculer l'effet de plan estimé du ratio. Pourquoi le ratio est-il si efficace pour suivre une évolution ?

1 Exercices du chapitre

Correction — Exercice 1

(a) $\overline{X} = 13\,500/3\,000 = 4{,}5$ personnes ; $R = 3\,600/4{,}5 = 800$ dh par personne ; $CV_x = 2/4{,}5 = 0{,}444$ et $CV_y = 1\,800/3\,600 = 0{,}5$. Le seuil vaut $CV_x/(2\,CV_y) = 0{,}444$: avec $\rho = 0{,}7$, la condition de gain est remplie.
(b) $S_e^2 = 1\,800^2 - 2 \times 800 \times 0{,}7 \times 1\,800 \times 2 + 800^2 \times 2^2 = 3\,240\,000 - 4\,032\,000 + 2\,560\,000 = 1\,768\,000$, soit $S_e = 1\,329{,}7$ dh. Effet de plan : $1\,768\,000/3\,240\,000 = 0{,}546$. Avec $(1-f)/n = 0{,}95/150 = 0{,}006333$:

- $V(\overline{y}) = 0{,}006333 \times 3\,240\,000 = 20\,520$, erreur-type 143,25 dh ;
- $V(\widehat{\overline{Y}}_R) \approx 0{,}006333 \times 1\,768\,000 = 11\,197{,}3$, erreur-type 105,82 dh.

(c) $CV_x^2 = 0{,}19753$ et $\rho\,CV_x\,CV_y = 0{,}7 \times 0{,}4444 \times 0{,}5 = 0{,}15556$; différence 0,04198. Biais approché : $3\,600 \times 0{,}006333 \times 0{,}04198 = 0{,}96$ dh, soit 0,9 % de l'erreur-type. Son carré ne fait que 0,008 % de l'erreur quadratique moyenne : avec 150 ménages, le biais est négligeable.
(d) Il faut $\left(\frac{1}{n} - \frac{1}{N}\right) S_y^2 = 11\,197{,}3$, soit

$$\frac{1}{n} = \frac{11\,197{,}3}{3\,240\,000} + \frac{1}{3\,000} = 0{,}0034560 + 0{,}0003333 = 0{,}0037893,$$

d'où $n = 263{,}9$: il faudrait 264 ménages au lieu de 150. La taille effective $150/0{,}546 = 274{,}9$ est un peu plus forte : elle ignore qu'en interrogeant plus de ménages, on augmente aussi le taux de sondage f, ce qui réduit la variance de $\overline{y}$.
(e) La régression aurait un effet de plan $1 - 0{,}7^2 = 0{,}51$, contre 0,546 pour le ratio. L'écart vaut $(\rho - k)^2 = (0{,}7 - 0{,}889)^2 = 0{,}036$, avec $k = CV_x/CV_y$. La droite des moindres carrés a pour pente $B = 0{,}7 \times 1\,800/2 = 630$ dh par personne et pour ordonnée à l'origine $3\,600 - 630 \times 4{,}5 = 765$ dh : un ménage dépense en moyenne 765 dh plus 630 dh par personne. Le ratio impose une droite qui passe par l'origine, de pente 800 : il s'écarte de la meilleure droite, et perd un peu de précision.

Correction — Exercice 2

(a) $\binom{5}{3} = 10$ échantillons, chacun de probabilité 1/10 ; $\overline{X} = 18/5 = 3{,}6$:

s	ABC	ABD	ABE	ACD	ACE	ADE	BCD	BCE	BDE	CDE
$\overline{y}$	5,667	6,333	8	6,667	8,333	9	7,333	9	9,667	10
$\overline{x}$	2,667	3	3,667	3	3,667	4	3,333	4	4,333	4,333
Ratio	7,650	7,600	7,855	8,000	8,182	8,100	7,920	8,100	8,031	8,308

(b) La somme des dix ratios vaut 79,745 : $E = 7{,}974$, biais $-0{,}0255$. Variance 0,0452 ; la somme des carrés des écarts à 8 vaut 0,4587, d'où une erreur quadratique moyenne de 0,0459. Pour $\overline{y}$: $V(\overline{y}) = (1 - 0{,}6) \times 14{,}5/3 = 1{,}933$, soit 42 fois l'erreur quadratique moyenne du ratio.
(c) Le biais passe de $-0{,}0557$ à $-0{,}0255$ (divisé par 2,18), la variance de 0,1078 à 0,0452 (divisée par 2,38). Le facteur $(1-f)/n$ passe de 0,3 à 0,1333 : divisé par 2,25. Biais et variance décroissent tous deux comme $(1-f)/n$, donc le biais décroît plus vite que l'**erreur-type** : le rapport biais/erreur-type passe de 0,17 à 0,12.
(d) $R = 40/18 = 2{,}2222$ et $R\,S_x^2 - S_{xy} = 2{,}2222 \times 2{,}3 - 5{,}75 = -0{,}6389$. Biais approché : $0{,}1333 \times (-0{,}6389)/3{,}6 = -0{,}0237$, pour un biais exact de $-0{,}0255$. Variance approchée : $0{,}1333 \times 0{,}3025 = 0{,}0403$, pour 0,0452. L'approximation de la variance s'améliore (89 % de la valeur exacte, contre 84 % pour $n = 2$) ; celle du biais reste du même ordre (93 %, contre 96 %). Avec si peu d'unités, ces formules ne donnent que des ordres de grandeur.

Correction — Exercice 3

(a) $\sum x_k = 40$ et $\sum y_k = 194$, d'où $\bar{x} = 4$, $\bar{y} = 19{,}4$ et $\widehat{R} = 19{,}4/4 = 4{,}85$ centaines de dh, soit 485 dh par personne et par mois. On n'a pas besoin de X : $\widehat{R}$ est le rapport de deux moyennes de l'échantillon.
(b) Résidus $\hat{e}_k = y_k - 4{,}85\,x_k$: 2,3 ; 0,45 ; 0,6 ; −1,4 ; −2,25 ; −1,1 ; 2,45 ; −0,25 ; −3,95 ; 3,15. Leur somme vaut $\sum y_k - \widehat{R}\sum x_k = 194 - 4{,}85 \times 40 = 0$. Somme des carrés 45,675, d'où $s_e^2 = 45{,}675/9 = 5{,}075$. Avec $f = 10/400 = 0{,}025$:

$$\widehat{V}(\widehat{R}) = \frac{1-f}{n} \cdot \frac{s_e^2}{\bar{x}^2} = \frac{0{,}975}{10} \times \frac{5{,}075}{16} = 0{,}0309,$$

erreur-type 0,176. Intervalle à 95 % : $4{,}85 \pm 0{,}345$, soit [4,505 ; 5,195] centaines de dh, de 450 à 520 dh par personne environ. Avec dix ménages seulement, l'approximation normale reste fragile.
(c) Les rapports y_k/x_k valent 6 ; 5 ; 5 ; 4,5 ; 4,4 ; 4,667 ; 5,667 ; 4,8 ; 4,286 ; 8, de moyenne 5,232. Cette moyenne compte autant le ménage d'une personne (8 par personne) que celui de sept personnes (4,29) ; or les petits ménages dépensent plus par personne. $\widehat{R} = 194/40$ compte chaque **personne** une fois : c'est lui qui estime $R = Y/X$.
(d) Quatre ménages ruraux (3, 5, 6 et 9) : $\bar{y}_d = (20 + 22 + 28 + 30)/4 = 25$ centaines de dh, soit 2 500 dh.

- Linéarisation : $e_k = \delta_k(y_k - 25)$ vaut −5, −3, 3, 5 pour les ruraux et 0 pour les six autres ménages ; somme des carrés 68, d'où $s_e^2 = 68/9 = 7{,}556$ et, avec $n_d/n = 0{,}4$, $\widehat{V}(\bar{y}_d) = \dfrac{0{,}975}{10} \times \dfrac{7{,}556}{0{,}4^2} = 4{,}604$: erreur-type 2,146, soit 215 dh.
- Chapitre 4 : $s_d^2 = 68/3 = 22{,}667$ et $(1-f)\,s_d^2/n_d = 0{,}975 \times 22{,}667/4 = 5{,}525$: erreur-type 2,351, soit 235 dh.

Le rapport des variances vaut $\dfrac{(n_d - 1)\,n}{n_d\,(n-1)} = \dfrac{3 \times 10}{4 \times 9} = 0{,}833$. Avec quatre ruraux, les deux approches diffèrent nettement ; elles se rejoignent quand le domaine compte quelques dizaines d'unités (0,984 pour les 48 ruraux de l'échantillon fil rouge).

2 Exercice cumulatif (chapitres 12, 11, 10 et 9)

Correction — Exercice 4

(a) Un commerce du quartier i est tiré si son quartier l'est (4/30), puis s'il est tiré dans son quartier $(10/M_i)$: $\pi_k = (4/30) \times (10/M_i)$ (chapitre 10).

Quartier tiré	1	2	3	4
π_k	0,01667	0,02963	0,02222	0,03810
Poids d_k	60	33,75	45	26,25

Somme des poids des 40 commerces : $10 \times (60 + 33{,}75 + 45 + 26{,}25) = 1\,650$, et non 1 500. La somme des poids est aléatoire : les quartiers tirés comptent en moyenne $220/4 = 55$ commerces, contre $1\,500/30 = 50$ pour l'ensemble.
(b) Totaux estimés des quartiers $M_i\,\bar{y}_i$: 3 360 ; 1 350 ; 2 160 ; 875, de somme 7 745 milliers de dh. $\widehat{Y} = (30/4) \times 7\,745 = 58\,087{,}5$ milliers de dh, soit 58,1 millions de dh par mois, et $\widehat{Y}/N = 38{,}73$ milliers de dh par commerce.
(c) $\widehat{N} = 7{,}5 \times 220 = 1\,650$ commerces. Ratio : $\widehat{Y}/\widehat{N} = 7\,745/220 = 35{,}20$ milliers de dh, et total $1\,500 \times 35{,}2045 = 52\,806{,}8$ milliers de dh, soit 52,8 millions. L'estimateur sans biais a tiré des quartiers plus grands que la moyenne : il surestime à la fois le chiffre d'affaires total et le nombre de commerces (1 650 au lieu de 1 500). Le ratio divise l'un par l'autre et corrige par le facteur $N/\widehat{N} = 0{,}909$: c'est l'estimateur par le ratio sur la variable auxiliaire M_i, de total N connu (chapitres 9 et 12).
(d) Résidus $M_i\,\bar{y}_i - 35{,}2045\,M_i$: 543,64 ; −234,20 ; 47,73 ; −357,16 ; $s_e^2 = 160\,077{,}6$. Avec $f_I = 4/30$ et $\overline{M} = 50$:

$$\widehat{V} = \frac{1 - 4/30}{4 \times 50^2} \times 160\,077{,}6 = 13{,}87,$$

erreur-type 3,72 milliers de dh, soit un coefficient de variation de 10,6 %. Négliger le second degré sous-estime un peu la variance ; et quatre quartiers seulement donnent une estimation de variance très fragile.
(e) Avec $p_i = M_i/1\,500$ et quatre quartiers, puis 10 commerces par quartier :

$$\pi_k = \frac{4\,M_i}{1\,500} \times \frac{10}{M_i} = \frac{40}{1\,500} = 0{,}02667$$

pour tous les commerces, soit un poids de 37,5 (chapitre 11). Le plan est **autopondéré** : la somme des poids vaut exactement $40 \times 37{,}5 = 1\,500 = N$, et la moyenne s'estime par la moyenne simple des 40 commerces, $(42 + 30 + 36 + 25)/4 = 33{,}25$ milliers de dh. Point commun avec le ratio : la taille des quartiers tirés ne déforme plus l'estimation. Le tirage proportionnel à la taille en tient compte **au tirage**, le ratio **à l'estimation**.

3 Applications en sciences économiques

Correction — Exercice 5

(a) $f = 40/800 = 0{,}05$ et $(1-f)/n = 0{,}02375$. $N\bar{y} = 800 \times 9{,}2 = 7\,360$ t ; erreur-type $800 \times \sqrt{0{,}02375 \times 5{,}5^2} = 800 \times 0{,}8476 = 678{,}1$ t ; intervalle $7\,360 \pm 1\,329$, soit [6 031 ; 8 689] t (± 18 %).
(b) $\widehat{R} = 9{,}2/4{,}6 = 2$ t par hectare, et $X\widehat{R} = 4\,000 \times 2 = 8\,000$ t. Avec $s_{xy} = 0{,}88 \times 3 \times 5{,}5 = 14{,}52$:

$$s_e^2 = 5{,}5^2 - 2 \times 2 \times 14{,}52 + 2^2 \times 3^2 = 30{,}25 - 58{,}08 + 36 = 8{,}17$$

Erreur-type $800 \times \sqrt{0{,}02375 \times 8{,}17} = 800 \times 0{,}4405 = 352{,}4$ t ; intervalle $8\,000 \pm 690{,}7$, soit [7 309 ; 8 691] t ($\pm 8{,}6$ %).
(c) Effet de plan estimé : $8{,}17/30{,}25 = 0{,}270$. $CV_x = 3/4{,}6 = 0{,}652$ et $CV_y = 5{,}5/9{,}2 = 0{,}598$: seuil $0{,}652/(2 \times 0{,}598) = 0{,}545$, largement dépassé par $r = 0{,}88$.
(d) Erreur-type du rendement : $0{,}4405/4{,}6 = 0{,}0958$ t/ha ; intervalle $2 \pm 0{,}188$, soit [1,81 ; 2,19] t/ha.
(e) Les exploitations tirées ont une superficie moyenne de 4,6 ha, contre $\bar{X} = 4\,000/800 = 5$ ha : elles sont 8 % plus petites que l'ensemble, et $N\bar{y}$ sous-estime la production. Le ratio multiplie par $5/4{,}6 = 1{,}087$: $7\,360 \times 1{,}087 = 8\,000$ t. On retient le ratio, dont la marge est deux fois plus étroite. Ce choix se fait **avant** l'enquête, sur la foi des enquêtes antérieures : choisir après coup l'estimation qui arrange fausserait les marges d'erreur annoncées.

Correction — Exercice 6

(a) $\widehat{R} = 1{,}32/1{,}25 = 1{,}056$: une hausse estimée de 5,6 %. Avec $s_{xy} = 0{,}95 \times 0{,}8 \times 0{,}9 = 0{,}684$:

$$s_e^2 = 0{,}9^2 - 2 \times 1{,}056 \times 0{,}684 + 1{,}056^2 \times 0{,}8^2 = 0{,}0791$$

$f = 0{,}1$ et $(1-f)/n = 0{,}015$: erreur-type de $\widehat{R}$, $\sqrt{0{,}015 \times 0{,}0791}/1{,}25 = 0{,}0276$; intervalle $1{,}056 \pm 0{,}054$, soit [1,002 ; 1,110]. L'intervalle exclut 1, de justesse : la hausse est établie, mais elle peut aller de 0,2 % à 11 %.
(b) $N\bar{y} = 600 \times 1{,}32 = 792$ millions de dh ; erreur-type $600 \times \sqrt{0{,}015 \times 0{,}81} = 600 \times 0{,}1102 = 66{,}1$; intervalle [662,4 ; 921,6]. Ratio : $X\widehat{R} = 720 \times 1{,}056 = 760{,}3$ millions de dh ; erreur-type $600 \times 0{,}03444 = 20{,}7$; intervalle [719,8 ; 800,8].
(c) $N\bar{x} = 600 \times 1{,}25 = 750$ millions de dh, contre $X = 720$: les entreprises tirées réalisaient en 2024 un chiffre d'affaires 4,2 % plus élevé que la moyenne. L'échantillon compte, par hasard, un peu trop de grandes entreprises, et $N\bar{y}$ surestime le total de 2025. Le ratio applique l'évolution observée au total connu : $X\widehat{R} = N\bar{y} \times X/(N\bar{x}) = 792 \times 0{,}96 = 760{,}3$.
(d) Effet de plan estimé : $0{,}0791/0{,}81 = 0{,}098$: l'erreur-type est divisée par 3,2. Le chiffre d'affaires d'une année est presque proportionnel à celui de l'année précédente ($r = 0{,}95$, pour un seuil de $0{,}64/(2 \times 0{,}682) = 0{,}469$) : les écarts de taille entre entreprises, qui font l'essentiel de la variance de $\bar{y}$, disparaissent du rapport $\bar{y}/\bar{x}$.