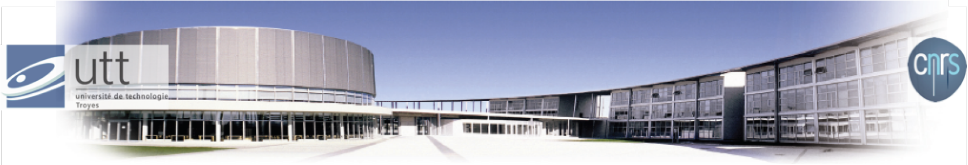


Construction de lois de probabilité à l'aide de l'itérée continue de fonctions exponentielles

Caractérisation d'un indice de traîne

Yann Dijoux

Equipe M2S
Université de Technologie de Troyes

Séminaire LJK-Probabilités & Statistique

23 janvier 2020

Motivations

Motivations en fiabilité

- On considère un système réparable dont la durée de vie pour un système neuf est une variable aléatoire X.
- On note R_X la fiabilité d'un système neuf et $\Lambda_X = -ln(R_X)$ le taux de hasard cumulé associé.
- Si on fait l'hypothèse que le système est soumis uniquement à des maintenances correctives (immédiates et instantanées) et que les maintenances sont minimales (As Bad As Old), alors le nombre moyen de défaillances à l'instant t est $\Lambda_X(t)$.

$\longrightarrow$ L'espérance d'un processus de comptage pour un Processus de Poisson Non Homogène est Λ_X.

Lois usuelles

- Pour une loi exponentielle (standard), le taux de hasard cumulé est linéaire: $\Lambda_{Exp}(t) = t$.
- Pour une loi de Gompertz (standard), le taux de hazard cumulé est exponentiel: $\Lambda_{Gomp}(t) = e^t - 1$.
- Pour une loi de Pareto (standard), le taux de hazard cumulé est logaritmique: $\Lambda_{Par}(t) = ln(1+t)$.

$\longrightarrow$ On peut concevoir que le nombre moyen de défaillances ait une croissance intermédiare: soit plus lente qu'exponentielle mais plus rapide que linéaire (polynomial), soit plus lente que polynomiale mais plus rapide que logarithmique, voire au-delà.

L'itérée continue de $t \mapsto e^t - 1$

$\longrightarrow$ Lien fonctionnel identique entre les taux de hasard cumulés:

$$\Lambda_{Exp} = exp(\Lambda_{Par}) - 1 \quad , \quad \Lambda_{Gomp} = exp(\Lambda_{Exp}) - 1$$

$\longrightarrow$ La fonction $t \mapsto e^t - 1$ a été longuement étudiée afin d'expliciter ses itérées entières, fractionnaires ou continues.

$\longrightarrow$ On note $\phi^{[\gamma]}(t)$ l'**itérée continue** de $t \mapsto e^t - 1$ qu'on nomme **fonction tetration**. γ est le **degré de tétration** de la fonction:

- $\gamma = 1$: $\phi^{[1]}(t) = e^t - 1$
- $\gamma = -1$: $\phi^{[-1]}(t) = ln(1 + t)$
- $\gamma = \frac{-1}{2}$: racine carrée fonctionnelle de $t \mapsto ln(1 + t)$

$\longrightarrow$ Définir une loi dont le taux de hasard cumulé est une paramétrisation de $\phi^{[\gamma]}(t)$.

Enjeux

$\longrightarrow$ Implémentation de la fonction $\phi^{[\gamma]}(t)$.

$\longrightarrow$ Paramétrisation de la **loi Tetration** et mise en évidence de premières propriétés.

$\longrightarrow$ Mise en oeuvre de méthodes d'inférence similaires à celles du graphe de Weibull.

$\longrightarrow$ Caractérisation d'un indice de traîne qui quantifie l'exponentialité de la décroissance de la fonction de survie.

$\longrightarrow$ Caractérisation d'un second indice de traîne qui quantifie une polynomialité éventuelle supplémentaire et généralise en particulier l'indice des valeurs extrêmes et l'indice de queue de Weibull.

$\longrightarrow$ Utilisation d'itérées continues (à partir de $t \mapsto e^t - 1$ ou $t \mapsto sinh(t)$) pour définir des lois continues sur un support infini.

Construction de $\phi^{[\gamma]}$

Calcul numérique actuel de $\phi^{[\gamma]}$

$\longrightarrow$ Existence d'une itérée continue de $t \mapsto e^t - 1$ car c'est une fonction schlicht (bijection avec point fixe et pente unitaire à l'origine).

$\longrightarrow$ Implémentations de $t \mapsto e^t - 1$ et $t \mapsto ln(1+t)$ spécifiques au voisinage de zéro ($expm1$ et $log1p$).

$\longrightarrow$ Calcul numérique de $\phi^{[\gamma]}$ effectué de plusieurs manières:

- Equations d'Abel (Fatou, Walker): utilisation des itérées entières, mais convergence lente.
- Premiers termes du développement en série et lecture de table (Szekeres).

$\longrightarrow$ Généralisation des travaux de Szekeres sur les séries formelles.

Série formelle

$\longrightarrow$ Ecriture sous forme de série formelle au voisinage de $(z=0,\gamma=0)$ en fonction d'une matrice $\{\mathcal{Y}_{i,j}\}_{i\geq 0, j\geq 0}$:

$$P(z,\gamma) = z + \gamma z^2 \sum_{i=0}^{\infty}\left(\sum_{j=0}^{i}\mathcal{Y}_{i,j}\gamma^j\right) z^i$$

$\longrightarrow$ La première colonne de $\mathcal{Y}$ est une mise à l'échelle des nombres de Koszul.

$\longrightarrow$ Le terme général de la matrice peut-être identifié par interpolation après identification entre les itérées entières de $t \mapsto ln(1+t)$ et $P(z,\gamma)$.

$\longrightarrow$ L'expression dépend des nombres de Stirling de première espèce $s(i,j)$ et du i-ème coefficient de la série entière de la j-ème itérée de $t \mapsto ln(1+t)$ $m(i,j)$.

$$\mathcal{Y}_{n,p} = \frac{(-1)^{n+p+1}}{(n+1)!} s(n+2,p+1)\mathcal{Y}_{n,0} + \sum_{\ell=1}^{n+1}\sum_{k=max(p,\ell)}^{n+1} \frac{(-1)^{k+\ell+p+1}}{\ell k!} s(k,p) C_k^{\ell} m(\ell,n+2)$$

Calcul pratique de $\phi^{[\gamma]}$

$\longrightarrow$ $\phi^{[\gamma]}$ est infiniment continue mais sa série formelle a un rayon de convergence nulle (Ecalle).

$\longrightarrow$ Pour calculer $\phi^{[\gamma]}(t)$ avec $|\gamma| \leq 0.5$, on applique à z 50 fois la fonction $t \mapsto ln(1+t)$, puis on y applique une version tronquée de la série formelle $P(.,\gamma)$ et enfin on y applique 50 fois la fonction $t \mapsto e^t - 1$.

$\longrightarrow$ Erreurs de l'ordre de la double précision en tronquant la série formelle à une quarantaine de termes.

La loi de durée de vie Tetration

Forme générale

$$F(t) = 1 - exp\left(\left(-\frac{\eta_0}{\eta}\left[\phi^{[\gamma]}\left(\left(\frac{t-\mu}{\eta_0}\right)^{\beta_0}\right)\right]^{\frac{1}{\beta_0}}\right)^{\beta}\right), \text{ pour } t \geq \mu$$

$\longrightarrow$ Un paramètre de position, deux paramètres d'échelle, deux paramètres de forme, un paramètre de tétration.

$\longrightarrow$ Deux paramétrisations simples sur $\mathbb{R}_+$:

$$F_{ITD}(t) = 1 - exp\left(-\phi^{[\gamma]}\left(\left(\frac{t}{\eta_0}\right)^{\beta_0}\right)\right) \quad , \quad F_{OTD}(t) = 1 - exp\left(-\left(\frac{\phi^{[\gamma]}(t)}{\eta}\right)^{\beta}\right)$$

Lois particulières

- $\gamma = 0$: loi de Weibull (loi exponentielle)
- $\gamma = -1$: lois de Pareto (I,II,III,IV) [log-logistique, Burr XII, Lomax], loi Shifted-Weibull-Pareto
- $\gamma = 1$: loi de Gompertz, loi Weibull-Exponential, loi Modified-Weibull-Extension

$\longrightarrow$ nouvelle loi "Half-Iterate Pareto" pour $\gamma = \frac{-1}{2}$ dont la croissance du taux de hasard cumulé est unique parmi toutes les lois de durée de vie usuelle.

Taux de hasard

$\longrightarrow$ L'allure du taux de défaillance $\lambda = \Lambda'$ dépend du paramètre de tétration γ et du paramètre de forme extérieur β.

$\longrightarrow$ Les monotonies classiques sont toutes présentes (TD,ITD,OTD): croissance (IFR), décroissance (DFR), constance (CFR), courbe en baignoire (BT), courbe en baignoire inversée (UBT).

	$\gamma < 0$	$\gamma = 0$	$\gamma > 0$
$\beta < 1$	DFR	DFR	BT
$\beta = 1$	DFR	CFR	IFR
$\beta > 1$	UBT	IFR	IFR

Exemples de taux de hasard

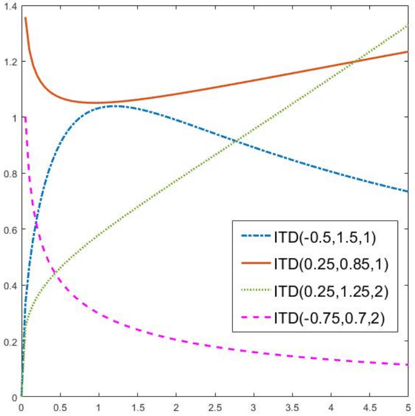


Figure: Taux de hasard ITD$(\gamma, \beta_0, \eta_0)$

Quantile et moments

$\longrightarrow$ Expression explicite de la fonction quantile du fait que la réciproque de $\phi^{[\gamma]}$ soit $\phi^{[-\gamma]}$.

$$Q_{ITD}(p) = \eta_0 \left[\phi^{[-\gamma-1]}\left(\frac{p}{1-p}\right)\right]^{\frac{1}{\beta_0}} \quad ,$$

$\longrightarrow$ Aucune expression analytique des moments.

- Existence des moments à tout ordre si $\gamma > -1$
- Aucun moment si $\gamma < -1$
- Pour $\gamma =$ -1, cela dépend essentiellementdes paramètresde forme.

Méthodes des moindres carrés

$\longrightarrow$ Généralisation de méthodes d'estimation de type "papier Weibull" pour les modèles ITD et OTD. Dans le cadre ITD:

$$ln(\phi^{[-\gamma]}(-ln(1-F_{ITD}(t)))) = \beta_0 ln(t) + \beta_0 ln(\eta_0)$$

$\longrightarrow$ Si γ est fixé, on peut calculer la droite $y = a(\gamma)x + b(\gamma)$ et la somme des carrés des résidus $SSR(\gamma)$ associé au modèle de régression linéaire pour:

$$(ln(x_i^*), ln(\phi^{[-\gamma]}(-ln(1-(i-0.3)/(n+0.4)))))_{i=1..n}$$

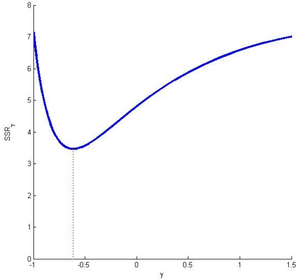
où x^* est le vecteur ordonné des n observations et où la fonction de répartition est estimée suivant la méthode des rangs médians.

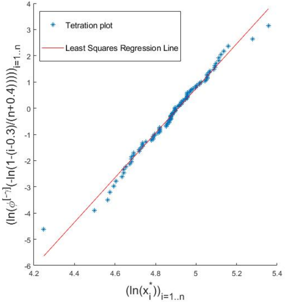
$\longrightarrow$ L'estimateur de γ est celui qui minimise $SSR(\gamma)$ sur l'intervalle [-2,2] et les paramètres de forme et d'échelle peuvent être obtenus à partir de la droite des moindre carrées correspondante.

$$\widehat{\gamma} = \underset{\gamma\in[-2,2]}{\text{argmin}}\, SSR(\gamma) \quad , \quad \widehat{\beta_0} = a(\widehat{\gamma}) \quad , \quad \widehat{\eta_0} = exp(\frac{b(\widehat{\gamma})}{\widehat{\beta_0}})$$

Exemple

$\longrightarrow$ 101 observations de limite de fatigue pour des coupons d'aluminium 6061-T6.





$\longrightarrow$ Estimation des paramètres ($\widehat{\gamma} = -0.613, \widehat{\beta_0} = 8.51, \widehat{\eta_0} = 135.8$)

$\longrightarrow$ Utilisation de ces valeurs pour initialiser la recherche de l'estimateur de maximum de vraisemblance ($\widehat{\gamma} = -0.799, \widehat{\beta_0} = 9.70, \widehat{\eta_0} = 134.1$)

Indices de traîne

Lourdeur de traîne

- Loi à queue lourde lorsque $\gamma < 0$.
- Loi à queue légère lorsque $\gamma > 0$.
- Pour $\gamma = 0$, la loi est à queue lourde si et seulement si $\beta_0 < 1$.
- On parle de loi à queue super-lourdes lorsque $\gamma < -1$

Les indices de tétration

$\longrightarrow$ Une loi de durée de vie associée à une fiabilité R a pour indice (principal) de tétration γ^* lorsque:

$$\gamma^* = \inf_{\gamma \in \overline{\mathbb{R}}} \left(\lim_{t \to +\infty} \frac{-ln(R(t))}{\phi^{[\gamma]}(t)} < \infty \right)$$

$\longrightarrow$ Une loi dont l'indice (principal) de tétration est fini et vaut γ^* admet un indice de tétration secondaire ρ^* si $\mathcal{R}(t) = \phi^{[-\gamma^*]}(-ln(R(t))$ est une fonction à variations régulières telle que $\mathcal{S}(t) = t^{-\rho^*}\mathcal{R}(t)$ est une fonction à variations lente.

$\longrightarrow$ Les lois continues définies sur un support compact ont un indice (principal) de tétration égal à $+\infty$

Indices de tétration et indices de queue

$\longrightarrow$ L'indice de tétration est une mesure nouvelle quantifiant la décroissance asymptotique de la fiabilité.

$\longrightarrow$ L'indice de tétration secondaire a été étudié spécifiquement lorsque γ^* vaut -1 ou 0:

- Pour $\gamma^* = -1$, l'indice de tétration secondaire est égal à l'**indice de Pareto** et est égal à l'inverse de **l'indice des valeurs extrêmes** ξ.
- Pour $\gamma^* = 0$, l'indice de tétration secondaire est l'inverse du **coefficient de queue de Weibull**.

Indices pour les lois usuelles

Loi	γ^*	ρ^*
TD	γ	$\beta_0(ITD)$
log-Cauchy	-2	$\frac{\pi}{\text{paramètre d'échelle}}$
Cauchy	-1	1
Pareto	-1	paramètre de forme
Fréchet	-1	paramètre de forme
Log-Normal	-1	Non défini
Lévy	-1	$\frac{1}{2}$
Birnbaum-Saunders	0	1
Weibull	0	paramètre de forme
Gumbel	0	1
Logistique	0	1
Normale	0	2
Gamma	0	1
Hypo/Hyper exponentielle	0	1
Gompertz	1	1
Uniforme	$+\infty$	$\times$
Beta	$+\infty$	$\times$

Extensions

Première généralisation sur $\mathbb{R}$

$\longrightarrow$ Réflexion de la loi Tétration, technique similaire à celle qui permet d'obtenir la loi de Laplace à partir de la loi exponentielle.

$$Z = (2B - 1)X \quad , X \sim TD \, , B \sim \mathcal{B}(\frac{1}{2}) \, , X \perp\!\!\!\perp B$$

$\longrightarrow$ Ajout d'un paramètre d'asymétrie $\kappa > 0$ (méthode Inverse Scale Factors).

$$Z = ((\frac{1}{\kappa} + \kappa)B - \kappa)X \quad , X \sim TD \, , B \sim \mathcal{B}(\frac{1}{\kappa^2 + 1}) \, , X \perp\!\!\!\perp B$$

- $\kappa < 1$: Asymétrie positive,
- $\kappa > 1$: Asymétrie négative,
- $\kappa = 1$: loi symétrique.

$\longrightarrow$ loi ATD (Asymmetric Tetration Distribution)

Exemples de densité

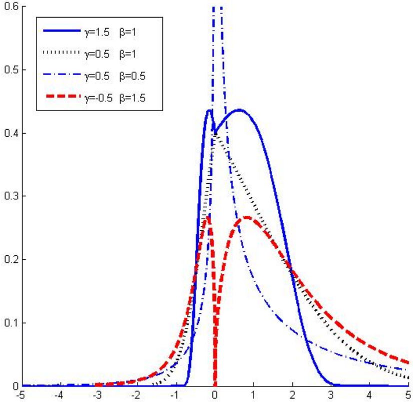


Figure: Densités ATD avec $\kappa = 0.5$

Indices de tetration en $-\infty$

$\longrightarrow$ Les indices de tétration en moins l'infini d'une variable aléatoire X sont définis comme les indices de tétration de $-X$ en plus l'infini.

$\longrightarrow$ Les indices de tétration d'une loi ATD sont identiques en plus et moins l'infini.

$\longrightarrow$ Mode parfois impropre ou fonction peu régulière au voisinage du mode pour les lois ATD

$\longrightarrow$ Proposer des lois plus régulières avec des indices de tétration distincts à l'infini.

- lois dont la fonction de répartition généralise la **loi logistique** et la **loi de Gumbel**:
$$F(t) = \frac{1}{1 + \phi^{[\gamma]}(e^{-t})}$$
- lois construites à partir de l'itérée continue de la fonction sinus hyperbolique.

Lois construites à partir de l'itérée continue de *sinh*

$\longrightarrow$ *sinh* a la propriété d'être une fonction schlicht simple et de type exponentiel sur $\mathbb{R}$.

$\longrightarrow$ Son itérée continue $\psi^{[\gamma]}$ peut être calculée numériquement à l'aide de sa série formelle.

$\longrightarrow$ Lois obtenues par une paramétrisation analogue à la transformation Johnson de type S_U

$\longrightarrow$ Exemple de paramétrisation sur $\mathbb{R}$.

$$F_{DHTD}(z) = \frac{1}{1 + \psi^{[\gamma-\delta]}\left(exp\left(-\psi^{[\delta]}\left(\frac{z-\mu}{\eta}\right)\right)\right)}, \quad z \in \mathbb{R}$$

$\longrightarrow$ Indices de tétration γ (1) en $-\infty$ et δ (1) en $+\infty$.

Exemples de densité

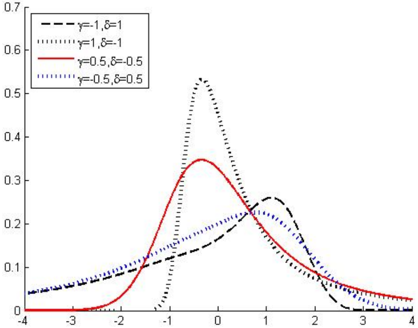


Figure: Densités DHTD

Applications

Applications en sûreté de fonctionnement

$\longrightarrow$ En sûreté de fonctionnement, la majorité des estimations pour des données de durées de fonctionnement/maintenance renvoient un paramètre de tétration dans l'intervalle [-0.9,0.1].

$\longrightarrow$ 150 durées de fonctionnement de valves thermoioniques pendant leur phase de test.

	ITD Estimations			KS p-value				
	$\widehat{\gamma}$	$\widehat{\beta_0}$	$\widehat{\eta_0}$	ITD	BS	Weibull	LogN	Gomp
TValves	-0.889	4.21	3.23	**0.971**	0.847	0.236	0.857	9.6e-5

Table: Données valves

Applications en sûreté de fonctionnement

⟶ Résistance de fibres de carbones soumises à différents stresses.

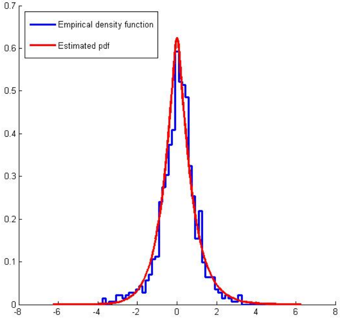
	ITD Estimations			KS p-value (statistic)				
	$\widehat{\gamma}$	$\widehat{\beta_0}$	$\widehat{\eta_0}$	ITD	BS	Weibull	LogN	Gomp
Fibres1	-0.4590	29.34	3.74	**0.564**	0.486	0.367	0.499	0.287
				(0.059)	(0.064)	(0.069)	(0.063)	(0.075)
Fibres2	-0.3025	6.07	4.41	**0.967**	0.835	0.803	0.851	0.716
				(0.063)	(0.079)	(0.082)	(0.078)	(0.089)
Fibres3	-0.8995	8.152	3.06	0.609	0.589	0.634	**0.672**	0.420
				(0.092)	(0.098)	(0.090)	(0.088)	(0.107)
Fibres4	-0.1765	5.901	2.595	**0.992**	0.761	0.970	0.791	0.829
				(0.049)	(0.077)	(0.056)	(0.075)	(0.072)
Fibres5	-0.3655	6.867	2.337	**0.972**	0.703	**0.972**	0.720	0.955
				(0.057)	(0.084)	(0.057)	(0.083)	(0.060)

⟶ La loi OTD (paramètre de forme extérieur) est encore mieux adaptée pour les données 1 et 3.

Rentabilité quotidienne du CaC40

$\widehat{\gamma}$	$\widehat{\eta_0}$	$\widehat{\beta_0}$	$\widehat{\kappa}$
0.047	0.7185	1.034	0.9739

Table: Estimations ATD période 2013-2019



$\longrightarrow$ loi de Laplace

Cours EUR/CHF

ATD	$\widehat{\gamma} = -0.4502$	$\widehat{\eta_0} = 0.1748$	$\widehat{\beta_0} = 1.101$	$\widehat{\kappa} = 1.001$

Table: Estimations 2003-2019 [$>$ 4000 observations]

$\longrightarrow$ loi symétrique à traîne lourde, dont la décroissance de la fonction de survie est à mi-chemin entre une décroissance exponentielle et polynomiale.

Eruptions de la Fournaise 1998-2011

$\longrightarrow$ Quarantaine d'éruptions volcaniques et activité stable entre deux longs repos [$>$ 3 ans].

durée des éruptions (j)	$\widehat{\gamma} = -0.05428$	$\widehat{\eta_0} = 18.009$	$\widehat{\beta_0} = 0.72625$
durée de repos(j)	$\widehat{\gamma} = 0.3019121$	$\widehat{\eta_0} = 140.960$	$\widehat{\beta_0} = 0.8737307$

Table: Estimations ITD

$\longrightarrow$ Prendre aussi en compte les longs repos.

Précipitations annuelles 1910-2019

$\longrightarrow$ Modèle Double-Hyperbolique-Tetration appliqué à des données de précipitations

Ville	$\widehat{\gamma}$	$\widehat{\delta}$	$\widehat{\eta}$	$\widehat{\mu}$	KS p-v
Paris	0.1436	0.390360	71.13	635.7	0.7286
Lyon	0.2851	0.153506	89.61	811.7	0.7870
Marseille	0.2742	0.232591	101.60	579.6	0.9677
Bordeaux	0.48928	0.480469	131.19	902.72	0.9794
Strasbourg	0.1745	0.060507	67.67	624.5	0.9599
Nantes	0.4575	0.524573	101.22	804.8	0.9601
Nancy	0.3521	0.629553	92.26	750.3	0.9798
Dijon	0.2293	0.138996	84.42	729.8	0.9751
Toulouse	0.2320	0.193135	76.86	658.7	0.9992
Montpellier	0.7412	0.246487	151.54	660.8	0.9818
Grenoble 1968-	0.6200	0.40712	114.28	910.8	0.9871
Ajaccio	0.4503	0.5669	107.82	662.96	0.753
Cayenne (973) 1953-	0.2633	0.020091	406.95	3556.0	0.7373
Papeete 1958-	0.3386	0.2753	251.8	1646.4	0.967
Gillot (974) 1953-	0.3685	-0.163579	275.60	1501.5	0.8707

$\longrightarrow$ Prendre aussi en compte les tendances.

Conclusion

Travaux en cours

$\longrightarrow$ Propriétés des estimateurs, sensibilité des paramètres [F. Bertrand]

$\longrightarrow$ Aide à la sélection de modèle.

$\longrightarrow$ Propriétés éventuelles d'inifinie divisibilité, calcul analytique des moments.

$\longrightarrow$ Etude spécifique de la loi Half-Iterate-Pareto

$\longrightarrow$ Approche Bayésienne

$\longrightarrow$ Approfondissement des cas d'étude et applications à d'autres domaines.

$\longrightarrow$ Dépôt d'un package R courant 2020.

$\longrightarrow$ Construction de plusieurs estimateurs spécifiques à l'indice de tétration.

$\longrightarrow$ Proposition de lois à support compact ou multidimensionnel.