Intervalles de Confiance

1 Préambule

$\mathscr{L}$ est une loi de probabilité dépendant d'un paramètre θ ; $X_1, \ldots, X_n$ ($n \geq 2$) est un échantillon aléatoire de $\mathscr{L}$; B_1 et B_2 sont deux statistiques de cet échantillon.

L'intervalle $[B_1; B_2]$ est un intervalle de confiance du paramètre θ au niveau $1-\alpha$ ($0<\alpha<1$) s'il contient θ avec probabilité $1-\alpha$; on note alors : $IC_{1-\alpha}(\theta) = [B_1; B_2]$.[a]

Les intervalles de confiance d'une moyenne, d'une proportion ou d'une variance donnés dans les sections suivantes reposent sur l'estimateur de la moyenne : $\bar{X} = \left(\sum_{i=1}^n X_i\right)/n$ et sur l'estimateur sans biais de la variance : $S'^2 = \left(\sum_{i=1}^n (X_i - \bar{X})^2\right)/(n-1)$.

z_β désigne le quantile d'ordre β ($0<\beta<1$) de la loi normale centrée réduite, $t_{\nu;\beta}$ ($\nu \in \mathbb{N} - \{0\}$) le quantile d'ordre β de la loi de Student à ν degrés de liberté et $\chi^2_{\nu;\beta}$ le quantile d'ordre β de la loi de χ^2 à ν degrés de liberté.

2 Intervalle de confiance d'une moyenne

$\mathscr{L}$ est une loi de probabilité de moyenne μ ; on cherche un intervalle de confiance de μ au niveau $1-\alpha$ reposant sur un échantillon $X_1, \ldots, X_n$ de $\mathscr{L}$. On distingue traditionnellement quatre cas.

□ *$\mathscr{L}$ est une loi normale de variance σ^2 connue et la taille n de l'échantillon est quelconque*

Un intervalle de confiance de μ au niveau $1-\alpha$ est :

$$IC_{1-\alpha}(\mu) = \bar{X} \pm z_{1-\alpha/2} \times \sigma/\sqrt{n} \tag{1}$$

□ *$\mathscr{L}$ est une loi normale de variance σ^2 inconnue et la taille n de l'échantillon est quelconque*

Un intervalle de confiance de μ au niveau $1-\alpha$ est :

$$IC_{1-\alpha}(\mu) = \bar{X} \pm t_{n-1;1-\alpha/2} \times S'/\sqrt{n} \tag{2}$$

□ *$\mathscr{L}$ est une loi inconnue de variance σ^2 connue et la taille n de l'échantillon est grande ($n \geq 30$)*

Un intervalle de confiance de μ au niveau $1-\alpha$ est :

$$IC_{1-\alpha}(\mu) = \bar{X} \pm z_{1-\alpha/2} \times \sigma/\sqrt{n} \tag{3}$$

□ *$\mathscr{L}$ est une loi inconnue de variance σ^2 inconnue et la taille n de l'échantillon est grande ($n \geq 30$)*

Un intervalle de confiance de μ au niveau $1-\alpha$ est :

$$IC_{1-\alpha}(\mu) = \bar{X} \pm z_{1-\alpha/2} \times S'/\sqrt{n} \tag{4}$$

Remarque. (1) et (3) sont identiques mais elles ne reposent pas sur la même justification. (1) est une conséquence de l'additivité de la loi normale et l'intervalle de confiance cité est exact. (3) résulte du théorème *central limit* et l'intervalle de confiance cité est approximatif.

3 Intervalle de confiance d'une proportion

$\mathscr{L}$ est une loi de Bernoulli de paramètre p ; on cherche un intervalle de confiance de p au niveau $1-\alpha$ reposant sur un échantillon $X_1, \ldots, X_n$ de $\mathscr{L}$. Dans ce contexte, la statistique $\bar{X}$ représente la proportion (ou la fréquence) de succès dans l'échantillon, on la note F et pour $n \geq 30$ un intervalle de confiance de p au niveau $1-\alpha$ est approximativement :

$$IC_{1-\alpha}(p) = F \pm z_{1-\alpha/2} \times \sqrt{F(1-F)/n} \tag{5}$$

A. Lourme, Faculté d'économie, gestion & AES, Université de Bordeaux `http://alexandrelourme.free.fr`

[a] Les bornes d'un intervalle de confiance sont aléatoires mais le paramètre visé ne l'est pas.

4 Intervalle de confiance d'une variance

$\mathscr{L}$ est une loi normale de variance σ^2 ; on cherche un intervalle de confiance de σ^2 au niveau $1-\alpha$ reposant sur un échantillon $X_1, \ldots, X_n$ de $\mathscr{L}$.

Un intervalle de confiance de σ^2 au niveau $1-\alpha$ est :

$$IC_{1-\alpha}(\sigma^2) = \left[(n-1)S'^2/\chi^2_{n-1;1-\alpha/2} ; (n-1)S'^2/\chi^2_{n-1;\alpha/2}\right] \tag{6}$$

5 Résumé

paramètre	*loi* $\mathscr{L}$	*contexte*	*intervalle de confiance*
moyenne μ	normale	σ^2 connu	$IC_{1-\alpha}(\mu) = \bar{X} \pm z_{1-\alpha/2} \times \sigma/\sqrt{n}$
	normale	σ^2 inconnu	$IC_{1-\alpha}(\mu) = \bar{X} \pm t_{n-1;1-\alpha/2} \times S'/\sqrt{n}$
	inconnue	σ^2 connue et $n \geq 30$	$IC_{1-\alpha}(\mu) = \bar{X} \pm z_{1-\alpha/2} \times \sigma/\sqrt{n}$
	inconnue	σ^2 inconnue et $n \geq 30$	$IC_{1-\alpha}(\mu) = \bar{X} \pm z_{1-\alpha/2} \times S'/\sqrt{n}$
proportion p	Bernoulli	n grand ($n \geq 30$)	$IC_{1-\alpha}(p) = F \pm z_{1-\alpha/2} \times \sqrt{F(1-F)/n}$
variance σ^2	normale	-	$IC_{1-\alpha}(\sigma^2) = \left[(n-1)S'^2/\chi^2_{n-1;1-\alpha/2} ; (n-1)S'^2/\chi^2_{n-1;\alpha/2}\right]$

Table 1: *Tableau synoptique pour l'estimation par intervalle de confiance d'une moyenne, d'une proportion, d'une variance.*

6 Exercices Type

Exercice 0.

Nommez, interprétez et déterminez la valeur des quantiles suivants : $z_{0,95}$, $t_{15;0,95}$, $\chi^2_{10;0,90}$.

Exercice 1.

La masse moyenne de seize américains choisis au hasard est de 78 kg. Déterminez un intervalle de confiance de la masse moyenne des américains au niveau 95% en supposant que la masse des américains a pour variance 2500 kg^2.

Exercice 2.

La masse moyenne de seize américains choisis au hasard est de 78 kg et la variance de leurs masses vaut 2500 kg^2. Déterminez un intervalle de confiance de la masse moyenne des américains au niveau 95%.

Exercice 3.

La taille moyenne de trente-six français choisis au hasard est de 175 cm. Déterminez un intervalle de confiance de la taille moyenne des français au niveau 90% en supposant que la variance de leurs taille vaut 100 cm^2.

Exercice 4.

La taille moyenne de trente-six français choisis au hasard est de 175 cm et la variance de leurs taille vaut 169 cm^2. Déterminez un intervalle de confiance de la taille moyenne des français au niveau 90%.

Exercice 5.

Il y a cent dix partisans de M. Cameron parmi deux cents anglais choisis au hasard ; déterminez un intervalle de confiance au niveau 90% du score de M. Cameron aux prochaines élections.

Exercice 6.

La variance des salaires de cinquante-et-un fonctionnaires suisses est de 10^6 (CHF^2). Déterminez un intervalle de confiance au niveau 90% de la variance des salaires des fonctionnaires suisses.

Exercice 7.

Table 2 donne le flux observé sur le compte de trente clients de la banque Paul & Simon ainsi que leur catégorie de risque (A/B).

client	*flux*	*risque*	*client*	*flux*	*risque*	*client*	*flux*	*risque*
1	19	A	11	21	B	21	16	A
2	14	B	12	16	A	22	15	A
3	18	B	13	16	A	23	19	B
4	20	A	14	12	A	24	21	B
5	12	A	15	14	B	25	12	A
6	11	A	16	20	A	26	22	B
7	29	B	17	19	B	27	28	B
8	42	B	18	17	A	28	19	B
9	11	A	19	10	A	29	17	A
10	18	A	20	22	A	30	15	A

Table 2: *Flux et catégorie de risque d'un échantillon de trente clients de la banque Paul & Simon*

1. Déterminez la valeur observée des statistiques suivantes :
   (a) fréquence du risque A
   (b) moyenne des flux
   (c) variance des flux
2. Donnez un intervalle de confiance au niveau 95% de la proportion de clients A chez Paul & Simon.
3. Déterminez un intervalle de confiance au niveau 95% de la moyenne de tous les flux de cette banque.
4. Déterminez un intervalle de confiance au niveau 95% de la variance des flux de la banque.

Exercice 8.

Montrez que la longeur d'un intervalle de confiance (i) diminue lorsque le risque α augmente et (ii) diminue également lorsque la taille n de l'échnatillon augmente.