Tests d'hypothèses via une **A**pproche **E**xpérimentale des **P**robabilités

CQLS : cqls@upmf-grenoble.fr
http://cqls.upmf-grenoble.fr

1 Introduction et généralités

1.1 Cadre d'estimation à un échantillon

Dans ce cadre, tout problème pratique doit se ramener à l'étude d'une unique variable d'intérêt notée ici Y (pouvant aussi être vue comme une future unique donnée). En pratique, nous disposerons d'un jeu de n données $\boldsymbol{y} = (y_1, \ldots, y_n)$ (i.e. un vecteur ou "paquet" de n observations "indépendantes" de Y) qui peut par conséquent être vu comme un résultat possible d'un futur jeu de n données $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$. Afin d'expliciter le paramètre d'intérêt intimement lié dans les problématiques du cours à la variable d'intérêt Y, nous imaginerons disposer d'une infinité de données virtuelles $y_{[1]}, \ldots, y_{[m]}, \ldots$ dont la notation en indice entre crochet (i.e. "$_{[.]}$") nous rappelle qu'il ne faut pas les confondre avec le jeu des n données $y_1, \ldots, y_n$ qui seront bien réelles. Rappelons aussi que dans ce cours les tailles n des données **réelles** et m des données **virtuelles** ont a priori des ordres de grandeur complètement différents, à savoir n plutôt raisonnablement grand et m aussi grand que possible voire infini.

1.1.1 Paramètres proportion et moyenne

La moyenne notée μ_Y ou plus simplement μ (plutôt appelée espérance de Y dans l'**A**pproche **M**athématique des **P**robabilités et notée $\mathbb{E}(Y)$) s'exprime via l'**A**pproche **E**xpérimentale des **P**robabilités par :

$$\boxed{\overline{(y_{[\cdot]})}_m = \frac{1}{m}\sum_{k=1}^{m} y_{[k]} \simeq \overline{(y_{[\cdot]})}_\infty = \mu_Y = \mathbb{E}(Y).}$$

Soulignons toutefois que si les données sont exclusivement à valeurs 0 ou 1, la moyenne devient une proportion (ou probabilité) et sera notée p plutôt que μ. Rappelons qu'une future estimation $\widehat{\mu_Y}(\boldsymbol{Y})$ de μ_Y est tout simplement $\overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$ (notée aussi $\overline{Y}$).

1.1.2 Paramètre variance

La variance notée σ_Y^2 ou plus simplement σ^2 (conservant la même dénomination dans l'**A.M.P.** et notée $\mathbb{V}\mathrm{ar}(Y)$) s'exprime via l'**A.E.P.** par :

$$\boxed{\left(\overleftrightarrow{(y_{[\cdot]})}_m\right)^2 = \frac{1}{m}\sum_{k=1}^{m}\left(y_{[k]} - \overline{(y_{[\cdot]})}_m\right)^2 \simeq \left(\overleftrightarrow{(y_{[\cdot]})}_\infty\right)^2 = \sigma_Y^2 = \mathbb{V}ar(Y) = \sigma(Y)^2.}$$

Dans le cadre de grands échantillons (voir plus loin), il est plus qu'intéressant de noter que la variance est aussi une moyenne. En effet, nous pouvons écrire $\sigma_Y^2 = \mu_{\ddot{Y}}$ puisque $\mathbb{V}\mathrm{ar}(Y) = \mathbb{E}((Y-\mu_Y)^2) = \mathbb{E}(\ddot{Y})$ où $\ddot{Y} = (Y-\mu_Y)^2$ est le carré de la variable aléatoire Y préalablement centrée. Le vecteur des futures données $((Y_1-\mu_Y)^2, \ldots, (Y_n-\mu_Y)^2)$ étant inaccessible puisque μ_Y est inconnu, nous le remplacerons par $\ddot{\boldsymbol{Y}} = ((Y_1-\overline{Y})^2, \ldots, (Y_n-\overline{Y})^2)$. Ainsi, nous pourrions aussi proposer $\widehat{\mu_{\ddot{Y}}}\left(\ddot{\boldsymbol{Y}}\right)$ comme future estimation de $\sigma_Y^2 = \mu_{\ddot{Y}}$ (plutôt lorsque la taille n des données sera suffisamment grande).

1.2 Cadre d'estimation à deux échantillons (indépendants)

ll y a dans ce cadre deux variables d'intérêts $Y^{(1)}$ et $Y^{(2)}$ ("indépendantes") dont on cherche soit à comparer les moyennes soit les variances à partir de deux échantillons, l'un $\mathbf{y}^{(1)} = (y_1^{(1)}, \ldots, y_{n^{(1)}}^{(1)})$ de taille $n^{(1)}$ et l'autre $\mathbf{y}^{(2)} = (y_1^{(2)}, \ldots, y_{n^{(2)}}^{(2)})$ de taille $n^{(2)}$. Il en découle deux futurs jeux de données $\mathbf{Y}^{(1)} = (Y_1^{(1)}, \ldots, Y_{n^{(1)}}^{(1)})$ et $\mathbf{Y}^{(2)} = (Y_1^{(2)}, \ldots, Y_{n^{(2)}}^{(2)})$. Pour homogénéiser ce cas avec celui à un seul échantillon, nous noterons $\mathbf{Y}$, le vecteur aggrégé de toutes les futures données $\mathbf{Y} = (\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)})$ de taille $n = n^{(1)} + n^{(2)}$. De manière analogue au cas d'un seul échantillon, nous imaginons disposer de deux infinités de données virtuelles, l'une $y_{[1]}^{(1)}, \ldots, y_{[m]}^{(1)}, \ldots$ relative à $Y^{(1)}$ et l'autre $y_{[1]}^{(2)}, \ldots, y_{[m]}^{(2)}, \ldots$ relative à $Y^{(2)}$. Pour $j = 1$ ou $j = 2$, on peut alors exprimer :

- la **moyenne** $\mu_{Y^{(j)}}$ ou plus simplement $\mu^{(j)}$ par $\boxed{\overline{\left(y_{[\cdot]}^{(j)}\right)}_m \simeq \overline{\left(y_{[\cdot]}^{(j)}\right)}_\infty = \mu^{(j)}}$.
- la **variance** $\sigma^2_{Y^{(j)}}$ ou plus simplement $\sigma^2_{(j)}$ définie par $\boxed{\left(\overleftrightarrow{\left(y_{[\cdot]}^{(j)}\right)}_m\right)^2 \simeq \left(\overleftrightarrow{\left(y_{[\cdot]}^{(j)}\right)}_\infty\right)^2 = \sigma^2_{(j)}}$

Nous sommes alors en mesure d'introduire les paramètres servant à comparer respectivement les moyennes et les variances.

- **comparaison de moyennes** s'étudiant soit à partir de la **différence de moyennes** $d_\mu = \mu^{(1)} - \mu^{(2)}$ soit à partir du **rapport de moyennes** $r_\mu = \mu^{(1)}/\mu^{(2)}$ (si $\mu^{(2)} \neq 0$).
- **comparaison de variances** s'étudiant soit à partir de la **différence de variances** $d_{\sigma^2} = \sigma^2_{(1)} - \sigma^2_{(2)}$ soit à partir du **rapport de variances** $r_{\sigma^2} = \sigma^2_{(1)}/\sigma^2_{(2)}$ (si $\sigma^2_{(2)} \neq 0$).

Insistons sur le fait que les utilisations d'une différence ou d'un rapport ne sont pas anodines puisqu'elles permettent de traiter des assertions d'intérêt différentes.

1.3 Les deux cadres usuels : asymptotique et gaussien

→ **Cadre asymptotique ou grand(s) échantillon(s)** : par grand échantillon, on entend dans ce cours une taille de données $n \geq 30$ pour le cas un seul échantillon et des tailles $n^{(1)} \geq 30$ et $n^{(2)} \geq 30$ pour celui de deux échantillons.

→ **Cadre gaussien** : si une variable d'intérêt est supposée suivre une loi Normale on dit que l'échantillon associé est gaussien. Ce cadre d'étude n'est a priori intéressant que s'il est possible de vérifier (éventuellement à partir d'un outil statistique) cette hypothèse de Normalité de la variable d'intérêt. Alors qu'il faudrait disposer d'un grand échantillon pour cette vérification, l'usage dans la littérature statistique est d'utiliser ce cadre d'étude même pour des petits échantillons. Les résultats reposent alors sur la validité de l'a priori que la variable d'intérêt suit une loi Normale. Cependant, certains phénomènes étudiés peuvent laisser penser que cette hypothèse sur la (ou les) variable(s) d'intérêt ne doit pas être aberrante.

1.4 Comparaison entre A.M.P., A.E.P. et Pratique

Dans le tableau suivant, le **jour J** désigne le jour où les données sont réellement récoltées (*Indic* : voir fin de document pour les différentes notations).

Avant le jour J *(θ fixé éventuellement à une valeur arbitraire pour l'expérimentation)*				
Mathématique	$\boldsymbol{Y}$	Y	$\widehat{\theta}(\boldsymbol{Y})$ ou $\widehat{\Theta}$	$t(\boldsymbol{Y})$ ou T
Expérimental	$\boldsymbol{y}_{[1]}$	$\left\{\begin{array}{c} y_{[1]} \\ \vdots \\ y_{[n]} \end{array}\right.$	$\widehat{\theta}(\boldsymbol{y}_{[1]})$ ou $\widehat{\theta}_{[1]}$	$t(\boldsymbol{y}_{[1]})$ ou $t_{[1]}$
	$\boldsymbol{y}_{[2]}$	$\left\{\begin{array}{c} y_{[n+1]} \\ \vdots \\ y_{[2n]} \end{array}\right.$	$\widehat{\theta}(\boldsymbol{y}_{[2]})$ ou $\widehat{\theta}_{[2]}$	$t(\boldsymbol{y}_{[2]})$ ou $t_{[2]}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$\boldsymbol{y}_{[m]}$	$\left\{\begin{array}{c} y_{[(m-1)\times n+1]} \\ \vdots \\ y_{[m\times n]} \end{array}\right.$	$\widehat{\theta}(\boldsymbol{y}_{[m]})$ ou $\widehat{\theta}_{[m]}$	$t(\boldsymbol{y}_{[m]})$ ou $t_{[m]}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Moyenne =		$\mu := \overline{(y_{[\cdot]})}_\infty = \mathbb{E}(Y)$	$\overline{\left(\widehat{\theta}(\boldsymbol{y}_{[\cdot]})\right)}_\infty = \mathbb{E}\left(\widehat{\theta}(\boldsymbol{Y})\right)$	$\overline{\left(t(\boldsymbol{y}_{[\cdot]})\right)}_\infty = \mathbb{E}(t(\boldsymbol{Y}))$
Ecart-Type =		$\sigma := \overleftrightarrow{(y_{[\cdot]})}_\infty$ $= \sigma(Y)$ $= \sqrt{\mathbb{V}ar(Y)}$	$\sigma_{\widehat{\theta}} := \overleftrightarrow{\left(\widehat{\theta}(\boldsymbol{y}_{[\cdot]})\right)}_\infty$ $= \sigma(\widehat{\theta}(\boldsymbol{Y}))$ $= \sqrt{\mathbb{V}ar\left(\widehat{\theta}(\boldsymbol{Y})\right)}$	$\overleftrightarrow{\left(t(\boldsymbol{y}_{[\cdot]})\right)}_\infty = \sigma(t(\boldsymbol{Y}))$ $= \sqrt{\mathbb{V}ar(t(\boldsymbol{Y}))}$
Proportion dans $[a,b[$ =		$\overline{(y_{[\cdot]} \in [a,b[)}_\infty$ $= \mathbb{P}(Y \in [a,b[)$	$\overline{\left(\widehat{\theta}(\boldsymbol{y}_{[\cdot]}) \in [a,b[\right)}_\infty$ $= \mathbb{P}(\widehat{\theta}(\boldsymbol{Y}) \in [a,b[)$	$\overline{\left(t(\boldsymbol{y}_{[\cdot]}) \in [a,b[\right)}_\infty$ $= \mathbb{P}(t(\boldsymbol{Y}) \in [a,b[)$
Histogramme à pas "zéro" =		f_Y	$f_{\widehat{\theta}(\boldsymbol{Y})}$ ou $f_{\widehat{\Theta}}$	$f_{g(\boldsymbol{Y})}$ ou f_T
Surface brique (m fini) =		$\frac{1}{mn}$	$\frac{1}{m}$	$\frac{1}{m}$

Après le jour J *(θ est égal à $\theta^\bullet$ qui est toujours inconnu)*				
Pratique	$\boldsymbol{y}$	$\left\{\begin{array}{c} y_1 \\ \vdots \\ y_n \end{array}\right.$	$\widehat{\theta}(\boldsymbol{y})$ ou $\widehat{\theta}$	$t(\boldsymbol{y})$ ou t

Le **jour J** ($\theta = \theta^\bullet$), si on essaye d'associer des temps de conjugaison aux différents concepts, nous pouvons dire :

→ le **jeu de données réel $\boldsymbol{y}$** représente le *présent*.

→ le **jeu de données aléatoire $\boldsymbol{Y}$** représente le *futur* (on pourra alors aussi l'appeler *futur jeu de données*)

→ les **jeux de données virtuels $\boldsymbol{y}_{[j]}$** représentent le *conditionnel* (ils représentent une infinité de jeux de données que l'on aurait pu avoir à la place de $\boldsymbol{y}$)

2 Test d'hypothèses

De manière générale, la rédaction standard d'un test d'hypothèses s'écrit toujours de la même façon. Elle est décrite ci-dessous pour un paramètre θ qui devra être remplacé par p pour une proportion, μ pour une moyenne, σ^2 pour une variance, d_μ (resp. r_μ) pour une différence (resp. rapport) de moyennes et enfin d_{σ^2} (resp. r_{σ^2}) pour une différence (resp. rapport) de variances. La valeur de référence θ_0 et la loi $\mathcal{L}_0$ devront être adaptée selon la problématique.

Rédaction standard d'un test d'hypothèses paramétrique

Hypothèses de test :

$$\mathbf{H_0} : \theta = \theta_0 \text{ contre } \mathbf{H_1} : \begin{cases} \theta > \theta_0 & (\text{cas } \mathbf{(a)} : \textit{test unilatéral droit}) \\ \theta < \theta_0 & (\text{cas } \mathbf{(b)} : \textit{test unilatéral gauche}) \\ \theta \neq \theta_0 & (\text{cas } \mathbf{(c)} : \textit{test bilatéral}) \end{cases}$$

Statistique de test sous $\mathbf{H_0}$:

$$\widehat{\delta_{\theta,\theta_0}}(\boldsymbol{Y}) \rightsquigarrow \mathcal{L}_0$$

où $\mathcal{L}_0$ est une loi standard à préciser (selon la problématique envisagée).

Règle de décision :

on accepte $\mathbf{H_1}$ si $\begin{cases} \boxed{p - valeur < \alpha} \\ \text{ou de manière équivalente} \\ \boxed{\begin{cases} \widehat{\delta_{\theta,\theta_0}}(\boldsymbol{y}) > \delta^+_{\lim,\alpha} & \mathbf{(a)} \\ \widehat{\delta_{\theta,\theta_0}}(\boldsymbol{y}) < \delta^-_{\lim,\alpha} & \mathbf{(b)} \\ \widehat{\delta_{\theta,\theta_0}}(\boldsymbol{y}) < \delta^-_{\lim,\alpha/2} \text{ ou } \widehat{\delta_{\theta,\theta_0}}(\boldsymbol{y}) > \delta^+_{\lim,\alpha/2} & \mathbf{(c)} \end{cases}} \end{cases}$

où $\delta^-_{\lim,\alpha} = q_\alpha$ et $\delta^+_{\lim,\alpha} = q_{1-\alpha}$ désignent respectivement les quantiles d'ordre α et $1-\alpha$ associés à la loi $\mathcal{L}_0$ et où la p−valeur est définie mathématiquement par :

$$p\text{-}valeur = \begin{cases} \mathbb{P}_{\theta=\theta_0}\left(\widehat{\delta_{\theta,\theta_0}}(\boldsymbol{Y}) > \widehat{\delta_{\theta,\theta_0}}(\boldsymbol{y})\right) & \mathbf{(a)} : \textit{p-valeur droite} \\ \mathbb{P}_{\theta=\theta_0}\left(\widehat{\delta_{\theta,\theta_0}}(\boldsymbol{Y}) < \widehat{\delta_{\theta,\theta_0}}(\boldsymbol{y})\right) & \mathbf{(b)} : \textit{p-valeur gauche} \\ 2\times\min\left(\mathbb{P}_{\theta=\theta_0}\left(\widehat{\delta_{\theta,\theta_0}}(\boldsymbol{Y}) < \widehat{\delta_{\theta,\theta_0}}(\boldsymbol{y})\right), \mathbb{P}_{\theta=\theta_0}\left(\widehat{\delta_{\theta,\theta_0}}(\boldsymbol{Y}) > \widehat{\delta_{\theta,\theta_0}}(\boldsymbol{y})\right)\right) & \mathbf{(c)} : \textit{p-valeur bilatérale} \end{cases}$$

Conclusion : Application de la règle de décision au vu des données $\boldsymbol{y}$.

Propriétés :

1. **La somme des p-valeur gauche et p-valeur droite est égale à 1**
2. **La p-valeur bilatérale est égale à deux fois la plus petite des p-valeurs gauche et droite**

Tableaux récapitulatifs :

Il sera aussi supposé que les données ont été saisies dans le logiciel `R` soit sous le nom `y` (pour un unique échantillon) soit sous les noms `y1` et `y2` (pour deux échantillons indépendants).

θ	$\widehat{\theta}(\boldsymbol{Y})$	$\widehat{\theta}(\boldsymbol{y})$ en R	$\sigma_{\widehat{\theta}}$	$\widehat{\sigma_{\widehat{\theta}}}(\boldsymbol{Y})$	$\widehat{\sigma_{\widehat{\theta}}}(\boldsymbol{y})$ en R
p	$\widehat{p}(\boldsymbol{Y}) = \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$	`mean(y)`	$\sigma_{\widehat{p}} = \sqrt{\frac{p(1-p)}{n}}$	$\sqrt{\frac{\widehat{p}(\boldsymbol{Y})(1-\widehat{p}(\boldsymbol{Y}))}{n}}$	`seMean(y)`
μ	$\widehat{\mu}(\boldsymbol{Y}) = \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$	`mean(y)`	$\sigma_{\widehat{\mu}} = \sqrt{\frac{\sigma^2}{n}}$	$\sqrt{\frac{\widehat{\sigma^2}(\boldsymbol{Y})}{n}}$	`seMean(y)`
σ^2	$\widehat{\sigma^2}(\boldsymbol{Y}) = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$	`var(y)`	$\sigma_{\widehat{\sigma^2}} = \sqrt{\frac{\sigma^2_{\ddot{Y}}}{n}}$	$\sqrt{\frac{\widehat{\sigma^2_{\ddot{Y}}}\left(\ddot{\boldsymbol{Y}}\right)}{n}}$	`seVar(y)`
$d_\mu = \mu^{(1)} - \mu^{(2)}$	$\widehat{d_\mu}(\boldsymbol{Y}) = \widehat{\mu^{(1)}}\left(\boldsymbol{Y^{(1)}}\right) - \widehat{\mu^{(2)}}\left(\boldsymbol{Y^{(2)}}\right)$	`mean(y1)-mean(y2)`	$\sigma_{\widehat{d_\mu}} = \sqrt{\frac{\sigma^2_{(1)}}{n^{(1)}} + \frac{\sigma^2_{(2)}}{n^{(2)}}}$	$\sqrt{\frac{\widehat{\sigma^2_{(1)}}\left(\boldsymbol{Y^{(1)}}\right)}{n^{(1)}} + \frac{\widehat{\sigma^2_{(2)}}\left(\boldsymbol{Y^{(2)}}\right)}{n^{(2)}}}$	`seDMean(y1,y2)`
			Cas Gaussien et $\sigma^2_{(1)} = \sigma^2_{(1)} = \sigma^2$: $\widehat{\sigma^2}(\boldsymbol{Y}) = \frac{(n^{(1)}-1)\widehat{\sigma^2_{(1)}}\left(\boldsymbol{Y^{(1)}}\right) + (n^{(2)}-1)\widehat{\sigma^2_{(2)}}\left(\boldsymbol{Y^{(2)}}\right)}{n^{(1)}+n^{(2)}-2}$		`seDMeanG(y1,y2)`
$d_{\sigma^2} = \sigma^2_{(1)} - \sigma^2_{(2)}$	$\widehat{d_{\sigma^2}}(\boldsymbol{Y}) = \widehat{\sigma^2_{(1)}}\left(\boldsymbol{Y^{(1)}}\right) - \widehat{\sigma^2_{(2)}}\left(\boldsymbol{Y^{(2)}}\right)$	`var(y1)-var(y2)`	$\sigma_{\widehat{d_{\sigma^2}}} = \sqrt{\frac{\sigma^2_{\ddot{Y}^{(1)}}}{n^{(1)}} + \frac{\sigma^2_{\ddot{Y}^{(2)}}}{n^{(2)}}}$	$\sqrt{\frac{\widehat{\sigma^2_{\ddot{Y}^{(1)}}}\left(\ddot{\boldsymbol{Y}}^{(1)}\right)}{n^{(1)}} + \frac{\widehat{\sigma^2_{\ddot{Y}^{(2)}}}\left(\ddot{\boldsymbol{Y}}^{(2)}\right)}{n^{(2)}}}$	`seDVar(y1,y2)`
$r_\mu = \frac{\mu^{(1)}}{\mu^{(2)}}$	$\widehat{r_\mu}(\boldsymbol{Y}) = \frac{\widehat{\mu^{(1)}}\left(\boldsymbol{Y^{(1)}}\right)}{\widehat{\mu^{(2)}}\left(\boldsymbol{Y^{(2)}}\right)}$	`mean(y1)/mean(y2)`	$\sigma_{\widehat{r_\mu}} = \frac{1}{\mu^{(2)}}\sqrt{\frac{\sigma^2_{(1)}}{n^{(1)}} + r^2_\mu \times \frac{\sigma^2_{(2)}}{n^{(2)}}}$	$\frac{1}{\widehat{\mu^{(2)}}\left(\boldsymbol{Y^{(2)}}\right)}\sqrt{\frac{\widehat{\sigma^2_{(1)}}\left(\boldsymbol{Y^{(1)}}\right)}{n^{(1)}} + \widehat{r_\mu}(\boldsymbol{Y})^2 \times \frac{\widehat{\sigma^2_{(2)}}\left(\boldsymbol{Y^{(2)}}\right)}{n^{(2)}}}$	`seRMean(y1,y2)`
$r_{\sigma^2} = \frac{\sigma^2_{(1)}}{\sigma^2_{(2)}}$	$\widehat{r_{\sigma^2}}(\boldsymbol{Y}) = \frac{\widehat{\sigma^2_{(1)}}\left(\boldsymbol{Y^{(1)}}\right)}{\widehat{\sigma^2_{(2)}}\left(\boldsymbol{Y^{(2)}}\right)}$	`var(y1)/var(y2)`	$\sigma_{\widehat{r_{\sigma^2}}} = \frac{1}{\sigma^2_{(2)}}\sqrt{\frac{\sigma^2_{\ddot{Y}^{(1)}}}{n^{(1)}} + r^2_{\sigma^2}\frac{\sigma^2_{\ddot{Y}^{(2)}}}{n^{(2)}}}$	$\frac{1}{\widehat{\sigma^2_{(2)}}\left(\boldsymbol{Y^{(2)}}\right)}\sqrt{\frac{\widehat{\sigma^2_{\ddot{Y}^{(1)}}}\left(\ddot{\boldsymbol{Y}}^{(1)}\right)}{n^{(1)}} + \widehat{r_{\sigma^2}}(\boldsymbol{Y})^2\frac{\widehat{\sigma^2_{\ddot{Y}^{(1)}}}\left(\ddot{\boldsymbol{Y}}^{(1)}\right)}{n^{(1)}}}$	`seRVar(y1,y2)`

			Cadre Asymptotique		Cadre Gaussien	
θ	θ_0	$\sigma_{\widehat{\theta}}$ sous $\mathbf{H_0}$	$\delta_{\theta,\theta_0} = (\theta - \theta_0)/\sigma_{\widehat{\theta}}$	$\widehat{\delta_{\theta,\theta_0}}(\boldsymbol{Y})$ et sa loi sous $\mathbf{H_0}$	δ_{θ,θ_0}	$\widehat{\delta_{\theta,\theta_0}}(\boldsymbol{Y})$ et sa loi sous $\mathbf{H_0}$
p	p_0	$\sqrt{\frac{p_0(1-p_0)}{n}}$	$\delta_{p,p_0} = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	$\widehat{\delta_{p,p_0}}(\boldsymbol{Y}) = \frac{\widehat{p}(\boldsymbol{Y}) - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$		
μ	μ_0	$\sigma_{\widehat{\mu}}$	$\delta_{\mu,\mu_0} = \frac{\mu - \mu_0}{\sigma_{\widehat{\mu}}}$	$\widehat{\delta_{\mu,\mu_0}}(\boldsymbol{Y}) = \frac{\widehat{\mu}(\boldsymbol{Y}) - \mu_0}{\widehat{\sigma_{\widehat{\mu}}}(\boldsymbol{Y})} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$	$\delta_{\mu,\mu_0} = \frac{\mu - \mu_0}{\sigma_{\widehat{\mu}}}$	$\widehat{\delta_{\mu,\mu_0}}(\boldsymbol{Y}) = \frac{\widehat{\mu}(\boldsymbol{Y}) - \mu_0}{\widehat{\sigma_{\widehat{\mu}}}(\boldsymbol{Y})} \rightsquigarrow \mathcal{S}t(n-1)$
σ^2	σ^2_0	$\sigma_{\widehat{\sigma^2}}$	$\delta_{\sigma^2,\sigma^2_0} = \frac{\sigma^2 - \sigma^2_0}{\sigma_{\widehat{\sigma^2}}}$	$\widehat{\delta_{\sigma^2,\sigma^2_0}}(\boldsymbol{Y}) = \frac{\widehat{\sigma^2}(\boldsymbol{Y}) - \sigma^2_0}{\widehat{\sigma_{\widehat{\sigma^2}}}(\boldsymbol{Y})} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$	$\delta_{\sigma^2,\sigma^2_0} = (n-1)\frac{\sigma^2}{\sigma^2_0}$	$\widehat{\delta_{\sigma^2,\sigma^2_0}}(\boldsymbol{Y}) = (n-1)\frac{\widehat{\sigma^2}(\boldsymbol{Y})}{\sigma^2_0} \rightsquigarrow \chi^2(n-1)$
$d_\mu = \mu^{(1)} - \mu^{(2)}$	d_0	$\sigma_{\widehat{d_\mu}}$	$\delta_{d_\mu,d_0} = \frac{d_\mu - d_0}{\sigma_{\widehat{d_\mu}}}$	$\widehat{\delta_{d_\mu,d_0}}(\boldsymbol{Y}) = \frac{\widehat{d_\mu}(\boldsymbol{Y}) - d_0}{\widehat{\sigma_{\widehat{d_\mu}}}(\boldsymbol{Y})} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$	$\delta_{d_\mu,d_0} = \frac{d_\mu - d_0}{\sigma_{\widehat{d_\mu}}}$	$\widehat{\delta_{d_\mu,d_0}}(\boldsymbol{Y}) = \frac{\widehat{d_\mu}(\boldsymbol{Y}) - d_0}{\widehat{\sigma_{\widehat{d_\mu}}}(\boldsymbol{Y})} \rightsquigarrow \mathcal{S}t(n^{(1)}+n^{(2)}-2)$
$d_{\sigma^2} = \sigma^2_{(1)} - \sigma^2_{(2)}$	d_0	$\sigma_{\widehat{d_{\sigma^2}}}$	$\delta_{d_{\sigma^2},d_0} = \frac{d_{\sigma^2} - d_0}{\sigma_{\widehat{d_{\sigma^2}}}}$	$\widehat{\delta_{d_{\sigma^2},d_0}}(\boldsymbol{Y}) = \frac{\widehat{d_{\sigma^2}}(\boldsymbol{Y}) - d_0}{\widehat{\sigma_{\widehat{d_{\sigma^2}}}}(\boldsymbol{Y})} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$		
$r_\mu = \frac{\mu^{(1)}}{\mu^{(2)}}$	r_0	$\sigma_{\widehat{r_\mu}}$	$\delta_{r_\mu,r_0} = \frac{r_\mu - r_0}{\sigma_{\widehat{r_\mu}}}$	$\widehat{\delta_{r_\mu,r_0}}(\boldsymbol{Y}) = \frac{\widehat{r_\mu}(\boldsymbol{Y}) - r_0}{\widehat{\sigma_{\widehat{r_\mu}}}(\boldsymbol{Y})} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$		
$r_{\sigma^2} = \frac{\sigma^2_{(1)}}{\sigma^2_{(2)}}$	r_0	$\sigma_{\widehat{r_{\sigma^2}}}$	$\delta_{r_{\sigma^2},r_0} = \frac{r_{\sigma^2} - r_0}{\sigma_{\widehat{r_{\sigma^2}}}}$	$\widehat{\delta_{r_{\sigma^2},r_0}}(\boldsymbol{Y}) = \frac{\widehat{r_{\sigma^2}}(\boldsymbol{Y}) - r_0}{\widehat{\sigma_{\widehat{r_{\sigma^2}}}}(\boldsymbol{Y})} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$	$\delta_{r_{\sigma^2},r_0} = \frac{r_{\sigma^2}}{r_0}$	$\widehat{\delta_{r_{\sigma^2},r_0}}(\boldsymbol{Y}) = \frac{\widehat{r_{\sigma^2}}(\boldsymbol{Y})}{r_0} \rightsquigarrow \mathcal{F}(n^{(1)}-1, n^{(2)}-1)$

3 Intervalle de confiance

3.1 Généralités

Le concept d'intervalle de confiance d'un paramètre quelconque θ consiste à proposer un encadrement (ou une "fourchette") représenté par un intervalle de variables aléatoires $\left[\widetilde{\theta}_{\inf}(\boldsymbol{Y}), \widetilde{\theta}_{\sup}(\boldsymbol{Y})\right]$ de sorte que le paramètre d'intérêt θ inconnu ait un niveau de confiance $1-\alpha$ (plutôt élevé si α raisonnablement petit) d'être à l'intérieur de cet intervalle. Mathématiquement cela s'exprime par :

$$\boxed{\mathbb{P}\left(\theta \in \left[\widetilde{\theta}_{\inf}(\boldsymbol{Y}), \widetilde{\theta}_{\sup}(\boldsymbol{Y})\right]\right) = 1-\alpha}$$

Par l'approche expérimentale, si nous pouvions imaginer répéter autant de fois que possible la conception d'intervalles de confiance $\left[\widetilde{\theta}_{\inf}(\boldsymbol{y}_{[1]}), \widetilde{\theta}_{\sup}(\boldsymbol{y}_{[1]})\right], \ldots, \left[\widetilde{\theta}_{\inf}(\boldsymbol{y}_{[m]}), \widetilde{\theta}_{\sup}(\boldsymbol{y}_{[m]})\right], \ldots$ obtenus respectivement sur une infinité de jeux de données virtuels $\mathbf{y}_{[1]}, \ldots, \mathbf{y}_{[m]}, \ldots$, nous constaterions alors qu'il n'y en aurait qu'une proportion $1-\alpha$ qui contiendraient le paramètre d'intérêt θ inconnu. La construction de ces intervalles dépend en général de la caractérisation du comportement aléatoire de la mesure d'écart standardisée $\delta_{\widehat{\theta},\theta}(\mathbf{Y})$ proposée dans toutes les problématiques dans le tableau suivant :

	Cadre Asymptotique	Cadre Gaussien
θ	$\delta_{\widehat{\theta},\theta}(\boldsymbol{Y})$	$\delta_{\widehat{\theta},\theta}(\boldsymbol{Y})$
p	$\delta_{\widehat{p},p}(\boldsymbol{Y}) = \dfrac{\widehat{p}(\boldsymbol{Y}) - p}{\widehat{\sigma_{\widehat{p}}}(\boldsymbol{Y})} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$	
μ	$\delta_{\widehat{\mu},\mu}(\boldsymbol{Y}) = \dfrac{\widehat{\mu}(\boldsymbol{Y}) - \mu}{\widehat{\sigma_{\widehat{\mu}}}(\boldsymbol{Y})} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$	$\delta_{\widehat{\mu},\mu}(\boldsymbol{Y}) = \dfrac{\widehat{\mu}(\boldsymbol{Y}) - \mu}{\widehat{\sigma_{\widehat{\mu}}}(\boldsymbol{Y})} \rightsquigarrow \mathcal{S}t(n-1)$
σ^2	$\delta_{\widehat{\sigma^2},\sigma^2}(\boldsymbol{Y}) = \dfrac{\widehat{\sigma^2}(\boldsymbol{Y}) - \sigma^2}{\widehat{\sigma_{\widehat{\sigma^2}}}(\boldsymbol{Y})} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$	$\delta_{\widehat{\sigma^2},\sigma^2}(\boldsymbol{Y}) = (n-1)\dfrac{\widehat{\sigma^2}(\boldsymbol{Y})}{\sigma^2} \rightsquigarrow \chi^2(n-1)$
$d_\mu = \mu^{(1)} - \mu^{(2)}$	$\delta_{\widehat{d_\mu},d_\mu}(\boldsymbol{Y}) = \dfrac{\widehat{d_\mu}(\boldsymbol{Y}) - d_\mu}{\widehat{\sigma_{\widehat{d_\mu}}}(\boldsymbol{Y})} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$	$\delta_{\widehat{d_\mu},d_\mu}(\boldsymbol{Y}) = \dfrac{\widehat{d_\mu}(\boldsymbol{Y}) - d_\mu}{\widehat{\sigma_{\widehat{d_\mu}}}(\boldsymbol{Y})} \rightsquigarrow \mathcal{S}t(n^{(1)} + n^{(2)} - 2)$
$d_{\sigma^2} = \sigma^2_{(1)} - \sigma^2_{(2)}$	$\delta_{\widehat{d_{\sigma^2}},d_{\sigma^2}}(\boldsymbol{Y}) = \dfrac{\widehat{d_{\sigma^2}}(\boldsymbol{Y}) - d_{\sigma^2}}{\widehat{\sigma_{\widehat{d_{\sigma^2}}}}(\boldsymbol{Y})} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$	
$r_\mu = \dfrac{\mu^{(1)}}{\mu^{(2)}}$	$\delta_{\widehat{r_\mu},r_\mu}(\boldsymbol{Y}) = \dfrac{\widehat{r_\mu}(\boldsymbol{Y}) - r_\mu}{\widehat{\sigma_{\widehat{r_\mu}}}(\boldsymbol{Y})} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$	
$r_{\sigma^2} = \dfrac{\sigma^2_{(1)}}{\sigma^2_{(2)}}$	$\delta_{\widehat{r_{\sigma^2}},r_{\sigma^2}}(\boldsymbol{Y}) = \dfrac{\widehat{r_{\sigma^2}}(\boldsymbol{Y}) - r_{\sigma^2}}{\widehat{\sigma_{\widehat{r_{\sigma^2}}}}(\boldsymbol{Y})} \overset{approx.}{\rightsquigarrow} \mathcal{N}(0,1)$	$\delta_{\widehat{r_{\sigma^2}},r_{\sigma^2}}(\boldsymbol{Y}) = \dfrac{\widehat{r_{\sigma^2}}(\boldsymbol{Y})}{r_{\sigma^2}} \rightsquigarrow \mathcal{F}(n^{(1)} - 1, n^{(2)} - 1)$

3.2 Cadre asymptotique

Tous les intervalles de confiance relatifs à toutes les problématiques du cours s'obtiennent **dans le cadre de grands échantillons** par la même méthode. Après substitution du paramètre d'intérêt de votre problématique (au choix parmi p, μ, σ^2, d_μ, d_{σ^2}, r_μ et r_{σ^2}) notée ici de manière générale θ, nous allons naturellement utiliser la caractérisation du comportement aléatoire de l'écart entre $\widehat{\theta}(\boldsymbol{Y})$ et θ exprimée via la mesure d'écart standardisée $\delta_{\widehat{\theta},\theta}(\boldsymbol{Y})$ suivant approximativement une loi Normale $\mathcal{N}(0,1)$. Très facilement, nous pouvons affirmer que :

$$1-\alpha \simeq \mathbb{P}\left(\left|\delta_{\widehat{\theta},\theta}(\boldsymbol{Y})\right| < \delta^+_{lim,\frac{\alpha}{2}}\right) = \mathbb{P}\Big(\underbrace{\widehat{\theta}(\boldsymbol{Y}) - \delta^+_{lim,\frac{\alpha}{2}} \times \widehat{\sigma_{\widehat{\theta}}}(\boldsymbol{Y})}_{\widetilde{\theta}_{\inf}(\boldsymbol{Y})} < \theta < \underbrace{\widehat{\theta}(\boldsymbol{Y}) + \delta^+_{lim,\frac{\alpha}{2}} \times \widehat{\sigma_{\widehat{\theta}}}(\boldsymbol{Y})}_{\widetilde{\theta}_{\sup}(\boldsymbol{Y})}\Big)$$

où $\delta^+_{lim,\frac{\alpha}{2}}$ est le quantile d'ordre $1-\frac{\alpha}{2}$ de la loi $\mathcal{N}(0,1)$

3.3 Cadre gaussien

Pour construire un intervalle de confiance dans un cadre gaussien du paramètre θ (au choix μ, σ^2, d_μ ou r_{σ^2}), nous allons naturellement utiliser la caractérisation du comportement aléatoire de

l'écart entre $\widehat{\theta}(\boldsymbol{Y})$ et θ exprimée via la mesure d'écart standardisée $\delta_{\widehat{\theta},\theta}(\boldsymbol{Y})$. Il s'agit alors de trouver $\widetilde{\theta}_{\text{inf}}(\boldsymbol{Y})$ et $\widetilde{\theta}_{\text{sup}}(\boldsymbol{Y})$ tels que

$$1-\alpha = \mathbb{P}\left(\widetilde{\theta}_{\text{inf}}(\boldsymbol{Y}) < \theta < \widetilde{\theta}_{\text{sup}}(\boldsymbol{Y})\right) \qquad \text{en utilisant le fait que} \qquad 1-\alpha = \mathbb{P}\left(q_{\frac{\alpha}{2}} < \delta_{\widehat{\theta},\theta}(\boldsymbol{Y}) < q_{1-\frac{\alpha}{2}}\right)$$

où $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1-\frac{\alpha}{2}$ de la loi de la mesure d'écart standardisée $\delta_{\widehat{\theta},\theta}(\boldsymbol{Y})$. L'exercice est plus difficile que dans le cadre asymptotique d'une part parce que la mesure d'écart standardisée $\delta_{\widehat{\theta},\theta}(\boldsymbol{Y})$ (θ étant au choix μ, σ^2, d_μ ou r_{σ^2}) ne se décline pas toujours sur le même schéma de construction et d'autre part parce que la loi de $\delta_{\widehat{\theta},\theta}(\boldsymbol{Y})$ n'est plus une loi Normale standard. Sans trop nous attarder, voici les différents intervalles de confiance pour les différents choix de θ :

- $\theta = \mu$:

$$\widetilde{\mu}_{\text{inf}}(\boldsymbol{Y}) = \widehat{\mu}(\boldsymbol{Y}) - q_{1-\frac{\alpha}{2}} \times \sqrt{\frac{\widehat{\sigma^2}(\boldsymbol{Y})}{n}} \quad \text{et} \quad \widetilde{\mu}_{\text{sup}}(\boldsymbol{Y}) = \widehat{\mu}(\boldsymbol{Y}) + q_{1-\frac{\alpha}{2}} \times \sqrt{\frac{\widehat{\sigma^2}(\boldsymbol{Y})}{n}}$$

  où $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1-\frac{\alpha}{2}$ de la loi $\mathcal{S}t(n-1)$.
- $\theta = \sigma^2$:

$$\widetilde{\sigma^2}_{\text{inf}}(\boldsymbol{Y}) = \widehat{\sigma^2}(\boldsymbol{Y}) \times \frac{n-1}{q_{1-\frac{\alpha}{2}}} \quad \text{et} \quad \widetilde{\sigma^2}_{\text{sup}}(\boldsymbol{Y}) = \widehat{\sigma^2}(\boldsymbol{Y}) \times \frac{n-1}{q_{\frac{\alpha}{2}}}$$

  où $q_{1-\frac{\alpha}{2}}$ (resp. $q_{\frac{\alpha}{2}}$) est le quantile d'ordre $1-\frac{\alpha}{2}$ (resp. $\frac{\alpha}{2}$) de la loi $\chi^2(n-1)$.
- $\theta = d_\mu$:

$$\widetilde{d_\mu}_{\text{inf}}(\boldsymbol{Y}) = \widehat{d_\mu}(\boldsymbol{Y}) - q_{1-\frac{\alpha}{2}} \times \sqrt{\widehat{\sigma^2}(\boldsymbol{Y})\left(\frac{1}{n^{(1)}} + \frac{1}{n^{(2)}}\right)}$$

  et

$$\widetilde{d_\mu}_{\text{sup}}(\boldsymbol{Y}) = \widehat{d_\mu}(\boldsymbol{Y}) + q_{1-\frac{\alpha}{2}} \times \sqrt{\widehat{\sigma^2}(\boldsymbol{Y})\left(\frac{1}{n^{(1)}} + \frac{1}{n^{(2)}}\right)}$$

  où $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1-\frac{\alpha}{2}$ de la loi $\mathcal{S}t(n^{(1)}+n^{(2)}-2)$. La quantité $\widehat{\sigma^2}(\boldsymbol{Y})$ est définie comme dans le tableau récapitulatif des tests d'hypothèses (partie cadre gaussien).
- $\theta = r_{\sigma^2}$:

$$\widetilde{r_{\sigma^2}}_{\text{inf}}(\boldsymbol{Y}) = \widehat{r_{\sigma^2}}(\boldsymbol{Y}) \times \frac{1}{q_{1-\frac{\alpha}{2}}} \quad \text{et} \quad \widetilde{r_{\sigma^2}}_{\text{sup}}(\boldsymbol{Y}) = \widehat{r_{\sigma^2}}(\boldsymbol{Y}) \times \frac{1}{q_{\frac{\alpha}{2}}}$$

  où $q_{1-\frac{\alpha}{2}}$ (resp. $q_{\frac{\alpha}{2}}$) est le quantile d'ordre $1-\frac{\alpha}{2}$ (resp. $\frac{\alpha}{2}$) de la loi $\mathcal{F}\left(n^{(1)}-1, n^{(2)}-1\right)$.

4 Langage mathématique et Systèmes de notation

- Dans ce cours, deux systèmes de notation sont utilisés pour décrire des expressions mathématiques dédiées à la statistique. Le premier, appelé *Norme CQLS* (ou *Norme CQLS Standard*) consiste en un système de notation riche (et peut-être un peu lourde) dont le principal avantage est qu'il est taillé sur mesure pour être traduisible dans le langage littéral. Le deuxième système, appelé *Norme **SSE*** (ou *Norme CQLS Simplifié*), a pour vocation à être **S**imple, **S**ynthétique et **E**xplicite (ou du moins le plus possible). Il demande cependant dans son utilisation un meilleur niveau d'expertise essentiellement dû au fait que sa traduction dans le langage littéral est moins explicite que celle pour la *Norme CQLS*.
- Notre conseil est de commencer par l'utilisation de la *Norme CQLS* pour, au fur et à mesure du cours, passer à la *Norme SSE*.
- Conventions communes aux deux Normes CQLS et SSE :
  1. Majuscule versus Minuscule : une variable aléatoire (ou susceptible de l'être) est notée en majuscule quand une variable dont on sait qu'elle est déterministe (i.e. non aléatoire) est noté en minuscule.
  2. Le Chapeau au dessus d'une quantité (par exemple, $\widehat{\theta}$) désigne généralement un remplaçant appelé plus communément estimation dans le cas où la quantité est un paramètre (ici θ).
  3. Un vecteur est noté en **caractères gras**.
     Remarque : une expression écrite sur un document imprimé en caractères gras (ex : "**expression en gras**") est subtituée sur un tableau ou sur une feuille papier par sa version soulignée (ex : "<u>expression en gras</u>").

4. "Delta" (δ en minuscule et Δ en majuscule) est utilisé pour désigner un écart le plus souvent additif (i.e. une soustraction) mais parfois multiplicatif (i.e. une division).

– La *Norme CQLS* a été introduite pour décrire le plus précisément possible l'Approche Expérimentale des Probabilités (A.E.P.). L'A.E.P. s'articulant sur une distinction des différents jeux de données, la *Norme CQLS* repose sur la convention suivante : Toute statistique (i.e. v.a. dépendant d'un jeu de données) s'écrit comme une fonction du jeu de données.
– Il n'y a pas vraiment de convention propre à la *Norme SSE*. Son objectif est cependant de ne pas respecter la convention spécifique (ci-dessus) à la *Norme CQLS* dans le but de rendre plus synthétique les notations mathématiques.
– Le tableau ci-dessous exprime plus clairement la spécificité des *Normes CQLS* et *SSE* en proposant les principales expressions utilisées en statistique dans les 2 normes.

Statistique	Aléatoire ou futur		Réalisé ou présent		Réalisable ou conditionnel	
(v.a. fonction de l'échantillon)	*CQLS*	*SSE*	*CQLS*	*SSE*	*CQLS*	*SSE*
Estimation de θ	$\widehat{\theta^\bullet}(\boldsymbol{Y})$	$\widehat{\Theta}^\bullet$	$\widehat{\theta^\bullet}(\boldsymbol{y})$	$\widehat{\theta}^\bullet$	$\widehat{\theta^\bullet}\left(\boldsymbol{y_{[k]}}\right)$	$\widehat{\theta}^\bullet_{[k]}$
Estimation de $p^\bullet$	$\widehat{p^\bullet}(\boldsymbol{Y})$	$\widehat{P}^\bullet$	$\widehat{p^\bullet}(\boldsymbol{y})$	$\widehat{p}^\bullet$	$\widehat{p^\bullet}\left(\boldsymbol{y_{[k]}}\right)$	$\widehat{p}^\bullet_{[k]}$
Estimation de $\mu^\bullet$	$\widehat{\mu^\bullet}(\boldsymbol{Y})$	$\widehat{M}^\bullet$	$\widehat{\mu^\bullet}(\boldsymbol{y})$	$\widehat{\mu}^\bullet$	$\widehat{\mu^\bullet}\left(\boldsymbol{y_{[k]}}\right)$	$\widehat{\mu}^\bullet_{[k]}$
Estimation de $\sigma^2_\bullet$	$\widehat{\sigma^2_\bullet}(\boldsymbol{Y})$	$\widehat{\Sigma}^2_\bullet$	$\widehat{\sigma^2_\bullet}(\boldsymbol{y})$	$\widehat{\sigma}^2_\bullet$	$\widehat{\sigma^2_\bullet}\left(\boldsymbol{y_{[k]}}\right)$	$\widehat{\sigma}^2_{\bullet,[k]}$
Erreur standard de $\widehat{\theta^\bullet}$	$\widehat{\sigma_{\widehat{\theta^\bullet}}}(\boldsymbol{Y})$	$\widehat{\Sigma}_{\theta^\bullet}$	$\widehat{\sigma_{\widehat{\theta^\bullet}}}(\boldsymbol{y})$	$\widehat{\sigma}_{\theta^\bullet}$	$\widehat{\sigma_{\widehat{\theta^\bullet}}}\left(\boldsymbol{y_{[k]}}\right)$	$\widehat{\sigma}_{\theta^\bullet,[k]}$
Ecart entre $\widehat{\theta^\bullet}$ et $\theta^\bullet$	$\delta_{\widehat{\theta^\bullet},\theta^\bullet}(\boldsymbol{Y})$	$\Delta_{\theta^\bullet}$	$\delta_{\widehat{\theta^\bullet},\theta^\bullet}(\boldsymbol{y})$	$\delta_{\theta^\bullet}$	$\delta_{\widehat{\theta^\bullet},\theta^\bullet}(\boldsymbol{y}_{[k]})$	$\delta_{\theta^\bullet,[k]}$
Estimation de $\delta_{\theta^\bullet,\theta_0}$ (ou δ_{θ_0})	$\widehat{\delta_{\theta^\bullet,\theta_0}}(\boldsymbol{Y})$	$\widehat{\Delta}_{\theta_0}$	$\widehat{\delta_{\theta^\bullet,\theta_0}}(\boldsymbol{y})$	$\widehat{\delta}_{\theta_0}$	$\widehat{\delta_{\theta^\bullet,\theta_0}}\left(\boldsymbol{y_{[k]}}\right)$	$\widehat{\delta}_{\theta_0,[k]}$

– Le tableau ci-dessous illustre comment convertir une notation en sa définition littérale ou mathématique pour des concepts de base de la statistique. La conversion dans le langage `R` y est aussi proposée permettant à l'utilisateur de savoir comment obtenir ces quantités en Pratique :

Notation	Définition littérale	Définition mathématique
$\boldsymbol{y}$ ou $(y_\cdot)_n$	Vecteur des réels $y_1, \cdots, y_n$ (y_i est la $i^{ème}$ composante de $\boldsymbol{y}$)	$(y_1, \cdots, y_n)$ (en `R` : `y <- c(`$y_1, \cdots, y_n$`)`)
$\#(\boldsymbol{y})$	Nombre de composantes de $\boldsymbol{y}$	$n \overset{R}{=}$ `length(y)`
$\overline{y}$ ou $\overline{(y_\cdot)_n}$	Moyenne (empirique) de $\boldsymbol{y}$	$\frac{1}{n}\sum_{i=1}^{n} y_i \overset{R}{=}$ `mean(y)`
$\overline{y=a}$ ou $\overline{(y_\cdot=a)_n}$	Proportion des $y_1, \cdots, y_n$ égaux à a	$\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{y_i=a} \overset{R}{=}$ `mean(y==a)`
$\overline{a\leq y\leq b}$ ou $\overline{(a\leq y_\cdot\leq b)_n}$	Proportion des $y_1, \cdots, y_n$ dans $[a,b]$ avec $(a\leq b)$	$\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{[a,b]}(y_i) \overset{R}{=}$ `mean(a<= y & y<= b)`
$\overleftrightarrow{y}$ ou $\overleftrightarrow{(y_\cdot)_n}$	Ecart-type (empirique) de $\boldsymbol{y}$	$\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\overline{y})^2} \overset{R}{=}$ `sd(y)`
$(\overleftrightarrow{y})^2$ ou $\left(\overleftrightarrow{(y_\cdot)_n}\right)^2$	Variance (empirique) de $\boldsymbol{y}$	$\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\overline{y})2 \overset{R}{=}$ `var(y)`
$q_\alpha(\boldsymbol{y})$ ou $q_\alpha\big((y_\cdot)_n\big)$	Quantile d'ordre α de `y` $(0<\alpha<1)$	$y_{[\alpha n]+1}$ (n impair) et $\frac{y_{[\alpha n]+1}+y_{[\alpha n]+1}}{2}$ (n pair) $\overset{R}{=}$ `quantile(y,alpha)`

5 Quelques instructions R

Instructions de base par l'exemple : des exemples (commentés) valent (peut-être) mieux que de longs discours !

```
> c(-1,1)                    # Création du vecteur (-1,1)
[1] -1  1
> 4+2*c(-1,0,1)              # Transformation 4+2*x appliqué pour chaque composante de y
[1] 2 4 6
> y<-c(1,3,2,4,7,6)
```

```
> y
[1] 1 3 2 4 7 6
> 4+2*y
[1]  6 10  8 12 18 16
> mean(y)                    # Moyenne de y
[1] 3.833333
> sd(y)                      # Ecart-type de y
[1] 2.316607
> yc <- y-mean(y)            # yc correspond au vecteur y centré
> yc
[1] -2.8333333 -0.8333333 -1.8333333  0.1666667  3.1666667  2.1666667
> mean(yc)                   # Moyenne nulle
[1] -1.480297e-16
> sd(yc)                     # Idem que l'écart-type de y
[1] 2.316607
> ycr <- (y-mean(y))/sd(y)   # ycr correspond au vecteur y centré et réduit
> mean(ycr)                  # Moyenne nulle
[1] -7.40239e-17
> sd(ycr)                    # Ecart-type à 1
[1] 1
> var(y)                     # Variance de y
[1] 5.366667
> sqrt(var(y))               # Ecart-type = racine carrée de variance
[1] 2.316607
> sd(y)^2                    # Variance = carré de l'écart-type
[1] 5.366667
```

Quantiles et fonctions de répartition avec R : Soit p un réel appartenant à $]0,1[$, on définit le quantile d'ordre p associée à une loi de probabilité le réel qui via l'approche expérimentale peut être vu comme le réel qui sépare l'infinité des observations (associée à la loi de probabilité) en deux, une proportion p à gauche et une proportion $1-p$ à droite. On définit également la fonction de répartition en un réel q, la proportion parmi l'infinité des observations qui se situent avant q. Ces deux notions sont illustrées dans la figure 1.

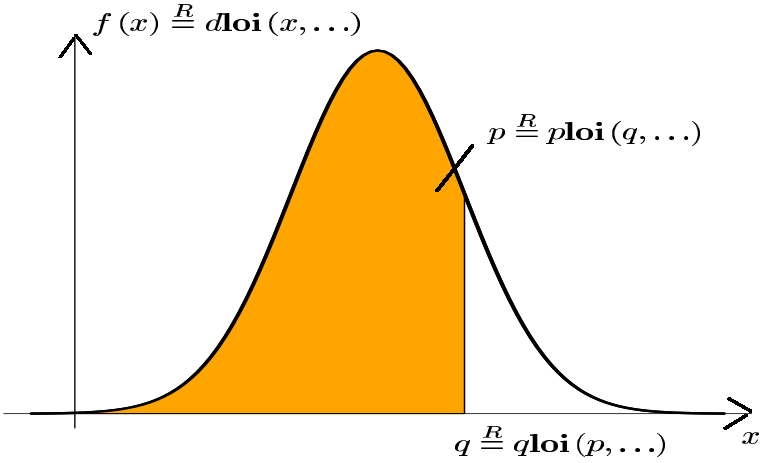


FIGURE 1 – Si $X \rightsquigarrow \mathbf{loi}(\ldots)$ (v.a. continue), alors $f(x) \overset{R}{=} d\mathbf{loi}(x,\ldots)$ représente sa densité de probabilité, $p = F(q) = P(X \leq q) \overset{R}{=} p\mathbf{loi}(q,\ldots)$ sa fonction de répartition et $q = F^{-1}(p) \overset{R}{=} q\mathbf{loi}(p,\ldots)$ son quantile d'ordre p.

Le tableau suivant résume les différentes lois de probabilités considérées dans ce cours de deuxième année ainsi que les instructions R permettant d'évaluer les quantiles et fonctions de répartitions associés à ces lois de probabilités.

lois de probabilités	loi `R`	quantile d'ordre `p`	fonction de répartition en `q`
Normale $\mathcal{N}(\mu,\sigma)$	`norm`	`qnorm(p,`μ,σ`)`	`pnorm(q,`μ,σ`)`
Normale $\mathcal{N}(0,1)$	`norm`	`qnorm(p)`	`pnorm(q)`
Chisquare $\chi^2(n)$	`chisq`	`qchisq(p,`n`)`	`pchisq(q,`n`)`
Fisher $\mathcal{F}(n_1,n_2)$	`f`	`qf(p,`n_1,n_2`)`	`pf(q,`n_1,n_2`)`
Student $\mathcal{S}t(n)$	`t`	`qt(p,`n`)`	`pt(q,`n`)`

Application :

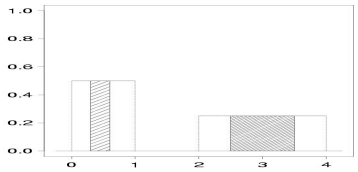
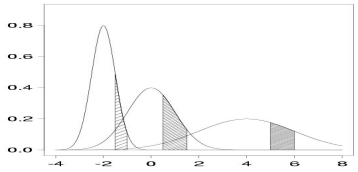
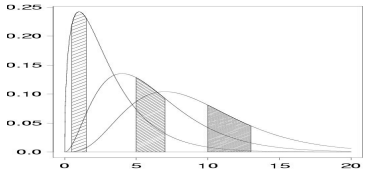
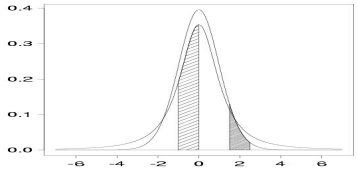
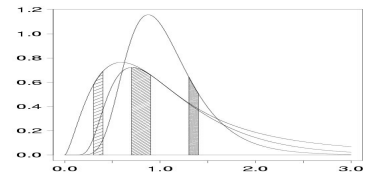
```
> pnorm(1.6449)                              # proba N(0,1) plus petit que 1.6449
[1] 0.9500048
> qnorm(0.95)                                # quantile N(0,1) d'ordre 95% proche de 1.6449
[1] 1.644854
> 1-pnorm(1.96)                              # proba N(0,1) plus grand que 1.96 proche de 2.5%
[1] 0.0249979
> qnorm(c(.95,.975,.99))                     # quantiles N(0,1) d'ordre 95%, 97.5% et 99%
[1] 1.644854 1.959964 2.326348
> qt(c(.95,.975,.99),10)                     # quantiles St(10) d'ordre 95%, 97.5% et 99%
[1] 1.812461 2.228139 2.763769
> pt(c(1.812461,2.228139,2.763769),10)       # les probas correspondantes
[1] 0.950 0.975 0.990
> qchisq(c(.95,.975,.99),10)                 # quantiles Khi2(10) d'ordre 95%, 97.5% et 99%
[1] 18.30704 20.48318 23.20925
> pchisq(c(18.30704,20.48318,23.20925),10)   # les probas correspondantes
[1] 0.950 0.975 0.990
> qf(c(.95,.975,.99),10,20)                  # quantiles F(10,20) d'ordre 95%, 97.5% et 99%
[1] 2.347878 2.773671 3.368186
> pf(c(2.347878,2.773671,3.368186),10,20)    # les probas correspondantes
[1] 0.950 0.975 0.990
```

Illustration du lien entre A.E.P. et A.M.P. : Une instruction `rloi(n,...)` (du même type que les intructions `ploi(q,...)` et `qloi(p,...)` présentées précédemment) permet de générer simultanément n réalisations $\boldsymbol{y} := (y_1, \cdots, y_n)$ d'une v.a. Y ayant pour loi `loi(...)`. Illustrons-le sur une vérification expérimentale (A.E.P.) d'obtention de probabilité, quantile, moyenne et variance relatifs à une loi $\mathcal{N}(1,2)$.

```
> yy<-rnorm(10000,1,2)         # les m=10000 réalisations ont stockées dans le vecteur yy
> yy                           # les 10 premières et 10 dernières composantes de yy
    [1]  4.279724e+00  8.447115e-01 -1.098879e+00  2.826055e+00 -1.356146e+00
    [6]  1.540536e+00  2.304664e+00 -3.084724e+00 -1.098613e+00  9.650271e-01
...
 [9991]  2.483222e+00  3.517996e-01 -1.382401e-02  2.162169e+00  4.103853e-01
 [9996] -1.998113e+00  5.178801e+00  6.135185e-01 -3.672471e-01  9.240147e-01
> mean(yy<0.5)                 # proportion des m=10000 composantes strictement inférieur à 0.5
[1] 0.4053
> pnorm(0.5,1,2)               # idem si m=infini
[1] 0.4012937
> mean(yy==0.5)                # proportion des m=10000 composantes égale à 0.5 (=0 si m=infini)
[1] 0
> mean(0.5<=yy && yy<=3)       # proportion des m=10000 composantes compris entre 0.5 et 3
[1] 0
> pnorm(3,1,2)-pnorm(.5,1,2)   # idem si m=infini
[1] 0.4400511
> quantile(yy,.95)             # quantile d'ordre 95% des m=10000 composantes
     95%
4.316004
> qnorm(.95,1,2)               # idem si m=infini
[1] 4.289707
> mean(yy)                     # moyenne des m=10000 composantes (=1 si m=infini)
[1] 0.9720361
```

```
> var(yy)                          # variance des m=10000 composantes (=2^2=4 si m=infini)
[1] 4.134311
```

Tables de lois usuelles de variables aléatoires continues (pour la statistique)

Nom	Graphe	Densité de probabilité	Espérance et Variance	Remarques
Uniforme $X \rightsquigarrow \mathcal{U}([a,b])$ $a<b$	 $\mathcal{U}([0,1])$ et $\mathcal{U}([2,4])$.	$f(x)=\begin{cases}\frac{1}{b-a} & \text{si } x\in[a,b]\\ 0 & \text{sinon}\end{cases}$	$E(X)=\frac{a+b}{2}$ et $Var(X)=\frac{(b-a)^2}{12}$	La densité de probabilité d'une loi uniforme est un histogramme à 1 classe.
Normale $X \rightsquigarrow \mathcal{N}(\mu,\sigma)$ μ réel et σ réel >0	 $\mathcal{N}(-2,0.5)$, $\mathcal{N}(0,1)$ puis $\mathcal{N}(4,2)$.	$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	$E(X)=\mu$ et $Var(X)=\sigma^2$	1) Si $X \rightsquigarrow \mathcal{N}(\mu,\sigma)$ alors $\frac{X-\mu}{\sigma} \rightsquigarrow \mathcal{N}(0,1)$ 2) Si $X \rightsquigarrow \mathcal{N}(\mu_X,\sigma_X)$ et $Y \rightsquigarrow \mathcal{N}(\mu_Y,\sigma_Y)$ sont des v.a. indépendantes alors $X+Y \rightsquigarrow \mathcal{N}\left(\mu_X+\mu_Y,\sqrt{\sigma_X^2+\sigma_Y^2}\right)$.
Chisquare $X \rightsquigarrow \chi^2(\nu)$ ν entier >0	 $\nu=3$, $\nu=6$ puis $\nu=9$.	$f(x)=\begin{cases}\frac{e^{-\frac{x}{2}}x^{\frac{(\nu-2)}{2}}}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} & \text{si } x>0\\ 0 & \text{sinon}\end{cases}$	$E(X)=\nu$ et $Var(X)=2\nu$	Si $X_1,\cdots,X_n$ sont n lois $\mathcal{N}(0,1)$ indépendantes alors $Y=\sum_{i=1}^{n}{X_i}^2 \rightsquigarrow \chi^2(n)$
Student $X \rightsquigarrow \mathcal{S}t(\nu)$ ν entier >0	 $\nu=2$ et $\nu=30$.	$f(x)=\frac{\left(1+\frac{x^2}{\nu}\right)^{-\frac{1}{2}(\nu+1)}}{\beta\left(\frac{1}{2},\frac{\nu}{2}\right)\sqrt{\nu}}$	$E(X)=0$ si $\nu\geq 2$ et $Var(X)=\frac{\nu}{\nu-2}$ si $\nu\geq 3$	Si $X \rightsquigarrow \mathcal{N}(0,1)$ et $Y \rightsquigarrow \chi^2(\nu)$ sont indépendantes alors $Z=\frac{X}{\sqrt{\frac{Y}{\nu}}} \rightsquigarrow \mathcal{S}t(\nu)$
Fisher $X \rightsquigarrow \mathcal{F}(\nu_1,\nu_2)$ ν_1,ν_2 entiers >0	 $\mathcal{F}(5,200)$, $\mathcal{F}(200,5)$ puis $\mathcal{F}(30,30)$.	$\begin{cases}f(x)=\frac{\nu_1^{\frac{1}{2}\nu_1}\nu_2^{\frac{1}{2}\nu_2}x^{\frac{\nu_1}{2}-1}}{(\nu_1x+\nu_2)^{\frac{1}{2}(\nu_1+\nu_2)}\beta\left(\frac{\nu_1}{2},\frac{\nu_2}{2}\right)}\\ \text{si } x>0\end{cases}$ $f(x)=0$ sinon	$\begin{cases}E(X)=\frac{\nu_2}{\nu_2-2}\\ \text{si } \nu_2\geq 3\end{cases}$ et $\begin{cases}Var(X)=\frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}\\ \text{si } \nu_2\geq 5\end{cases}$	Si $X_1 \rightsquigarrow \chi^2(\nu_1)$ et $X_2 \rightsquigarrow \chi^2(\nu_2)$ sont indépendantes alors $Y=\frac{X_1/\nu_1}{X_2/\nu_2} \rightsquigarrow \mathcal{F}(\nu_1,\nu_2)$