Bregman superquantiles. Estimation methods and applications

Tatiana Labopin-Richard

Institut de mathématiques de Toulouse

2 juin 2014

Joint work with F. Gamboa, A. Garivier (IMT) and B. Iooss (EDF R&D).

Let X be a real-valued random variable and let F_X be its cumulative distribution function. We denote for $u \in]0,1[$, the quantile function

$$F_X^{-1}(u) := \inf\{x : F_X(x) \geq u\}.$$

A usual way to quantify the risk associated with X is to consider, for a given number $\alpha \in]0,1[$ close to 1, its lower quantile

$$q_\alpha := F_X^{-1}(\alpha).$$

The quantile is not subadditive
(in some examples $q_\alpha^{X+Y} > q_\alpha^X + q_\alpha^Y$)

- Subadditivity is interesting in finance.
- Rockafellar introduces a new quantity which is subadditive, the superquantile.

Definition of the superquantile

The superquantile Q_α of a law X is defined in this way

$$Q_\alpha := \mathbb{E}(X|X \geq q_\alpha) = \mathbb{E}(X|X \geq F_X^{-1}(\alpha)) = \mathbb{E}\left(\frac{X\mathbf{1}_{X \geq F_X^{-1}(\alpha)}}{1-\alpha}\right)$$

Subadditivity is not the sole interesting property for a measure of risk.

Definition of a coherent measure of risk

Let $\mathcal{R}$ be a measure of risk and X and X' be two real-valued random variables. We say that $\mathcal{R}$ is coherent if, and only if, it satisfies the five following properties :

i) **Constant invariant** : let $C \in \mathbb{R}$, If $X = C$ (a.s.) then $\mathcal{R}(C) = C$.

ii) **Homogeneity** : $\forall \lambda > 0$, $\mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$.

iii) **Subaddidivity** : $\mathcal{R}(X + X') \leq \mathcal{R}(X) + \mathcal{R}(X')$.

iv) **Non decreasing** : If $X \leq X'$ (a.s.) then $\mathcal{R}(X) \leq \mathcal{R}(X')$.

v) **Closeness** : Let $(X_h)_{h \in \mathbb{R}}$ be a collection of random variables. If $\mathcal{R}(X_h) \leq 0$ and $\lim_{h \to 0} ||X_h - X||_2 = 0$ then $\mathcal{R}(X) \leq 0$.

Bregman divergence

Let γ be a **strictly convex function**, $\overline{\mathbb{R}}$-valued on $\mathbb{R}$. We assume that domγ is a non empty open set and that γ is closed proper and differentiable on the interior of domγ.
The **Bregman divergence** d_γ associated to γ is a function defined on dom$\gamma \times$ domγ by

$$d_\gamma(x, x') := \gamma(x) - \gamma(x') - \gamma'(x')(x - x') \ \ ; (x, x' \in \text{dom}\gamma).$$

Definition of the Bregman mean

Let μ be a probability measure whose support is included in $\text{dom}\gamma$ such that $\mu(\overline{\text{dom}\gamma}\setminus\text{dom}\gamma) = 0$ and γ' is integrable with respect to μ. **The Bregman mean is the unique point b in the support of μ satisfying**

$$\int d_\gamma(b,x)\mu(dx) = \min_{m\in\text{dom}\gamma} \int d_\gamma(m,x)\mu(dx). \qquad (1)$$

By differentiating it's easy to see that

$$b = \gamma'^{-1}\left[\int \gamma'(x)\mu(dx)\right].$$

Examples

- Euclidean. $\gamma(x) = x^2$ on $\mathbb{R}$, we obviously obtain, for $x, x' \in \mathbb{R}$,

$$d_\gamma(x, x') = (x - x')^2$$

  and b is the **classical mean**.
- Geometric. $\gamma(x) = x \ln(x) - x + 1$ on $\mathbb{R}_+^*$ we obtain, for $x, x' \in \mathbb{R}_+^*$,

$$d_\gamma(x, x') = x \ln \frac{x}{x'} + x' - x$$

  and b is the **geometric mean**.
- Harmonic. $\gamma(x) = -\ln(x) + x - 1$ on $\mathbb{R}_+^*$ we obtain, for $x, x' \in \mathbb{R}_+^*$,

$$d_\gamma(x, x') = -\ln \frac{x}{x'} + \frac{x}{x'} - 1$$

  and b is the **harmonic mean**.

Definition of the Bregman superquantile

Let $\alpha \in]0,1[$, the Bregman superquantile $Q_\alpha^{d_\gamma}$ is defined by

$$Q_\alpha^{d_\gamma} := \gamma'^{-1}\left(\mathbb{E}(\gamma'(X)|X \geq F_X^{-1}(\alpha))\right) = \gamma'^{-1}\left[\mathbb{E}\left(\frac{\gamma'(X)\mathbf{1}_{X \geq F_X^{-1}(\alpha)}}{1-\alpha}\right)\right].$$

In words $Q_\alpha^{d_\gamma}$ satisfies (1) taking for μ the distribution of X conditionally to $X \geq F_X^{-1}(\alpha)$.

Proposition

Fix α in $]0, 1[$.

i) Any Bregman superquantile always satisfies the properties of constant invariance and non decreasing.

ii) The Bregman superquantile associated to γ is homogeneous if and only if $\gamma'(x) = \ln(x)$ or $\gamma'(x) = (x^\beta - 1)/\beta$ for some $\beta \neq 0$.

iii) If γ' is concave and subadditive, then subadditivity and closeness axioms both hold.

Examples

- **Bregman geometric function** : $\gamma'(x) = x \mapsto \ln(x)$.
  - $\rightarrow$ i) and ii) satisfied.
  - $\rightarrow$ γ' is concave but subadditive only on $[1, +\infty[$. Then it satisfies iii) only for couples (X, X') such that,

$$\min\left(q_\alpha^X(\alpha), q_\alpha^{X'}(\alpha), q_\alpha^{X+X'}(\alpha)\right) > 1$$

- **The subadditivity and the homogeneity are not true in the general case**. For $\gamma(x) = \exp(x)$ and $X \sim \mathcal{U}([0,1])$:

$$\mathcal{R}(2X) - 2\mathcal{R}(X) = 0.000107 > 0,$$

and

$$\frac{\mathcal{R}(4X)}{4\mathcal{R}(X)} = 1,000321 > 1$$

Plug-in estimator

Assume that we have at hand $(X_1, \ldots, X_n)$ an i.i.d sample with same distribution as X. If we wish to estimate $Q_\alpha^{d_\gamma}$, we may use the following empirical estimator :

$$\hat{Q_\alpha^{d_\gamma}} = \gamma'^{-1}\left[\frac{1}{1-\alpha}\left(\frac{1}{n}\sum_{i=\lfloor n\alpha\rfloor+1}^{n}\gamma'(X_{(i)})\right)\right]$$

where $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ is the re-ordered sample built with $(X_1, \ldots, X_n)$.

Theorem : consistency

Let $\alpha \in]0,1[$ be close to 1 and X be a real-valued random variable. Let $(X_1, \dots X_n)$ be an independant sample with the same distribution as X.

We assume that

H1) γ is twice differentiable.

H2) F_X is $\mathcal{C}^1$ on $]0,1[$ and $f_X > 0$ on its support.

H3) The derivative of $\gamma' \circ F_X^{-1}$ that we denote l_γ is non-decreasing and $o\left((1-t)^{-2}\right)$ in the neighborhood of 1.

Then the plug-in estimator is consistent.

Theorem : asymptotic normality

Strongly, we assume that

H1) γ is three times differentiable.

H2) F_X is $\mathcal{C}^2$ and $f_X > 0$ on its support.

H3) The second derivative of $\gamma' \circ F_X^{-1}$ that we denote L_γ is non decreasing and $O\left((1-t)^{-m_L}\right)$ for an $1 < m_L < \frac{5}{2}$, in the neighborhood of 1.

Then the estimator is asymptotically normal.

$$\sqrt{n}\left(\gamma'^{-1}\left(\frac{1}{n(1-\alpha)}\sum_{i=\lfloor n\alpha\rfloor+1}^{n}\gamma'(X_{(i)})\right) - Q_\alpha^{d\gamma}(X)\right) \Longrightarrow \mathcal{N}\left(0, \frac{\sigma_\gamma^2}{\left(\gamma''\circ\gamma'^{-1}(Q_\alpha(\gamma'(X)))\right)^2(1-\alpha)}\right)$$

where

$$\sigma_\gamma^2 := \int_\alpha^1\int_\alpha^1 \frac{(\min(x,y)-xy)}{f_Z(F_Z^{-1}(x))f_Z(F_Z^{-1}(y))} \quad \textit{and } Z := \gamma'(X)$$

Remark

The key point of the proof is to notice that we have the fundamental link between these two quantities

$$Q_{\alpha}^{d\gamma}(X) = \gamma'^{-1}\left(Q_{\alpha}(\gamma'(X))\right).$$

Indeed, as $\gamma'\left(F^{-1}_{(\gamma')^{-1}(Z)}(\alpha)\right) = F_Z^{-1}(\alpha)$, we have

$$\mathbb{E}\left(\gamma'(X)\mathbf{1}_{X>F_X^{-1}(\alpha)}\right) = \mathbb{E}\left(Z\mathbf{1}_{Z>F_Z^{-1}(\alpha)}\right).$$

Then we have to study the asymptotic behaviour of the superquantile.

Plug-in estimator of the superquantile

For $\gamma' = id$, the estimator becomes $(n(1-\alpha))^{-1} \sum_{i=\lfloor n\alpha \rfloor}^{n} X_{(i)}$.

Proposition : consistency of the plug-in estimator of the superquantile

Let $\alpha \in]0,1[$ be close to 1 and X be a real-valued random variable. Let $(X_1, \ldots X_n)$ be an independant sample with the same distribution as X. We assume that

H1) F_X is $\mathcal{C}^1$ on $]0,1[$ and $f_X > 0$ on its support.

H2) The derivative of the quantile function F_X^{-1} denoting l is non-decreasing and $o\left((1-t)^{-2}\right)$ in the neighborhood of 1.

Then, the plug-in estimator is **consistent**.

Proposition : asymptotic normality of the plug-in estimator of the superquantile

Strongly, we assume that

H1) F_X is $\mathcal{C}^2$, $f_X > 0$ on its support.

H2) The second derivative of the quantile function that we denote L is non decreasing and $O\left((1-t)^{-m_L}\right)$ for an $1 < m_L < \frac{5}{2}$, in the neighborhood of 1.

Then the estimator is **asymptotically normal.**

$$\sqrt{n}\left(\frac{1}{n(1-\alpha)}\sum_{i=\lfloor n\alpha\rfloor+1}^{n} X_{(i)} - Q_\alpha\right) \Longrightarrow \mathcal{N}\left(0, \frac{\sigma^2}{1-\alpha}\right)$$

where $\sigma^2 := \int_\alpha^1 \int_\alpha^1 \frac{(\min(x,y)-xy)}{f(F^{-1}(x))f(F^{-1}(y))}$.

Sketch of proof

- Step 1 : Using properties on ordered statistics, we show that our proposition is equivalent to show the convergence in law of

$$\sqrt{n}\left[\frac{1}{n}\sum_{i=\lfloor n\alpha\rfloor}^{n} X_{(i)} - \frac{1}{n}\sum_{i=\lfloor n\alpha\rfloor}^{n} F^{-1}\left(\frac{i}{n+1}\right)\right].$$

Then we use Taylor-Lagrange formula

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=\lfloor n\alpha\rfloor+1}^{n}\left[X_{(i)} - F^{-1}\left(\frac{i}{n+1}\right)\right]\right) \overset{=}{\mathcal{L}} \sqrt{n}\left[\frac{1}{n}\sum_{i=\lfloor n\alpha\rfloor+1}^{n}\left(U_{(i)} - \frac{i}{n+1}\right)\frac{1}{f\left(F^{-1}\left(\frac{i}{n+1}\right)\right)}\right]$$
$$+\frac{1}{\sqrt{n}}\sum_{i=\lfloor n\alpha\rfloor+1}^{n}\left[\int_{\frac{i}{n+1}}^{U_{(i)}}\frac{f'(F^{-1}(t))}{\left(f(F^{-1}(t))\right)^{3}}\left(U_{(i)}-t\right)dt\right].$$

- Step 2 : Convergence of the second order term to 0 in probability (Markov's inequality).
- Step 3 : Identification of the limit in law of the first order using a corollary of the Lindenberg-Feller theorem.

Examples : Exponential distribution

On $\mathbb{R}_*^+$, $f(t) = \exp(-t)$ and $F^{-1}(t) = -\ln(1-t)$.

- **Superquantile.**

→ **Consistency.**

$$l(t) = (1-t)^{-1} = o\left((1-t)^{-2}\right)$$

Then **the estimator of the superquantile is consistent**.

→ **Asymptotic normality.**

$$L(t) = (1-t)^{-2} = O\left((1-t)^{-m_L}\right)$$

for $2 < m_L < \frac{5}{2}$. **The estimator is asymptotically gaussian.**

Examples : Exponential distribution

- **Bregman superquantile with Bregman geometric function.**

For $\gamma(x) = x\ln(x) - x + 1$, $F_Z^{-1}(t) = 1 + \frac{1}{\ln(1-t)}$.

→ **Consistency.**

$$l_\gamma(t) = \frac{1}{(1-t)(\ln(1-t))^2} = o\left((1-t)^{-2}\right),$$

for $2 < m_L < \frac{5}{2}$. **The estimator is consistent.**

→ **Asymptotic normality.**

$$L_\gamma(t) = \frac{(\ln(1-t))^2 + 2\ln(1-t)}{(1-t)^2(\ln(1-t))^4} = O\left((1-t)^{-m_L}\right),$$

for $2 < m_L < \frac{5}{2}$. **Our estimator is asymptotically gaussian.**

Examples : Pareto law of parameter $a > 0$

On $\mathbb{R}_*^+$, $f(t) = ax^{-a-1}$ and $F^{-1}(t) = (1-t)^{\frac{-1}{a}}$.

- **Superquantile**

→ **Consistency.**

$$l(t) = (a(1-t)^{-1-\frac{1}{a}}) = o\left((1-t)^{-2}\right)$$

as soon as $a > 1$. **The consistency is true when $a > 1$.**

→ **Asymptotic normality.**

$$L(t) = C(a)(1-x)^{-\frac{1}{a}-2} = O\left(\frac{1}{(1-t)^{m_L}}\right)$$

for $\frac{3}{2} < m_L < \frac{5}{2}$ as soon as $a > 2$. **The asymptotic normality is true if and only if $a > 2$.**

Examples : Pareto law

- **Bregman superquantile with the Bregman harmonic function**

For $\gamma(x) = -\ln(x) + x - 1$, $F_Z^{-1}(t) = -\frac{1}{a}\ln(1-t)$.

→ **Consistency**.

$$l_\gamma(t) = \frac{1}{a}\frac{1}{1-t} = o\left(\frac{1}{(1-t)^2}\right).$$

The estimator is consistent for every $a > 0$.

→ **Asymptotic normality.**

$$L_\gamma(t) = \frac{1}{a}\frac{1}{(1-t)^2} = O\left((1-t)^{-m_L}\right),$$

for $2 < m_L < \frac{5}{2}$. **The estimator is normally asymptotic for every $a > 0$.**

Conclusion and perspectives

Conclusion :

We introduce a new measure of risk, the Bregman superquantile. We have a large choice for the very regular functions γ this is why this quantity is powerful.

Perspectives :

- Find a NSC for the property of subadditivity.
- Well understand in which case we have to choose such-and-such functions γ.

References I

C. Acerbi and D. Tasche.
On the coherence of expected shortfall.
Preprint, available at http ://arxiv.org/pdf/cond-mat/0104295.pdf.

P. Artzner, F. Delbaen, J.M Eber, and D. Heath.
Coherent measures of risk.
1998.

References II

L. M. Bregman.
The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming.
USSR computational mathematics and mathematical physics, 7(3) :200–217, 1967.

David and Nagaraja.
Order Statistics.
Wiley, 2003.

F. Delbaen.
A remark on the structure of expectiles.
Preprint, available at http ://arxiv.org/abs/1307.5881.

References III

F. Gamboa, A. Garivier, B. Iooss, and T. Labopin-Richard.
Bregman superquantiles. Estimations methods and applications.
Preprint, available at http ://arxiv.org/abs/1405.6677, 2014.

R. Tyrrell Rockafellar.
Convex analysis.
Princeton University Press, 1997.

R.T. Rockafellar.
Coherent approches to risk in optimization under uncertainty.
Tutorials in operation research Informs, 2017.

References IV

R.T. Rockafellar and J.O. Royset.
Random variables, monotone relations, and convex analysis.
Mathematical Programming B, forthcoming, 2013.

R.T. Rockafellar, J.O. Royset, and S.I. Miranda.
Superquantile regression with applications to buffered reliability, uncertainty quantification, and conditional value-at-risk.
Preprint, available at http ://www. math. washington. edurtr/papers. html, 2013.

R.T. Rockafellar and S.Uryasev.
Optimization of conditional value-at-risk.
Journal of risk, 2 :21–42, 2000.

References V

R.T. Rockafellar and S. Uryasev.
Optimization of conditional value at risk.
Portfolio Safeguard, 1999.

R.T. Rockafellar and S. Uryasev.
The fundamental risk quadrangle in risk management, optimization and statistical estimation.
Surveys in Operations Research and Management Science, 18(1) :33–53, 2013.

R.T. Rockafellarl.
The fundamental quadrangle of risk in optimization and estimation.
Risk, 2010.

The quantile is not subadditive

A classical counter-example from "Coherent measures of risk", P. Artzner, F. Delbaen, J.M Eber and D. Health. We consider two independent identically distributed random variables X_1 and X_2 having the same density 0.90 on the interval $[0, 1]$ and the same density 0.05 on the interval $[-2, 0]$. Then

$$q_{X_1}^{10\%} = q_{X_2}^{10\%} = 0 \ \textit{and} \ q_{X_1+X_2}^{10\%} > 0$$