

Méthodes statistiques Solutions Chapitre 9

Problème 9.1

Tableau des résultats

a)		b)		c)	
Estimation par la moyenne	Total estimé	Estimation par la différence	Total estimé	Estimation par le quotient	Total estimé
$T = N\bar{y}$	13 90101	$T_{y_d} = N\hat{\mu}_{y_d}$	136 940	$T_{y_q} = N\hat{\mu}_{y_q}$	136 999
$\hat{\sigma}_T = N\hat{\sigma}_{\bar{y}}$	6 459,227	$\hat{\sigma}_{T_{y_d}} = N\hat{\sigma}_{\hat{\mu}_{y_d}}$	3 301,12	$\hat{\sigma}_{T_{y_q}} = N\hat{\sigma}_{\hat{\mu}_{y_q}}$	3 294,61
Limite inférieure	126 092	Limite inférieure	130 338	Limite inférieure	130 410
Limite supérieure	151 929	Limite supérieure	143 543	Limite supérieure	143 588

Problème 9.2

$$\hat{R} = 0,974326$$

Tableau des résultats

	Moyenne	Différence	Quotient
Estimation	309 813	383 036	381 156
Écart-type de l'estimateur	7 870,30	6 480,99	6 296,80
Limite inférieure	294 073	370 074	368 563
Limite supérieure	325 554	395 998	393 750

Problème 9.3

	Estimation par		
	la moyenne	la différence	le quotient
Estimation	2 335	2 404	2 441
Écart-type	277,56	208,32	241,79
Limite inférieure	1 780	1 987	1 958
Limite supérieure	2 890	2 821	2 925

Problème 9.4

On trouve $\hat{\sigma}_{\hat{\mu}_{y_q}} = 32,40$ et $\hat{\sigma}_{\hat{\mu}_{y_d}} = 9545$. L'estimateur par la différence est tout à fait inacceptable. En revanche, si on divise tous les x par 1 000, on trouve alors $\hat{\sigma}_{\hat{\mu}_{y_q}} = 32,40$ et $\hat{\sigma}_{\hat{\mu}_{y_d}} = 30,69$. L'estimation par la différence est légèrement meilleure cette fois-ci. On remarque que l'estimateur par le quotient n'est pas affecté par l'unité de mesure alors que l'estimateur par la différence l'est. Ce défaut de l'estimateur par la différence invite à la prudence. C'est une raison pour laquelle il vaut mieux limiter son utilisation aux cas où x et y s'expriment dans les mêmes unités.

Problème 9.5

- a)
 - (i) On utiliserait alors l'estimation par le quotient, en particulier si dans la région la plupart des agriculteurs cultivaient du blé.
 - (ii) J'utiliserais l'estimation par la moyenne.
 - (iii) Il faudrait employer l'estimation par la différence ou par le quotient, de préférence ce dernier si l'échantillon est assez grand.
- b) Une estimation par la différence devrait se révéler efficace.
- c) La corrélation entre le salaire et le revenu du ménage est probablement trop faible pour que la variable auxiliaire soit utile. À moins que la population comprend une forte proportion de ménages à un seul revenu, auquel cas on pourra employer l'estimateur par la différence (ou peut-être par le quotient).
- d) On utilise l'estimateur par le quotient si les commerces sont de même nature – disons, tous des épiceries. Sinon, on utilisera une estimation par la moyenne.
- e) Si les commerces sont de tailles très variables, la corrélation sera forte, et l'estimation par le quotient sera avantageuse.
- f) La variable auxiliaire ne se révélera probablement pas très utile.
- g) On a intérêt à utiliser soit l'estimateur par la différence, soit l'estimateur par le quotient.

Problème 9.6

Voici quelques données brutes

μ_y	54,875	R	1,876 068
μ_x	29,25	Moyenne des estimations	55,592 76
S_y	33,193 96	Écart-type calculé des estimations	5,087 095
S_x	19,091 88	Écart-type réel	6,172 028
S_{xy}	602,464 3		

- a) L'estimateur est biaisé. Le biais est de 1,3%.
- b) L'écart-type réel est 6,172. La formule a donc sous-estimé l'écart-type.
- c) Voici les écarts-types : $\sigma_{\bar{y}} = 15,1509$, $\sigma_{\hat{\mu}_{\text{est}}} = 7,3797$. Les estimateurs par la différence et par le quotient sont comparables.
- d) Résumé des réponses :

	Estimation par la moyenne	Estimation par le quotient
Probabilité de se tromper de plus de 10%	0,6964	0,3571
Probabilité de se tromper de plus de 20%	0,5357	0,536

- e) L'espérance de $\hat{\sigma}_{\hat{\mu}_{\text{est}}}^2$ est 26,618 alors que variance réelle est 38,094, un biais de 31 %.

Problème 9.7

- a) On estime la proportion d'adultes à 0,709; l'écart-type de l'estimateur est estimé à 0,0343. Intervalle de confiance: le pourcentage d'adultes est entre 64 % et 78 %.
- b) On estime la proportion d'adultes dans les maisons à 0,6378; l'écart-type de l'estimateur est estimé à 0,0378. Intervalle de confiance: le pourcentage d'adultes dans les maisons est entre 56 % et 71 %.
- c) On estime le nombre d'adultes à 3810; l'écart-type de l'estimateur est estimé à 184. Intervalle de confiance: le nombre d'adultes est entre 3442 et 4178.
- d) On estime la proportion des habitations qui sont des maisons sans enfants à 2430; l'écart-type de l'estimateur est estimé à 271,45. Intervalle de confiance: le nombre d'adultes se situe entre 1887 et 2973.
- e) Il s'agit d'une proportion. $\hat{p} = 0,12$ et $\hat{\sigma}_{\hat{p}} = 0,0456428$. Intervalle de confiance: entre 2,9 % et 21,1 %.
- f) On estime la proportion de maisons sans enfants à 0,1875; l'écart-type de l'estimateur est estimé à 0,06852715. Intervalle de confiance: la proportion de maisons sans enfants se situe entre 5,0 % et 32,5 %.