Economics of Risk and Insurance

November 20, 2018

Part I
Risk

All economic activity involves on almost daily basis, to make choices between risky decisions, ie, that are those for which the consequences are not known with certainty. The most obvious case is, for example, that of investment choices: the same amount of money can be invested either in a fixed-income government bond or in a portfolio of stocks whose returns are random. For an insurance company, of course, the risk estimate (as the liabilities to assets) is obviously central.

The purpose of this section is to present a number of tools that compare and quantify risks. The view taken here is deliberately "statistical". It is clear, however, that that any risk measure necessarily refers to an assessment: the choice between several risky situations necessarily meets a criterion, a preference and therefore an attitude towards risk. In the second part of this chapter, we will therefore introduce the hypotheses of behavior that allow for the "classification" of risks.

1 Assumptions on risk

Throughout this section we assume that the variable of interest is the income associated with a decision. It is customary to distinguish risky situations depending on the level of information the decision maker has on the possible values and on the probabilities of these values. The first wellknown dictinction, due to Knight in his work Risk, Uncertainty, and Profit, is between ”risk” and ”uncertainty”.

”Uncertainty must be taken in a sense radically distinct from the familiar notion of Risk, from which it has never been properly separated.... The essential fact is that risk means in some cases a quantity susceptible of measurement, ... It will appear that a measurable uncertainty, or risk proper, as we shall use the term, is so far different from an unmeasurable one...”

We can classify ”uncertainty situations” as follows :

Definition 1. Uncertainty : There is radical uncertainty when the decision maker knows neither the events (possible values, consequences...) nor the probabilities of these. There is uncertainty when one knows the events (possible values, consequences...) but have no estimate (subjective or not) of the probabilities.

Whereas risky situations are those for which probabilities can be (more or less) defined.

Definition 2. We call risk the case when one knows the events (amounts) and have some estimate (subjective or objective) of the probability. Risk is said to be ambiguous when there is some uncertainty on the probability distribution.

We represent here the risk of a real random variable representing the uncertain income of the decision maker. Initially we assume that the support of the studied random variables is bounded $[a, b]$ The probability is zero outside this interval. A first question is "scope" of the risk. Such variable income is it more or less risky than another? The issue of risk measurement is an important question.

2 Risk Analysis

In the sequel the risk will be represented by a real random variables which will be noted by a tilded letter such as $\widetilde{w}$. We restrict ourselves to "non singular" random variable that is to random variables that are mixtures of absolute continuous random variable (with density) and discrete random variables (on countable sets).

One of the first ways to represent the risk is to draw up the histogram of frequencies or more generally density of probability: each possible value (or each range of values) is associated with its frequency (its probability). An equivalent way consists in representing the distribution function (or cumulative density). Set $\widetilde{w}$ a random variable whose values belong to a closed interval included in $]a, b[$. The distribution function gives the probability that the variable is less than a given threshold.

2.1 Cumulative Distribution function

Definition 3. The distribution function F (CDF) of a real random variable $\widetilde{w}$ is given by $F(x) = \Pr[\widetilde{w} \leq x]$. This function is increasing positive right continuous at any point $[-\infty, +\infty]$ and such that $x \leq a \Rightarrow F(x) = 0$, and $x \geq b \Rightarrow F(x) = 1$.

The points of discontinuity correspond to atoms (Dirac masses). At these points the probability density is infinitely concentrated.

2.2 Stochastic dominance of degree 1

One way to compare two random incomes is to compare the likelihood of adverse events. Implicitly, the decision maker prefers random variables that are more likely to be large. That is, he prefers $\widetilde{w}$ to $\widetilde{v}$ if for all threshold u, the probability that income is smaller than u is greater with $\widetilde{v}$ than with $\widetilde{w}$. We can somehow conclude that $\widetilde{v}$ is less "favorable" than $\widetilde{w}$, since the bad valuesare more frequent. This observation leads to the following definition.

Definition 4. $\widetilde{w}$ stochastically dominates at the first degree $\widetilde{v}$ if and only if $\forall x \in [a, b]$ $F_{\widetilde{v}}(x) \geq F_{\widetilde{w}}(x)$

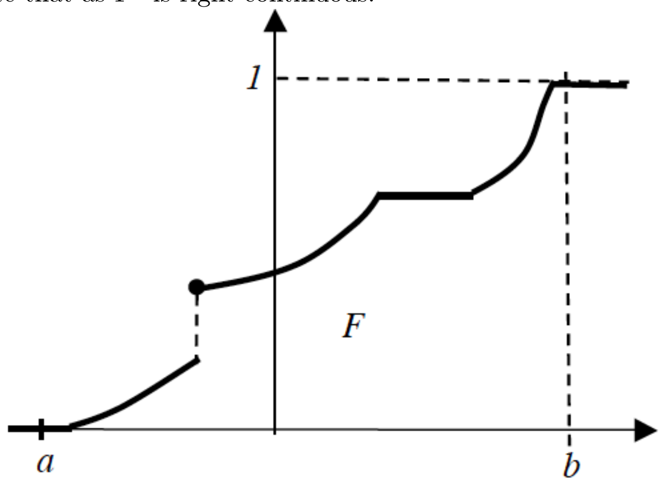
2.3 Quantile Function and Value at Risk

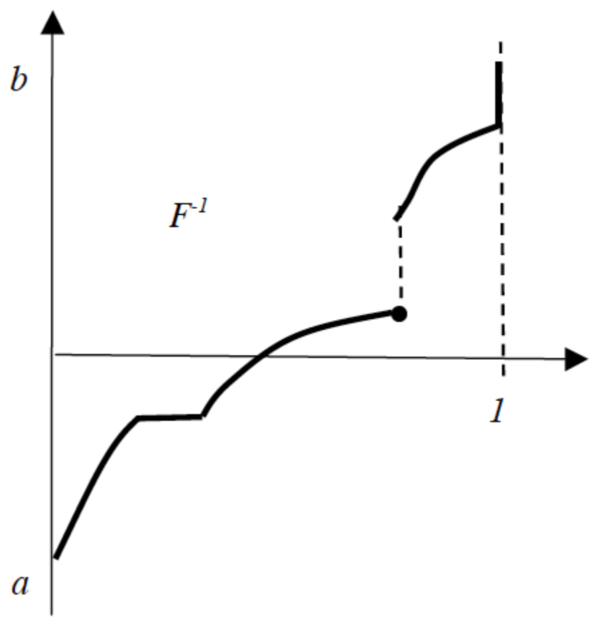
To that distribution function corresponds the quantile function which is simply its inverse function

Definition 5. the quantile function is defined by:

$$F_{\widetilde{w}}^{-1}(u) = \inf \{x, F_{\widetilde{w}}(x) \geq u\}$$

Note that as F is right continuous.





The quantile function reads as follows: in $u \times 100$ percent of the cases, the variable is less than $F_{\tilde{w}}^{-1}(u)$.

Note that the horizontal parts of F are "jumps" of F^{-1} and correspond to intervals of zero probability. F jumps, on the contrary, corresponding to the flat parts ofF^{-1}, are atoms , ie discrete values having non-zero probabilities.

The quantile function gives a first idea of the extent of the risk. In particular, it estimates the mattress to absorb losses (negative earnings) without using external inflows. For example we know that the random variable, hence the income, has a 10% probability to be less than $F^{-1}(0.1)$. So if the individual or institution has a reserve larger than $-F^{-1}(0.1)$, bankruptcy probability is less than

0.1. Formally, if K is the reserve:
$\Pr\left(K+\widetilde{w}\leq 0\right)\leq 0.1\Rightarrow \Pr\left(\widetilde{w}\leq -K\right)\leq 0.1\Rightarrow -K\leq F^{-1}(0.1)\Rightarrow K\geq -F^{-1}(0.1)$

Definition 6. If $\widetilde{w}$ is the random income with quantile function $F_{\widetilde{w}}^{-1}$, one calls Value at Risk at level u : $\mathrm{VaR}_{\widetilde{w}}(u)=-F_{\widetilde{w}}^{-1}(u)$. This is the needed reserve to absorb losses with at least probability$1-u$.

Remark 7. Note that when the random variable is the loss (and not income), i.e. $\widetilde{\ell}=-\widetilde{w}$, then $F_{\widetilde{\ell}}(s)=1-F_{\widetilde{w}}(-s)$ and VaR is: $\mathrm{VaR}_{\widetilde{\ell}}(u)=F_{\widetilde{\ell}}^{-1}(1-u)$

Value at Risk is a "risk measure" very commonly used in practice, including, moreover, random variables of infinite support.

For example, for a Gaussian random variable with mean μ and variance σ^2 , we have :
$F^{-1}(u)=\mu+N^{-1}(u)\sigma$
Where N^{-1} is the quantile function of the standard Gaussian distribution.
For a Cauchy distribution the quantile function is :
$F^{-1}\left(u\right)=\mu+\sigma\tan\left(\frac{\pi}{2}\left(2u-1\right)\right)$

3 Spread Analysis

First degree stochastic dominancee does not compare "concentrations" of random variables. In this section we will try to give a meaning to the comparisons of concentrations.

3.1 Expected Shortfall, Lorenz function

Some properties of the quantile function will allow us to analyze quite naturally the "concentration" of risks around the average.

The first result is interesting:

Proposition 8. *If $\widetilde{w}$ is a random variable with finite expectated value (which is the case here, since its support is bounded) we have:*

$$\mathbb{E}\left(\widetilde{w}\right)=\int_0^1 F^{-1}(t)dt$$

this result is obtained immediately,when F is strictly monotone and continuous, by change of variable $x=F^{-1}(t)$, $dt=dF(x)$

It extends easily when F has jumps and flats.

To better understand the risk we may be tempted to focus attention on the adverse event. Thus we know that with probability u The income is less than $F^{-1}(u)$

The average income in these unfavorable conditions, that is to say $\mathbb{E}(\widetilde{w}/\widetilde{w}\leq F^{-1}(\alpha))$,is equal to:

$$\frac{1}{u}\int_0^u F^{-1}(t)dt$$

Thus, if we focus on the 5% of most unfavorable case, the average income is $20\int_0^{0.05}F^{-1}(t)dt$.

Definition 9. On calls "expected shortfall" (or aVaR) at α the expected value of losses in $\alpha \times 100$ % of worst case:

$$\mathrm{ES}_{\widetilde{w}}\,(u) = -\mathbb{E}(\widetilde{w}/\widetilde{w} \leq F^{-1}(u)) = -\frac{1}{u}\int_0^u F^{-1}(t)dt$$

We see that ES is the average VaR. We will see later that we can generalize this kind of measure.

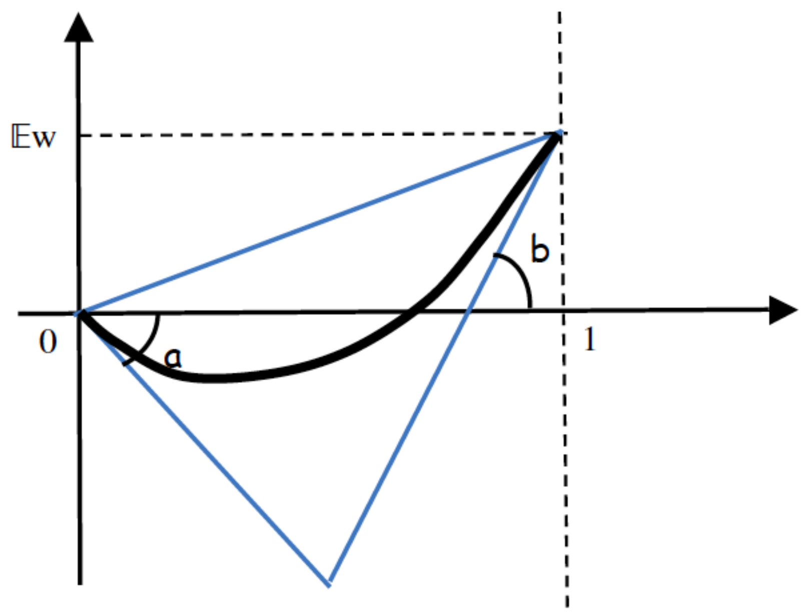
In fact, the cumulative function of the quantile function gives an idea of the concentration of the probability distribution.

Definition 10. One calls absolute Lorenz function the function defined by:

$$L_{\widetilde{w}}(u) = \int_0^u F^{-1}(t)dt$$

Proposition 11. *The Lorenz function* L *is a convex function (therefore continuous and left and right differentiablet at each point) satisfying* $L(0) = 0$, $L(1) = \mathbb{E}\,[\widetilde{w}]$ *. The derivative increases at any point, taking values between* a *and* b *.*

Conversely, every convex function L defined between 0 and 1, such that $L(0) = 0$, $L(1) = c$, with a slope greater than a at $0+$ and less than b at 1 is in the Lorenz function of a real random variable with support included in $[a, b]$ and with expected value c.



In the graph above the Lorenz function a random variable is represented. The slopes of the triangle are $a, E\,(\widetilde{w})\,, b$.

Remark 12. Extension to support$]-\infty, +\infty[$. The Lorenz function can be defined for a random variable with support $]-\infty, +\infty[$ provided that its expected valuee is$> -\infty$. In this case the Lorenz function is convex with slope $-\infty$ at (0,0) and with slope $+\infty$ at 1, or with a vertical asymptote in 1 if $\mathbb{E}\,(\widetilde{w}) = +\infty$.

3.2 Mean Preserving Concentration

Given a random variable with support $[a, b]$, and expected value$c = \mathbb{E}(\widetilde{w})$ The position of the Lorenz function in the triangle $a \leq c \leq b$ gives a fairly accurate idea of the concentration.

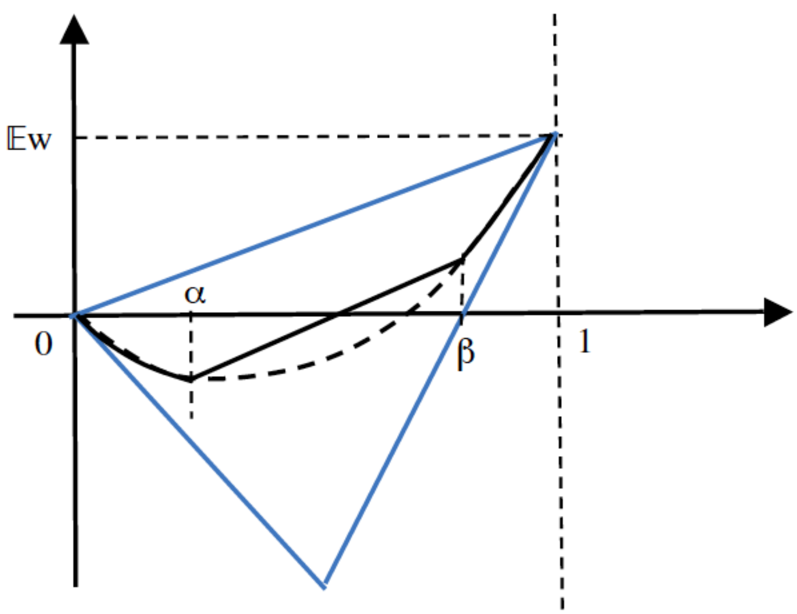
Consider for example the linear Lorenz function (thus the slope is $c = \mathbb{E}(\widetilde{w})$)): $L^{(2)}(u) = cu$. For that Lorenz function F^{-1} is equal to a constant c . The random variable is completely concentrated at c . The value is c with probability 1.

In contrast consider the Lorenz function formed y the segment with slope a and the segment with slope b, whose equation is $L(u) = \max\{au, -b(1-u)+c\}$. This is the Lorenz function of the random variable with support $[a, b]$ and expected value c which is the more "eccentric" since the only non-zero probability values are extreme values a and b .

Remark 13. Note 11. note that Lorenz function of a uniform variable on $[a, b]$ With $c = \frac{a+b}{2}$ is $L(u) = \frac{b-a}{2}u^2 + au$. This is the parabola through (0,0) with slope a at 0 and b at 1.

More generally easy handling on the Lorenz function to "condense" the risk. Specifically, consider a random variable characterized by its Lorenz function$L$. Define a new random variable as follows. We fixed α and β between 0 and 1, and built a new convex function that coincides with$L^{(2)}$ outside $]\alpha, \beta[$ and equal to the chord (whose equation is $L(\alpha) + \frac{L(\beta)-L(\alpha)}{\beta-\alpha}(u-\alpha)$) between α and β . In other words, the new convex function is defined by:

$$T_{[\alpha,\beta]}(L)(u) = \max\left\{L(u), L(\alpha) + \frac{L(\beta)-L(\alpha)}{\beta-\alpha}(u-\alpha)\right\}$$



Proposition 14. *mean preserving elemental concentration. Given a Lorenz function L of a random vairiable . The convex function*

$$T_{[\alpha,\beta]}(L)(u) = \max\left\{L(u), L(\alpha) + \frac{L(\beta)-L(\alpha)}{\beta-\alpha}(u-\alpha)\right\}$$

is the Lorenz function of the random variable (with the same expectation) obtained by replacing $\widetilde{w}$ *between* $F^{-1}(\alpha)$ *and* $F^{-1}(\beta)$ *by a constant value equal to* $\frac{1}{\beta-\alpha}\int_{\alpha}^{\beta} F^{-1}(u)$ *. This value is simply equal to the conditional expectation*

$$\mathbb{E}\left(\widetilde{w}/\widetilde{w} \in \left[F^{-1}(\alpha), F^{-1}(\beta)\right]\right)$$

Indeed, betweenα and β the "new" quantile function (which is the derivative of the Lorenz function) is constant. The new distribution function is the one of $\tilde{w}$ outside $\left[F^{-1}(\alpha), F^{-1}(\beta)\right]$, equals α between $F^{-1}(\alpha)$ and $\frac{L(\beta)-L(\alpha)}{\beta-\alpha}$, has an amplitude jump $\beta-\alpha$ in $\frac{L(\beta)-L(\alpha)}{\beta-\alpha}$, And is constant equal to β between $\frac{L(\beta)-L(\alpha)}{\beta\alpha}$ and$F^{-1}(\beta)$.

In a sense, this manipulation of the Lorenz function "concentrates" distribution while maintaining the constant expected value.

This mean preserving concentration will allow us to compare two random variables with the same expected value. Let then $\widetilde{v}$ and$\widetilde{w}$ two random variables with support included in $[a,b]$ and expected value c . If for any value of α $L_{\widetilde{v}}(\alpha) \leq L_{\widetilde{w}}(\alpha)$ we can conclude unambiguously that $\widetilde{w}$ is more concentrated than$\widetilde{v}$. Indeed, it is easy to see graphically that we can approach the curve $L_{\widetilde{w}}$, by operating from $L_{\widetilde{v}}$ a succession of elementary concentrations. At each step we take a chord based on the curve and tangent to $L_{\widetilde{w}}$.

Definition 15. Given two random variables $\widetilde{v}$ and $\widetilde{w}$ with support included in $[a,b]$ and even expected value c. It is said that $\widetilde{w}$ is more concentrated than $\widetilde{v}$, $\widetilde{w} \succsim \widetilde{v}$, if we can get the Lorenz function $\widetilde{w}$ as the limit of elemental concentrations from the Lorenz function $\widetilde{v}$.

Proposition 16. *Given two random variables* $\widetilde{v}$ *and* $\widetilde{w}$ *with support included in* $[a,b]$ *and with the same expected value c, then*

$$\forall u, L_{\widetilde{v}}(u) \leq L_{\widetilde{w}}(u) \Longleftrightarrow \widetilde{w} \succsim \widetilde{v}$$

Note that this proposition is also valid in the case of distributions with infinite support provided that the expected value remains finite.

So when the Lorenz curve of $\widetilde{w}$ is above that of $\widetilde{v}$ with the same endings, we can say unambiguously that $\widetilde{w}$ is more concentrated than $\widetilde{v}$.

3.3 Measures based on the quantile function.

There are other equivalent ways of writing the Lorenz curve for $\widetilde{v}$ is below that of $\widetilde{w}$: i.e. : $\forall u, L_{\widetilde{v}}(u) \leq L_{\widetilde{w}}(u)$.

Let φ a real function from $[0,1]$ to $[0,1]$ integrable.

define:

$$U_{\varphi}(\widetilde{v}) = \int_0^1 \varphi(t) F_{\widetilde{v}}^{-1}(t) dt$$

We can show that as soon as φ is taken decreasing, U_{φ} is a "utility measure" :

Indeed, integrating by parts yields:

$$U_\varphi(\widetilde{v}) = \mathbb{E}(\widetilde{v})\varphi(1) - \int_0^1 L(u)d\varphi(u\ )$$

Thus, for two random variables with the same expected value:

$$U_\varphi(\widetilde{v}) \le U_\varphi(\widetilde{w}) \Longleftrightarrow \int_0^1 L_{\breve{v}}(u)d\varphi(u) \ge \int_0^1 L_{\breve{w}}(u)d\varphi(u)$$

We then see that when φ is decreasing, $\widetilde{w} \succsim \widetilde{v} \Longrightarrow U_\varphi(\widetilde{w}) \ge U_\varphi(\widetilde{v})$.
Conversely, if $\forall\varphi$ decreasing, $U_\varphi(\widetilde{w}) \ge U_\varphi(\widetilde{v})$, then we must have $\forall u,\quad L_{\widetilde{v}}(u) \le L_{\widetilde{w}}(u)$.
We can thus state the following proposition :

Proposition 17. *Given two random variables $\widetilde{v}$ and $\widetilde{w}$ with support included in $[a,b]$ and the same expected value c.*

$$\widetilde{w} \succsim \widetilde{v} \Longleftrightarrow \forall\varphi \text{ decreasing, } U_\varphi(\widetilde{w}) \ge U_\varphi(\widetilde{v})$$

When φ is decreasing, $U_\varphi(\widetilde{w})$ can thus be regarded as an index, a "measure" of the concentration risk of $\widetilde{w}$.

When ones takes $\varphi(u) = \frac{1}{\alpha}1_{[0,\alpha]}(u)$, which is a decreasing step function, $\rho_\varphi(\widetilde{w}) = -U_\varphi(\widetilde{w})$ is then just the opposite of the expected shortfall at the level α .

Remark 18. The index $\rho_\varphi(\widetilde{w}), = -\int_0^1 \varphi(t)F_{\widetilde{w}}^{-1}(t)dt$ can be regarded as a “risk index” for the random income $\widetilde{w}$. A decresaing φ means that the decision maker does not like risk. We will come back on this in the sequel.

3.4 Coherent risk measures

Measures of risk of type $\delta = -\rho_\varphi$ defined in the previous paragraph are related to the concentration of risk "around the mean."

In a more general way, we can try to define a risk indicator axiomatically. In this approach, a risk measure is an application ρ that associates a real value to each random variable with support$[a,b]$, satisfying some coherence properties.

The literature distinguishes four desirable properties:

Definition 19. positive homogeneity: for all random variable $\widetilde{w}$ with bounded support $\forall\lambda \ge 0\ \ \delta(\lambda\widetilde{w}) = \lambda\delta(\widetilde{w})$

The extent of the risk is independent of the unit of account.

It must also reflect the fact that the "diversification" reduces the risk (when the expected values are finite):

Definition 20. For all random variables $\widetilde{v}$ and $\widetilde{w}$ with support included in $[a,b]$ and with the same expected value.

$$\delta\left(\frac{\widetilde{v}+\widetilde{w}}{2}\right) \le \frac{1}{2}\delta(\widetilde{v}) + \frac{1}{2}\delta(\widetilde{v})$$

Increasing the expected value reduces the risk additively:

Definition 21. $\forall \widetilde{w},\ \forall s \geq 0\ \ \delta(\widetilde{w}+d) = \delta(\widetilde{w}) - d$

And last monotony :

Definition 22. If $\widetilde{w} \geq \widetilde{v}$ then $\delta(\widetilde{w}) \leq \delta(\widetilde{v})$

One can show quer alone risk measurements that verify these four properties are precisely those of the form:
$\delta(\widetilde{w}) = -\int_0^1 \varphi(u) F_{\widetilde{w}}^{-1}(u)du$

4 Second degree Stochastic dominance

The preceding notions can also be defined by analyzing the integral of the distribution function or bi-cumulative.

Definition 23. We call bicumulative $F_{\widetilde{w}}^{(2)}$ the function defined for $x \in [a,b]$ by $F_{\widetilde{w}}^{(2)}(x) = \int_a^x F_{\widetilde{w}}(y)dy$

There is a remarkable relationship between the bicumulative function and the Lorenz function.

Proposition 24. *Given a random variable with support* $[a,b]$ *we have :*

$$F_{\widetilde{w}}^{(2)}(x) = \max_{\alpha}\left[\alpha x - L_{\widetilde{w}}(\alpha)\right]$$

$$L_{\widetilde{w}}(\alpha) = \max_{x}\left[x\alpha - F_{\widetilde{w}}^{(2)}(x)\right]$$

(We say that the functions $F_{\widetilde{w}}^{(2)}$ and $L_{\widetilde{w}}$ are dual in the sense of Fenchel Moreau)

Indeed, as a maximum of affine functions in x $x \to \max_{\alpha}\left[\alpha x - L_{\widetilde{w}}^{(2)}(\alpha)\right]$ is convex. Moreover, since L is convex, the first order condition $x = L'(\alpha^*) \Longleftrightarrow \alpha^* = F(x)$ is necessary and sufficient.

and so :

$$\max_{\alpha}\left[\alpha x - L_{\widetilde{w}}^{(2)}(\alpha)\right] = xF(x) - \int_0^{F(x)} F^{-1}(u)\,du = xF(x) - \int_a^x ydF(y) = \int_a^x F(y)dy$$

The second equality is obtained in the same way.

Note that, in particular, we have

$$F_{\widetilde{w}}^{(2)}(b) = \max_{\alpha}\left[\alpha b - L_{\widetilde{w}}(\alpha)\right] = b - E(\widetilde{w})$$

From this we deduce the following result:

Proposition 25. *Let two random variables* $\widetilde{v}$ *And* $\widetilde{w}$ *with support included in* $[a,b]$ *with of the same expected value c*

$$\widetilde{w} \succsim \widetilde{v} \Longleftrightarrow \forall x,\ F_{\widetilde{w}}^{(2)}(x) \leq F_{\widetilde{v}}^{(2)}(x)$$

Indeed :
$\widetilde{w} \succsim \widetilde{v} \Longleftrightarrow \forall\alpha,\, L_{\widetilde{w}}(\alpha) \geq L_{\widetilde{v}}(\alpha)$
$\Rightarrow \forall x, \alpha\ \alpha x - L_{\widetilde{w}}(\alpha) \leq \alpha x - L_{\widetilde{v}}(\alpha)$
$\Rightarrow \forall x,\, \max_{\alpha}[\alpha x - L_{\widetilde{w}}(\alpha)] \leq \max_{\alpha}[\alpha x - L_{\widetilde{v}}(\alpha)]$

Definition 26. Let two random variables $\widetilde{v}$ and $\widetilde{w}$ with support included in $[a, b]$. It is said that $\widetilde{w}$ second degree stochastically dominates $\widetilde{v}$, $\widetilde{w}\, SDD\, \widetilde{v}$, if $\forall x, F^{(2)}_{\widetilde{w}}(x) \leq F^{(2)}_{\widetilde{v}}(x)$

Proposition 27. *Consider two random variables $\widetilde{v}$ and $\widetilde{w}$ with support included in $[a, b]$ and with the same expected value c.*

$$\widetilde{w}\, SDD\, \widetilde{v} \Longleftrightarrow \forall x,\ F^{(2)}_{\widetilde{w}}(x) \leq F^{(2)}_{\widetilde{v}}(x) \Longleftrightarrow \forall\alpha, L_{\widetilde{w}}(\alpha) \geq L_{\widetilde{v}}(\alpha) \Longleftrightarrow \widetilde{w} \succsim \widetilde{v}$$

4.1 Mean preserving spread

The notion of mean preserving spread can be defined using the distribution function and the bicumulative function. Let two variables $\widetilde{v}$ and $\widetilde{w}$ with support included in $[a, b]$ and with the same expected value. We note $F_{\widetilde{v}}$ and $F_{\widetilde{w}}$ their cumulative distribution functions.

Definition 28. $\widetilde{v}$ is an elementary mean preserving spread of $\widetilde{w}$ if and only if

1. there is an interval $[x, y]$ such that $\forall s \in [x, y]\,, F_{\widetilde{v}}(x) = F_{\widetilde{v}}(s) = F_{v}(y)$, that is $\Pr[\widetilde{v} \in [x, y]] = 0$
2. On the other hand,$\mathbb{E}[\widetilde{v}] = \mathbb{E}[\widetilde{w}]$, that is $\int_a^b F_{\widetilde{v}}(s)ds = \int_a^b F_{\widetilde{w}}(s)ds$

An elemental mean preserving spread consists in taking the mass out of an interval $[x, y]$ and to distribute it outside while maintaining constant expected value.

Definition 29. $\widetilde{v}$ is mean preserving spread of $\widetilde{w}$ if and only if $\widetilde{v}$ is obtained by a sequence (possibly infinite) of elementary spreads of $\widetilde{w}$.

It is interesting to understand the effect of an elementary spread on the bi-cumulative.

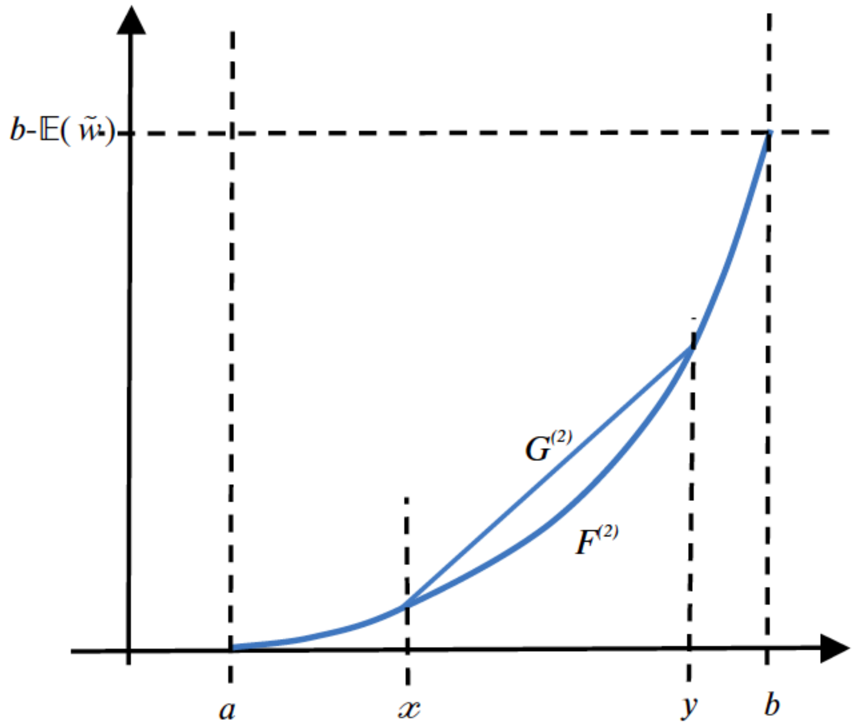
We know that $F^{(2)}_{\widetilde{w}}(x) = \int_a^x F_{\widetilde{w}}(s)ds$. $F^{(2)}_{\widetilde{w}}$ is a convex function (since its derivative is increasing). Given an interval $[x, y]$ define the following affine function:

$$s \rightarrow \frac{F^{(2)}_{\widetilde{w}}(y) - F^{(2)}_{\widetilde{w}}(x)}{y - x}(s - x) + F^{(2)}_{\widetilde{w}}(x)$$

This is the string of$F^{(2)}_{\widetilde{w}}$ between x and y. Then let the function:

$$G^{(2)}(s) = \max\left\{F^{(2)}_{\widetilde{w}}(s), \frac{F^{(2)}_{\widetilde{w}}(y) - F^{(2)}_{\widetilde{w}}(x)}{y - x}(s - x) + F^{(2)}_{\widetilde{w}}(x)\right\}$$

This is the function that replaces $F^{(2)}_{\widetilde{w}}$ by its string between x and y . This is obviously a convex function. It is easy to see that it is the bicumulate of the random variable with the same expectated value but that never takes values in $[x, y]$. Indeed, between x and y $G'(s)$ is constant, which means that the measure of $[x, y]$ is zero.



Let us now take two random variables of the same expectation such that $F_{\widetilde{w}}^{(2)}(x) \leq F_{\widetilde{v}}^{(2)}(x)$. they are therefore two increasing convex functions which have the same values in a , (0) and b , $(b - \mathbb{E}(\widetilde{w})))$.

It is intuitive to see that one can obtain $F_{\widetilde{v}}^{(2)}$ As the limit of a sequence of functions $F_k^{(2)}$ Where $F_1^{(2)} = F_{\widetilde{w}}^{(2)}$ And $F_{k+1}^{(2)}$ is an elementary spreading $F_k^{(2)}$ where the string is tangent to $F_{\widetilde{v}}^{(2)}$.

We thus have the following result:

Proposition 30. *Consider two random variables $\widetilde{v}$ and $\widetilde{w}$ with support included in $[a, b]$ and with the same expected value c.*

$$\forall x,\ F_{\widetilde{w}}^{(2)}(x) \leq F_{\widetilde{v}}^{(2)}(x) \Longleftrightarrow \widetilde{v} \text{ is a mean preserving spread of } \widetilde{w}$$

5 Expected utility hypothesis

Definition 31. Expected utility Hypothesis. We say that the decision is based on expected utility if it exits a real increasing function u, such that for all random variable of income $\widetilde{w}_1$ and $\widetilde{w}_2$, $\widetilde{w}_1$ is preferred to $\widetilde{w}_2$ if and only if $\mathbb{E}\left[u\left(\widetilde{w}_1\right)\right] \geq \mathbb{E}\left[u\left(\widetilde{w}_2\right)\right]$

Let F_1 and F_2 the distribution functions and $F_1^{(2)}$ and $F_2^{(2)}$ The bicumulates of $\widetilde{w}_1$ and $\widetilde{w}_2$, we have :

$$\mathbb{E}\left[u\left(\widetilde{w}_1\right)\right] \geq \mathbb{E}\left[u\left(\widetilde{w}_2\right)\right] \Longleftrightarrow \int_a^b u\left(x\right)\left[dF_1(x) - dF_2(x)\right] \geq 0$$

By integrating by parts:

$$\mathbb{E}\left[u\left(\widetilde{w}_1\right)\right] \geq \mathbb{E}\left[u\left(\widetilde{w}_2\right)\right] \Longleftrightarrow \left[u(x)\left(F_1(x)-F_2(x)\right)\right]_a^b - \int_a^b u'\left(x\right)\left[F_1(x)-F_2(x)\right] \geq 0$$

Since F_1 And F_2 have the same values in a and b :

$$\mathbb{E}\left[u\left(\widetilde{w}_1\right)\right] \geq \mathbb{E}\left[u\left(\widetilde{w}_2\right)\right] \Longleftrightarrow -\int_a^b u'\left(x\right)\left[F_1(x)-F_2(x)\right] \geq 0$$

By integrating in parts a second time:

$$\mathbb{E}\left[u\left(\widetilde{w}_1\right)\right] \geq \mathbb{E}\left[u\left(\widetilde{w}_2\right)\right] \Longleftrightarrow -\left[u'(x)\left(F_1^{(2)}(x)-F_2^{(2)}(x)\right)\right]_a^b + \int_a^b u''\left(x\right)\left[F_1^{(2)}(x)-F_2^{(2)}(x)\right]dx \geq 0$$

$$\mathbb{E}\left[u\left(\widetilde{w}_1\right)\right] \geq \mathbb{E}\left[u\left(\widetilde{w}_2\right)\right] \Longleftrightarrow u'(b)\left(\mathbb{E}\left(\widetilde{w}_1\right)-\mathbb{E}\left(\widetilde{w}_2\right)\right) + \int_a^b u''\left(x\right)\left[F_1^{(2)}(x)-F_2^{(2)}(x)\right]dx \geq 0$$

We can deduce :

Proposition 32. *Consider two random variables $\widetilde{v}$ and $\widetilde{w}$ with support included in $[a,b]$*

$$\forall x,\ F_1^{(2)}(x) \leq F_2^{(2)}(x) \Longleftrightarrow \forall u, u' \geq 0, u'' \leq 0, \mathbb{E}\left[u\left(\widetilde{w}_1\right)\right] \geq \mathbb{E}\left[u\left(\widetilde{w}_2\right)\right]$$

Thus, if we have two random varaibles such that $\forall x,\ F_1^{(2)}(x) \leq F_2^{(2)}(x),$ and therefore in particular $\mathbb{E}\left(\widetilde{w}_1\right) \geq \mathbb{E}\left(\widetilde{w}_2\right),$

Then any concave increasing criterion will prefer $\widetilde{w}_1$ to $\widetilde{w}_2$

6 Dual criterion

The previous approach gives the opportunity to define duality between expected utility and rank dependent expected utility. Consider the following criterion :

$$U_{\varphi}\left(\widetilde{v}\right) = \int_0^1 \varphi(t) F_{\widetilde{v}}^{-1}(t)dt$$

where φ is a given real function.

An integration by part gives :

$$U_{\varphi}\left(\widetilde{v}\right) = \varphi\left(1\right)\mathbb{E}\widetilde{v} - \int_0^1 \varphi'(t) L_{\widetilde{v}}(t)dt$$

If φ is decreasing :

$$\forall t \in [0,1], L_{\widetilde{v}}(t) \leq L_{\widetilde{w}}(t) \quad \text{and} \quad \mathbb{E}\widetilde{v} = \mathbb{E}\widetilde{w} \Rightarrow U_{\varphi}\left(\widetilde{v}\right) \leq U_{\varphi}\left(\widetilde{w}\right)$$

That means that a decreasing φ corresponds to a risk averse behaviour just like a concave utility function does so in the expected utility approach.

If moreover φ is positive (which is necessary if the decision maker has an increasing utility for sure increasing income!) we get :

$$\forall t \in [0,1], L_{\widetilde{v}}(t) \leq L_{\widetilde{w}}(t) \Rightarrow U_{\varphi}(\widetilde{v}) \leq U_{\varphi}(\widetilde{w})$$

The term $U_{\varphi}(\widetilde{v})$ can be rewritten (changing variable) :

$$U_{\varphi}(\widetilde{v}) = \int_a^b x\varphi(F(x))dF(x)$$

Taking ϕ one integral of φ and $\phi(F(x)) = \psi(x)$:

$$U_{\varphi}(\widetilde{v}) = \int_a^b x d\psi(x)$$

If one takes $\phi(0) = 0$ and $\phi(1) = 1$ this amounts to replace the distribution F by the distorted distribution ψ and take the expected value.

Definition 33. RDEU hypothesis. The decision is said to be based on the RDEU hypothesis if there exists a function ϕ mapping $[0,1]$ to$[0,1]$ such that for all income random variables $\widetilde{w}$ and $\widetilde{v}$, $\widetilde{w}$ is preferred to $\widetilde{v}$ if and only if $\int_a^b x\phi'(F_w(x))dF_w(x) \geq \int_a^b x\phi'(F_v(x))dF_v(x)$. When ϕ when is concave, the decision maker is risk averse in the sense that his satisfaction decreases with mean preserving spread

Remark that , when ϕ is concave, the decision-maker overweighs adverse events.

7 Ambiguity

We say that we have ambiguity when the probability distribution is itself fraught with uncertainty. So we can assume that $\widetilde{w}$ follows a law that depends on a θ parameter : $dF(\theta, x)$. The parameter itself is distributed according to a known G. Take χ an increasing function. The utility associated with $\widetilde{w}$ is then defined in the following manner:

$$U_{\varphi,\chi}(\widetilde{v}) = \chi^{-1}\left[\int_{\Theta} \chi\left(\int_a^b x\varphi(F(\theta,x))dF(\theta,x)\right) dG(\theta)\right]$$

or,

$$U_{\varphi,\chi}(\widetilde{v}) = \chi^{-1}\left[\int_{\Theta} \chi\left(\int_0^1 \varphi(u)F^{-1}(\theta,u)du\right) dG(\theta)\right]$$

8 Appendix

8.1 Using Quantile function to sample

A very simple way to generate a sample of a variable with a given probability distribution is to use its quantile function.

If U is a random varialble on $[0, 1]$ uniformly distributed : $\Pr(U \leq x) = x$, and if F is a given CDF then $X \equiv F^{-1}(U)$ is a random variable with CDF F. Indeed :

$$\Pr(X \leq x) = \Pr\left(F^{-1}(U) \leq x\right) = \Pr(U \leq F(x)) = F(x)$$

In order to generate a sample of a random variable, on draw (uniformly) a random number in $[0, 1]$ and apply the quantile function F^{-1}.

8.2 Generalization with copulas

A copula is the CDF of a d-dimensionnal variable $(U_1, U_2, ..., U_d)$on $[0, 1]^d$ with uniform marginal distributions. The analytic characterization is :

1. $C(u_1, u_2, ..., 0, ..u_d) = 0$
2. $C(1, 1, ...u, 1, ..1)) = u$
3. C is d-non-decreasing, i.e., for each hyperrectangle $B = \prod_{i=1}^{d}[x_i,y_i] \subseteq [0, 1]^d$ } the C-volume of B is non-negative:

   $$\int_B dC(u) = \sum_{\mathbf{z} \in \times_{i=1}^{d}\{x_i,y_i\}} (-1)^{N(\mathbf{z})} C(\mathbf{z}) \geq 0,$$

   where $N(\mathbf{z}) = \#\{k : z_k = x_k\}$.

For $d = 2$ the 2 non dcreasing condition amounts to :

$$\Pr(\{v \leq U_1 \leq v', w \leq U_2 \leq w'\}) = [C(v', w') - C(v, w')] - [C(v', w) - C(v, w)] \geq 0$$

The copula gives the structure of "dependence" : given d univariate CDF F_i, and a copula C , then $H(x_1, x_2, ..., x_d) \equiv C(F_1(x_1), F_2(x_2), ..., F_d(x_d))$ is a CDF of a random vector with marginals F_i.

Provided that the marginals are continuous the converse is true :

Sklar's theorem states that every multivariate cumulative distribution function $F(x_1, \ldots, x_d) = \Pr[X_1 \leq x_1, \ldots, X_d \leq x$ of a random vector $(X_1, X_2, \ldots, X_d)$ can be expressed in terms of its marginals $F_i(x_i) = \Pr[X_i \leq x_i]$ and the copula C :$C(u_1, u_2, ...u_d) \equiv F\left(F_1^{-1}(u_1), F_2^{-1}(u_2), ..., F_d^{-1}(u_d)\right)$

Then in order to generate a sample of a random vector X_i with marginals F_i, one generates a sample $(U_1, .., U_d)$ according to C and then take $X_i = F_i^{-1}(U_i)$. .