Introduction to risk theory and mathematical finance

Ł. Stettner*, Institute of Mathematics
Polish Academy of Sciences, Sniadeckich 8, 00-956 Warsaw

1 Introduction

As risk is an inevitable part of all human activity, we naturally seek to be able to better measure it, estimate it, and even reduce it. The notion of risk carries connotations of chaos, the unexpected and undesired behavior of an observed phenomenon. It is difficult to anticipate how a chaotic model will behave, since as the name itself indicates such a model does not have predictable dynamics. But while it is hard to say how a chaotic process will behave at any specific moment, things are quite different if we take a longer-term perspective, chiefly looking at the mean value of certain functions that hinge on the process. For example, the investors of a bank or insurance company are much more interested in the overall market situation than in individual transactions they might stand to lose on. The point is to be sure they come out on top in the appropriate long-term perspective, regardless of short-term fluctuations.

The lecture notes below consists mainly of the extended transparences prepared for the series of lectures held with SSDNM at the Maria Curie-Sklodowska University on November 21, 23 and 30th, 2011.

2 Risk in the history of finance and insurance

2.1 What is risk?

We can say that risk is a failure, an unexpected result. Risk is the probability that a hazard will turn into a disaster.
The probability that a disaster will happen.
Vulnerability and hazards are not dangerous, taken separately. But if they come together, they become a risk.
On the other hand risks can be reduced or managed.

*SSDNM, Maria Curie-Sklodowska University, November 21, 23 and 30th, 2011

Therefore the problem is first to model risk, measure it and then to think about its minimization, its control or management.
From financial point of view we can have two approaches to risk: for one group, let say speculators it is an opportunity to create extra profit (gain), for the other group the main problem is how to eliminate risk in our investments. Consequently there are investors who are looking for risk and investors who are fighting with risk introducing saver (more secure) financial instruments.

We start lectures with the historical part which is mainly based on the data taken from Wikipedia.

2.2 History of insurance and trade

Already 2000 B.C. Egyptian stone workers organized a kind of insurance based on common sharing of funeral costs.
The caravan trade of Asia and the Arabian penisula was subject to various risks and there was a necessity to find a kind of risk sharing guarantee that loosing cargo the trader will be able to survive.
The trade very soon moved from land to sea stimulated construction and development of trading vessels, oars and sails, which again were subject to various risk.
The above factor required suitable codification. In Babylonian times, around 2100 B.C., the Code of Hammurabi contained the system insurance policies used later on by Mediterranean sailing merchants. If a merchant received a loan to fund his shipment he paid also the lender an additional sum in exchange for the lender's guarantee to cancel the loan in the case of the shipment to be stolen.
Greeks and Romans introduced the origin of health and live insurance 600 B.C. creating guilds which cared for the families of decease members and paying funeral expenses.
In Middle Ages the guilds protected their members from loss by fire and shipwreck, paid ransoms to pirates, and provided respectable burials as well as support in times of sickness and poverty
First actual insurance contract was signed in Genoa in 1347. Policies were signed by individuals, either alone or in a group. They each wrote their name and the amount of risk they were willing to assume under the insurance proposal. A new profession has appeared - underwriters, who prepared insurance contracts.
In 1693, the astronomer Edmond Halley created a basis for underwriting life insurance by developing the first mortality table An Estimate of the Degrees of the Mortality of Mankind, drawn from curious Tables of Births and Funerals at the City of Breslau; with an Attempt to ascertain the Price of Annuities upon lives. Edmond Halley English astronomer, geophysicist, mathematician, meteorologist, and physicist who is best known for computing the orbit of the Halley's comet.

In 1756, Joseph Dodson made it possible to scale the premium rate to age.
In the meantime the insuring cargo while being shipped was widespread.
In London Lloyds Coffee House started to be popular place where merchants, ship-owners, and underwriters met to transact their business. Lloyds grew into one of the first modern insurance companies, Lloyds of London.

The first American insurance company was founded in the British colony of Charleston, SC. In 1787 and 1794 respectively, the first fire insurance companies were formed in New York City and Philadelphia. The first American insurance corporation was sponsored by a church the Presbyterian Synod of Philadelphia for their ministers and their dependents. Then other needs for insurance were discovered and, in the 1830s, the practice of classifying risks was begun.
People accepted the fact that they needed to pay premiums to protect themselves and their loved ones in case of loss, including major losses like fires. The insurance companies had a rude awakening to this fact in 1835 when the New York fire struck. The losses were unexpectedly high and they had no reserves prepared for such a situation. As a result of this, Massachusetts lead the states in 1837 by passing a law that required insurance companies to maintain such reserves. The great Chicago fire in 1871 reiterated the need for these reserves, especially in large dense cities.
Insurance companies had to work together to find a solution to the challenge of large losses. So they got together and devised a system called reinsurance whereby losses were distributed among many companies.
The Workmens Compensation Act of 1897 in Britain required employers to insure their employees against industrial mishaps. This also fostered what we know today as public liability insurance, which came strongly into play when the automobile arrived on the scene.
In the 19th century, many societies were founded to insure the life and health of their members
Now insurance was the accepted thing to do. Everybody needed to protect themselves against the many risks in life. Farmers wanted crop insurance. People wanted deposit insurance at their banks. Travelers wanted travel insurance. Everybody turned to insurance companies to give them peace of mind. And really, isnt that what insurance is the paying of a premium to protect against some form of loss.(Gareth Marples)

2.3 History of banking

In Egypt and Mesopotamia gold is deposited in temples for safe-keeping. But it lies idle there, while others in the trading community or in government have desperate need of it.

In Babylon at the time of Hammurabi, in the 18th century BC, there are records of loans made by the priests of the temple. The concept of banking has arrived.
Greek and Roman financiers: from the 4th century BC hey take deposits, make loans, change money from one currency to another and test coins for weight and purity; book transactions. Moneylenders can be found who will accept payment in one Greek city and arrange for credit in another.
In Rome by the 2nd century AD a debt can officially be discharged by paying the appropriate sum into a bank, and public notaries are appointed to register such transactions.
During the 13th century bankers from north Italy, collectively known as Lombards, gradually replace the Jews in their traditional role as money-lenders to the rich and powerful. The business skills of the Italians are enhanced by their invention of double-entry bookkeeping. Creative accountancy enables them to avoid the Christian sin of usury; interest on a loan is presented in the accounts either as a voluntary gift from the borrower or as a reward for the risk taken.
Florence is well equippped for international finance thanks to its famous gold coin, the florin. First minted in 1252, the florin is widely recognized and trusted. It is the hard currency of its day.
By the early 14th century two families in the city, the Bardi and the Peruzzi, have grown immensely wealthy by offering financial services. They arrange for the collection and transfer of money due to great feudal powers, in particular the papacy. They facilitate trade by providing merchants with bills of exchange, by means of which money paid in by a debtor in one town can be paid out to a creditor presenting the bill somewhere else (a principle familiar now in the form of a cheque).

2.4 History of stock exchange

A stock exchange is a corporation or mutual organization which provides the facilities for stock brokers to trade company stocks and other securities. Stock exchanges also provide facilities for the issue and redemption of securities, as well as other financial instruments and capital events including the payment of income and dividends.
In 12th century France the courratiers de change were concerned with managing and regulating the debts of agricultural communities on behalf of the banks. As these men also traded in debts, they could be called the first brokers.
In the late 13th century commodity traders in Bruges gathered inside the house of a man called Van der Bourse, and in 1309 they institutionalized this until now informal meeting and became the "Bruges Bourse". The idea spread quickly around Flanders and neighbouring counties and "Bourses" soon opened in Ghent and Amsterdam.
In the middle of the 13th century Venetian bankers began to trade in government securities. In 1351 the Venetian Government outlawed spreading rumors intended to lower the price

of government funds. There were people in Pisa, Verona, Genoa and Florence who also began trading in government securities during the 14th century. This was only possible because these were independent city states not ruled by a duke but a council of influential citizens.
The Dutch later started joint stock companies, which let shareholders invest in business ventures and get a share of their profits - or losses. In 1602, the Dutch East India Company issued the first shares on the Amsterdam Stock Exchange. It was the first company to issue stocks and bonds.
In 2010, a study published by job search website CareerCast ranked *actuary as the number 1* job in the United States (Needleman 2010). The study used five key criteria to rank jobs: environment, income, employment outlook, physical demands and stress.

3 Modeling in finance and actuary

3.1 Precursors of the mathematical finance and actuary -part 1

We start with two precursors of mathematical actuary:
Ernst Filip Oskar Lundberg (31 December 1876 - 2 June 1965), Swedish actuary, founder of mathematical risk theory and managing director of several insurance companie 1903 doctoral thesis, Approximations of the Probability Function/Reinsurance of Collective Risks, Uppsala
Harald Cramer (September 25, 1893 October 5, 1985), Stockholm University,
They have introduced a basic model of surplus of the insurance company.

3.2 Modeling of a surplus of insurance company

For simplicity we assume first that we a *discrete time* model
The surplus of the company X_n at time n satisfies the following recursive equation
$X_{n+1} = X_n - A_n + Y_{n+1}$
where Y_n represents premia minus payout of claims at time n and it is a sequence of (i.i.d.) random variables with distribution G. The term A_n above stands for an action of the company - possible investments or payout of the dividends.
More popular model is a continuous time model introduced by Lundberg and Cramer called **Cramer-Lundberg (C-L) model** in their papers (Filip Lundberg 1903, 1926, Harald Cramer 1930, 1955). Then the surplus of the company X_t at time t is of the form

$$X_t = x + ct - \sum_{i=1}^{N_t} Y_i \tag{1}$$

where $(N_t) = \sup\left\{n \geq 0 : \sum_{i=1}^{n} T_i \leq t\right\}$ Poisson process with intensity λ (random variables T_i are i.i.d. exponential law with parameter λ), and $0 \leq Y_i$ is again i.i.d. with distribution G. We denote $\mu = EY_i$, We denote by c - the premium rate and as is written above the time between the arrivals of claims (T_i) is exponential with parameter λ. Notice that the process $\sum_{i=1}^{N_t} Y_i$ is called compound Poisson process. (X_t) is an example of the Levy process (the process with stationary and independent increments)
Let $\tau = \inf\{n : X_n < 0\}$ be a bankruptcy, ruin time. As a measure of risk in insurance we frequently consider $P\{\tau \leq T\}$.

3.3 Precursors of the mathematical finance and actuary -part 2

The precursor of mathematics of finance is Louis Bachelier, (born 11 marca 1870 roku, died 26 April 1946 roku), who defended at Sorbonne in 1990 his PhD called "Theorie de la Speculation", later on published in Annales de l'Ecole Normale Superieure (trans. Random Character of Stock Market Prices). Louis Bachelier for the first time used Brownian motion (the process with stationary and independent Gaussian increments) to model fluctuations of asset prices. The way he modeled prices he could have them both positive nad negative. His work was for a long time forgot.
Paul Samuelson 15.05.1915-13.12.2009, 1970 Nobel Prize winner improved Bachelier model considering exponent of Brownian motion and acknowledged the contribution of Louis Bachelier, who after his papers started to be cited.

3.4 Asset prices modeling

We consider to approach to modelin of asset prices:
discrete time: the price of the asset $S_i(t)$, for $i = 1, 2 \ldots, d$ at time t has the following random rate of return ζ

$$\frac{S_i(t+1) - S_i(t)}{S_i(t)} = \zeta_i(z(t+1), \xi(t+1)) \tag{2}$$

which may depend on economical factors $(z(t))$, and i.i.d. noise $\xi(t)$.
An alternative is a *continuous time:* model, where $S_i(t)$ is the price of the ith asset at time t and is of the form

$$S_i(t) = S_i(0) e^{\int_0^t a_i(z(s), S(s)) ds + \int_0^t \sigma_i(z(s), S(s)) \cdot dL(s)} \tag{3}$$

in which L stands for Levy or Wiener process and $(z(s))$ is again a process of economical factors.

Both models are in dynamic, the asset prices change in discrete or continuous time, and we are thinking for a longer time horizon T.

3.5 Precursors of the mathematical finance and actuary -part 3

The notion of risk in financial investments was not recognized among economists for a long time (although all practitioner usually took it into account. The major step was contribution of Harry Markowith in his PHD thesis. Harry Max Markowitz (born 24 August 1927 in Chicago, Illinois) - an American economist and a recipient of the John von Neumann Theory Prize and the Nobel Memorial Prize in Economic Sciences. Markowitz is a professor of finance at the Rady School of Management at the University of California, San Diego (UCSD). He is best known for his pioneering work in Modern Portfolio Theory, studying the effects of asset risk, return, correlation and diversification on probable investment portfolio returns. His main result was published in "Portfolio Selection" The Journal of Finance, Vol. 7, No. 1. (Mar., 1952), pp. 77-91, and later appeared in his book Portfolio Selection: Efficient Diversification of Investments (1959)
Here is what *Harry Markowitz, Nobel prize 1990* told during Nobel price ceremony
The basic concepts of portfolio theory came to me one afternoon in the library while reading John Burr Williams's Theory of Investment Value. Williams proposed that the value of a stock should equal the present value of its future dividends. Since future dividends are uncertain, I interpreted Williams's proposal to be to value a stock by its expected future dividends. **But if the investor were only interested in expected values of securities, he or she would only be interested in the expected value of the portfolio; and to maximize the expected value of a portfolio one need invest only in a single security. This, I knew, was not the way investors did or should act. Investors diversify because they are concerned with risk as well as return. Variance came to mind as a measure of risk.** The fact that portfolio variance depended on security covariances added to the plausibility of the approach. Since there were two criteria, risk and return, it was natural to assume that investors selected from the set of Pareto optimal risk-return combinations.

4 Markowitz portfolio theory

4.1 Portfolio analysis - basics

Let $S_i(t)$ be the price of the ith asset at time t ($i = 1, 2, \dots, d$). Since we consider so called static approach we let $t = 0, 1$.

Random rate of return ith asset is of the form

$$\zeta_i := \frac{S_i(1) - S_i(0)}{S_i(0)} = \frac{S_i(1)}{S_i(0)} - 1$$

The expected rate of return of the ith asset is equal to $\mu_i := E\zeta_i$
We denote by $\mu = [\mu_1, \ldots, \mu_d]^T$ the vector of the expected rates of return, and by $\zeta := [\zeta_1, \ldots, \zeta_d]^T$, the vector of the random rates of return.
Let Σ the covariance matrix of ζ defined as follows:
$\Sigma = E\left\{(\zeta - \mu)(\zeta - \mu)^T\right\} = (\Sigma^{ij})$
with T standing for transponse of the matrix.
Let $\theta := [\theta_1, \ldots, \theta_d]^T$ be the vector of investment strategies (at time 0) - portions of capital invested in d assets, i.e. θ_i portion of capital invested in the i-th asset.
Then **random portfolio rate of return** $R(\theta)$ is equal to $\theta^T\zeta$.
In fact, since: if x is initial capital then $\frac{x\theta_i}{S_i(0)}$ is the amount of purchased ith assets at time 0 and $\frac{x\theta_i}{S_i(0)}S_i(1)$ is the value of the capital invested in the i-th asset at time 1 We clearly have that

$$R(\theta) = \frac{\sum_{i=1}^d \frac{x\theta_i}{S_i(0)}S_i(1) - x}{x} =$$
$$\sum_{i=1}^d \frac{\theta_i}{S_i(0)}S_i(1) - 1 = \sum_{i=1}^d \frac{\theta_i}{S_i(0)}(S_i(1) - S_i(0)) = \sum_{i=1}^d \theta_i\zeta_i$$

Consequently the **expected portfolio rate of return** is of the form

$$E\{R(\theta)\} = \theta^T\mu \tag{4}$$

and the **variance of the portfolio rate of return** is equal to

$$Var(R(\theta)) = \theta^T\Sigma\theta. \tag{5}$$

In fact, we show the last formula - we have

$$Var(R(\theta)) = E\left\{\left(\sum_{i=1}^d \theta_i(\zeta_i - \mu_i)\right)^2\right\} =$$
$$E\left\{(\theta^T(\zeta - \mu))((\zeta - \mu)^T\theta)\right\} = \theta^T E\left\{(\zeta - \mu)(\zeta - \mu)^T\right\}\theta$$

4.2 Markowitz theory - foundations

We maximize expected portfolio rate of return taking into account the risk (minimizing it?)

What is risk?, how to measure portfolio risk?

We introduce the portfolio risk function $Risk(R(\theta))$, which in the Markowitz theory is equal to $Risk(R(\theta)) = Var(R(\theta))$

If fact we have two criterion (minimax) problem: we want ot maximize the expected portfolio rate of return and minimize its variance.

Consequently we introduce so called *Markowitz order* $\succeq$.

The strategy θ (portfolio rate of return $R(\theta)$ is better than θ' (rate of return of $R(\theta')$), we write as $\theta \succeq \theta'$ lub $R(\theta) \succeq R(\theta')$, if $E\{R(\theta)\} \geq E\{R(\theta')\}$, and $Risk(R(\theta)) \leq Risk(R(\theta'))$.

The strategy θ is better than θ' if the rate of return of θ is greater than that of θ' and the risk corresponding to the strategy θ is not greater than the risk corresponding to θ'.

With each strategy one can associate a point on the plane R^2, $(Risk(R(\theta)), E\{R(\theta)\})$

The strategy θ is maximal, if there are no strategy θ' different than θ such that $\theta' \succeq \theta$.

The maximal elements (strategies) in this order form on the plane R^2 the set called **efficient frontier**.

4.3 Classical Markowitz theory - analytic approach

The problem we formulated is a convex analysis problem: to minimize

$$\theta^T \Sigma \theta \tag{6}$$

under the constraints

$\theta^T \mu = \mu_p$ i $\theta^T \mathcal{J} = 1$,

where $\mathcal{J} = [1, \ldots, 1]^T$, while μ_p is the fixed portfolio expected rate of return. For fixed portfolio rate of return we minimize risk understood as portfolio variance. Such problem can be solved using the results of the book

Markowitz, H., Portfolio Selection Efficient Diversification of Investments, Wiley, 1959.

When the matrix Σ is nonsingular, the problem can be is solved using Lagrange multipliers.

Consider the function

$G(\theta, \kappa) = \theta^T \Sigma \theta + \kappa_1(\theta^T \mu - \mu_p) + \kappa_2(\theta^T \mathcal{J} - 1)$

A necessary condition for optimality is of the form

$\frac{\partial G}{\partial \theta_i} = 0 \quad \frac{\partial G}{\partial \kappa_j} = 0$ dla $j = 1, 2$

We have

Theorem 1 *If μ does not have the same coordinates and Σ is nonsingular then*

$$\inf_{\theta} Var(R(\theta)) = B^T H^{-1} B, \tag{7}$$

where $H = A^T\Sigma^{-1}A,\ A = [\mu, \mathcal{J}],\ B = \begin{bmatrix} \mu_p \\ 1 \end{bmatrix}.$

Proof. By necessary conditions with $\kappa := [\kappa_1, \kappa_2]^T$ we have

$$2\Sigma\theta + A\kappa = 0 \quad A^T\theta = B \tag{8}$$

we solve the first equation with respect to θ:
$\theta = \frac{-1}{2}\Sigma^{-1}A\kappa$
From the second equation of (8) we have
$A^T\Sigma^{-1}A\kappa = -2B$
i.e. assuming that H is invertible we have
$\kappa = -2(A^T\Sigma^{-1}A)^{-1}B := -2H^{-1}B.$
Note that H is symmetric:
$H = A^T\Sigma^{-1}A$ therefore $H^T = A^T\Sigma^{-1T}A = H$ since Σ and Σ^{-1} are symmetric.
We are now in position to calculate the portfolio variance using the obtained formulae for θ i κ
$\theta^T\Sigma\theta = \frac{-1}{2}\theta^T\Sigma\Sigma^{-1}A\kappa = \theta^T AH^{-1}B = (A^T\theta)^T H^{-1}B = B^T H^{-1}B$
which completes the proof, assuming that H is invertible.
Let $H = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, and $\upsilon = detH$. H is 2×2 matrix and $a = \mu^T\Sigma^{-1}\mu$, $b = \mu^T\Sigma^{-1}\mathcal{J}$, $c = \mathcal{J}^T\Sigma^{-1}\mathcal{J}$.
If *matrix Σ is nonsingular and all coordinates of μ are nonidentical then* $\upsilon > 0$.
In fact, from the definition of H we have
$a = \mu^T\Sigma^{-1}\mu$, $b = \mu^T\Sigma^{-1}\mathcal{J} = \mathcal{J}^T\Sigma^{-1}\mu$, $c = \mathcal{J}^T\Sigma^{-1}\mathcal{J}$
also $\upsilon = ac - b^2$.
Since Σ is positive definite ($y^T\Sigma y > 0$ when $y \neq 0$), the matrix Σ^{-1} is also positive definite i.e. $a > 0$ whenever $\mu \neq 0$ and $c > 0$.
Furthermore
$(b\mu - a\mathcal{J})^T\Sigma^{-1}(b\mu - a\mathcal{J}) = bba - abb - abb - aac = a(ac - b^2) = a\upsilon > 0$
whenever μ does not have the same coordinates. The proof of Theorem 1 is completed.
□

The formula (7) characterizes the efficient frontier for variance as a measure of risk. This will be the upper part of the parabola in coordinate system $(Risk(R(\theta)), E\{R(\theta)\})$
Using matrix $H = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ we have

$$\min_\theta Var(R(\theta)) = \frac{1}{\upsilon}(c\mu_p^2 - 2b\mu_p + a) \tag{9}$$

which gives us formula for the above mentioned parabola. The coordinates of the origin of the parabola are of the form:
$\mu_{mv} = \frac{b}{c}$ and $Var(R(\theta_{mv})) = \frac{1}{c}$.

4.4 Optimal strategy

The formula for optimal strategy obtained in the proof of Theorem 1 is

$$\theta = \frac{-1}{2}\Sigma^{-1}A\kappa = \Sigma^{-1}AH^{-1}B \tag{10}$$

Since

$$H^{-1} = \frac{1}{v}\begin{bmatrix} c & -b \\ -b & a \end{bmatrix} \tag{11}$$

(note that $A = [\mu, \mathcal{J}]$, $B = \begin{bmatrix} \mu_p \\ 1 \end{bmatrix}$). we finally obtain

$$\theta_{opt} = \frac{1}{v}\Sigma^{-1}A\begin{bmatrix} c\mu_p - b \\ -b\mu_p + a \end{bmatrix} = \frac{1}{v}\Sigma^{-1}\left(\mu(c\mu_p - b) + \mathcal{J}(-b\mu_p + a)\right) =$$
$$\frac{1}{v}\Sigma^{-1}\left((a\mathcal{J} - b\mu) + (c\mu - b\mathcal{J})\mu_p\right)$$

which is the formula for an optimal strategy.

4.5 Minimal variance portfolio

The formula for optimal strategy obtained above is parametrized by expected portfolio rate of return μ_p i.e. we have

$$\theta_{opt} = \frac{1}{v}\Sigma^{-1}\left((a\mathcal{J} - b\mu) + (c\mu - b\mathcal{J})\mu_p\right) \tag{12}$$

If now $\mu_p = \frac{b}{c} = \mu_{mv}$, which corresponds to the vertex of efficient frontier parabola we obtain that
$\theta_{opt} = \frac{1}{v}\Sigma^{-1}\left(a - \frac{b^2}{c}\right)\mathcal{J} = \Sigma^{-1}\mathcal{J}\frac{1}{c}$
is the strategy minimizing variance we denote θ_{mv} for which we have $Var(R(\theta_{mv})) = \frac{1}{c}$.

4.6 Sharpe coefficient - tangent portfolio

We study now the approach considered by another 1990 Nobel prize winner (*William F. Sharpe (born in Boston in 1934).*
Minimal variance strategy usually provides relatively small portfolio rate of return. Hence it is quite natural to look for other portfolios from the efficient frontier.
One opportunity is to maximize so called Sharp coefficient

$$\frac{\mu_p}{\sqrt{Var(R(\theta))}}, \tag{13}$$

We are looking for the greatest intersection coefficient of the line starting from the origin with efficient frontier (upper part of hiperbola) i.e. for linear part of the line tangent to efficient frontier (upper part of hiperbola) starting from the origin (we tacitly assume here that $\mu_{mv} > 0$)
In fact, $Var(R(\theta_{opt})) = \frac{1}{\upsilon}(c\mu_p^2 - 2b\mu_p + a)$. Hence $\sqrt{Var(R(\theta_{opt}))} = \sqrt{\frac{1}{\upsilon}(c\mu_p^2 - 2b\mu_p + a)}$
To find optimal tangent portfolio we calculate

$$\frac{d\mu_p}{d\sqrt{Var(R(\theta_{opt}))}} = \left(\frac{d\sqrt{Var(R(\theta_{opt}))}}{d\mu_p}\right)^{-1} =$$
$$\left(\frac{1}{2\sqrt{\frac{1}{\upsilon}(c\mu_p^2 - 2b\mu_p + a)}}\frac{1}{\upsilon}(2c\mu_p - 2b)\right)^{-1} = \frac{\upsilon\sqrt{\frac{1}{\upsilon}(c\mu_p^2 - 2b\mu_p + a)}}{c\mu_p - b}$$

We would like to have $\frac{d\mu_p}{d\sqrt{Var(R(\theta_{opt})}} = \frac{\mu_p}{\sqrt{Var(R(\theta_{opt}))}}$, i.e.
$\mu_p(c\mu_p - b) = \upsilon Var(R(\theta_{opt})) = c\mu_p^2 - 2b\mu_p + a$ which gives $b\mu_p = a$ and

$$\mu_p = \frac{a}{b} = \mu_{tg}. \tag{14}$$

If $\mu_{tg} = \frac{a}{b}$ we have
$Var(R(\theta_{tg})) = \frac{1}{\upsilon}(c\left(\frac{a}{b}\right)^2 - 2b\frac{a}{b} + a) = \frac{1}{\upsilon}(\frac{ca^2}{b^2} - a) = \frac{a}{b^2}$.
Consequently using (9) we obtain

$$\theta_{tg} = \frac{1}{\upsilon}\Sigma^{-1}\left(a\mathcal{J} - b\mu) + (c\mu - b\mathcal{J})\frac{a}{b}\right) = \Sigma^{-1}\frac{\mu}{b}. \tag{15}$$

We have looked for the line tangent to $\sqrt{Var(R(\theta_{opt}))}$. One could could also look for the line tangent to the graph of $Var(R(\theta_{opt}))$, which we shall study in the next subsection.

4.7 Maximization of the tangency coefficient to the efficient frontier parabola

Now we want to maximize

$$\frac{\mu_p}{Var(R(\theta))} \tag{16}$$

i.e. tangency coefficient to the upper part of efficient frontier parabola
$c\mu_p^2 - 2b\mu_p + a - \upsilon Var(R(\theta_{opt})) = 0$.
The upper part of the parabola is given by $\mu_p = \frac{2b+\sqrt{\Delta}}{2c}$ with $\Delta = 4b^2 - 4c(a - \upsilon Var(R(\theta_{opt}))) = 4\upsilon(cVar(R(\theta_{opt})) - 1)$
The tangency coefficient should satisfy
$\frac{d\mu_p}{dVar(R(\theta_{min}))} = \frac{1}{2c}\frac{1}{2\sqrt{\Delta}}4\upsilon c = \frac{\upsilon}{\sqrt{\Delta}}$
and therefore the tangent line is of the form $\mu = \frac{\upsilon}{\sqrt{\Delta}}Var + z$.
It intersect the origin when $z = 0$. We look for μ and Var from the upper parabola part i.e. $2c\mu = 2b + \sqrt{\Delta}$ and $\mu\sqrt{\Delta} = \upsilon Var$, which gives $2c\mu^2 = 2b\mu + \upsilon Var$. Therefore $c\mu^2 + a - 2c\mu^2 = 0$, and $\mu = \sqrt{\frac{a}{c}} = \mu_{st}$.
Hence
$Var(R(\theta_{st})) = \frac{1}{\upsilon}(c\left(\frac{a}{c}\right)^2 - 2b\frac{a}{c} + a) = \frac{1}{\upsilon}(\frac{a^2}{c} - 2\frac{ba}{c} + a) = \frac{1}{\upsilon c}a(a - 2b + c)$
and
$\theta_{st} = \frac{1}{\upsilon}\Sigma^{-1}\left((a\mathcal{J} - b\mu) + (c\mu - b\mathcal{J})\frac{a}{c}\right) = \frac{1}{\upsilon}\Sigma^{-1}\Big(a(1 - \frac{b}{c})\mathcal{J} + (a - b)\mu\Big)$.

4.8 The role of minimal variance and tangent portfolios

The optimal strategy is of the form
$\theta_{opt} = \frac{1}{\upsilon}\Sigma^{-1}\left((a\mathcal{J} - b\mu) + (c\mu - b\mathcal{J})\mu_p\right)$
The optimal tangent strategy $\theta_{tg} = \Sigma^{-1}\frac{\mu}{b}$ and the optimal minimum variance strategy is $\theta_{mv} = \Sigma^{-1}\mathcal{J}\frac{1}{c}$.
Hence

$$\theta_{opt} = \frac{-b^2 + bc\mu_p}{\upsilon}\theta_{tg} + \frac{ac - bc\mu_p}{\upsilon}\theta_{mv} \tag{17}$$

which means that **optimal strategy is a linear (affine) combination of tangent and minimal variance strategies**.

4.9 Value functionals

Two cost functional problem (minmax) we considered so far leads us to an idea to *replace two criterions problem by one criterion.*
We maximize two parameter functional

$$F\left(E\left\{R(\theta)\right\},Risk(R(\theta))\right) \tag{18}$$

over all admissible portfolio strategies θ. The choice of the point of efficient frontier is replaced by the choice of risk parameter λ.
The function F should be increasing with respect to the first coordinate and decreasing with respect to the second. The risk aversion is measured by the parameter $\lambda \geq 0$. The most natural form of the function F is

$$F(x,y) = x - \frac{1}{2}\lambda y. \tag{19}$$

We maximize $F\left(E\left\{R(\theta)\right\},Var(R(\theta))\right) = E\left\{R(\theta)\right\} - \frac{1}{2}\lambda Var(R(\theta)) = \theta^T\mu - \frac{1}{2}\lambda\theta^T\Sigma\theta$ with respect to such theta θ that $\theta^T\mathcal{J} = 1$. Again we use Lagrange multipliers method. We form the function
$G(\theta,\kappa) = \theta^T\mu - \frac{1}{2}\lambda\theta^T\Sigma\theta + \kappa(\theta^T\mathcal{J} - 1)$
Necessary condition of optimality is:
$\frac{\partial G}{\partial\theta_i} = 0 \quad \frac{\partial G}{\partial\kappa} = 0$
We have

Theorem 2 *If μ does not have the same coordinates and matrix Σ is nonsingular then*

$$\theta_{opt} = \frac{1}{\lambda}\Sigma^{-1}\left(\mu + \mathcal{J}\frac{\lambda - b}{c}\right). \tag{20}$$

Proof.
From the necessary condition we get $\mu - \lambda\Sigma\theta + \mathcal{J}\kappa = 0$ and $\mathcal{J}^T\theta = 1$. We solve the first equation with respect to θ and obtain
$\theta = \frac{1}{\lambda}\left(\Sigma^{-1}\left(\mu + \mathcal{J}\kappa\right)\right)$
Substituting this to the second equation we obtain
$\frac{1}{\lambda}\mathcal{J}^T\Sigma^{-1}\left(\mu + \mathcal{J}\kappa\right) = 1.$
Hence (using nonsingularity of Σ) we have
$\kappa = \frac{\lambda - \mathcal{J}^T\Sigma^{-1}\mu}{\mathcal{J}^T\Sigma^{-1}\mathcal{J}} = \frac{\lambda - b}{c}.$
Recalling that $c = \mathcal{J}^T\Sigma^{-1}\mathcal{J}$ i $b = \mathcal{J}^T\Sigma^{-1}\mathcal{J}$ we obtain
$\theta_{opt} = \frac{1}{\lambda}\Sigma^{-1}\left(\mu + \mathcal{J}\frac{\lambda - b}{c}\right) = \frac{b}{\lambda}\theta_{tg} + (1 - \frac{b}{\lambda})\theta_{mv}.$
Since $\theta_{tg} = \Sigma^{-1}\mu\frac{1}{b}$ and $\theta_{mv} = \Sigma^{-1}\mathcal{J}\frac{1}{c}$ we finally have

$$\theta_{opt} = \frac{b}{\lambda}\theta_{tg} + (1 - \frac{b}{\lambda})\theta_{mv} \tag{21}$$

which is an analogy (14) for Markowitz model.
When $\lambda = b$ ($b = \mathcal{J}^T\Sigma^{-1}\mu$) we have $\theta_{opt} = \theta_{tg}$ and when $\lambda \to \infty$ we have $\theta_{opt} \to \theta_{mv}$.

Furthermore $\mu_{opt} = \theta_{opt}^T \mu = \frac{1}{\lambda}\mu^T \Sigma^{-1}\mu + \frac{1}{\lambda}\mathcal{J}^T \Sigma^{-1}\mu \frac{\lambda - b}{c} = \frac{a}{\lambda} + \frac{b}{\lambda}\frac{\lambda - b}{c} = \frac{a}{\lambda} + \frac{b}{c} - \frac{b^2}{c\lambda} = \frac{\upsilon}{c\lambda} + \mu_{mv}$ since $a = \mu^T \Sigma^{-1}\mu$, $b = \mathcal{J}^T \Sigma^{-1}\mu$, and $\mu_{mv} = \frac{b}{c}$. Furthermore (since $c = \mathcal{J}^T \Sigma^{-1}\mathcal{J}$ and $Var(R(\theta_{mv})) = \frac{1}{c}$) we have

$$Var(R(\theta_{opt})) = \theta_{opt}^T \Sigma \theta_{opt} = \frac{1}{\lambda}(\mu^T + \mathcal{J}^T \frac{\lambda - b}{c})\Sigma^{-1}\Sigma \frac{1}{\lambda}\Sigma^{-1}(\mu + \mathcal{J}\frac{\lambda - b}{c}) =$$
$$\frac{1}{\lambda^2}(a + 2b\frac{\lambda - b}{c} + (\frac{\lambda - b}{c})^2 c) = \frac{1}{\lambda^2}(a + \frac{2b\lambda - 2b^2 + \lambda^2 - 2b\lambda + b^2}{c}) =$$
$$\frac{1}{\lambda^2}(a + \frac{\lambda^2 - b^2}{c}) = \frac{\upsilon}{c\lambda^2} + \frac{1}{c} = \frac{\upsilon}{c\lambda^2} + Var(R(\theta_{mv})).$$

4.10 Form of the variance corresponding to the optimal strategy

We have $\mu_{opt} = \frac{\upsilon}{c\lambda} + \mu_{mv}$, $\mu_{mv} = \frac{b}{c}$ and $Var(R(\theta_{mv})) = \frac{1}{c}$.
Moreover

$$Var(R(\theta_{opt})) = \frac{\upsilon}{c\lambda^2} + Var(R(\theta_{mv})) = \frac{(\mu_{opt} - \mu_{mv})^2 c}{\upsilon} + Var(R(\theta_{mv})) =$$
$$\frac{1}{\upsilon}\left(c\mu_{opt}^2 - 2b\mu_{opt} + \frac{b^2}{c} + \frac{\upsilon}{c}\right) = \frac{1}{\upsilon}\left(c\mu_{opt}^2 - 2b\mu_{opt} + a\right)$$

This is an analogy to (9) ($\min_\theta Var(R(\theta)) = \frac{1}{\upsilon}(c\mu_p^2 - 2b\mu_p + a)$) obtained for Markowitz model.
We now can summarize the obtained results:
Summary: *Using optimal strategies for* $F(E\{R(\theta)\}, Var(R(\theta))) = E\{R(\theta)\} - \frac{1}{2}\lambda Var(R(\theta)) = \theta^T\mu - \frac{1}{2}\lambda\theta^T\Sigma\theta$ *for different* λ *we obtain the whole part of the efficient frontier from the origin of the parabola (minimal risk point) to the Sharpe point (tangency point with the effective frontier hiperbola).*

4.11 Comments and remarks

We complete this part of the lecture with row comments:
1. The methodology considered (both Markowitz and one criterion aim functional) does not impose any constraints on the portfolio strategies (we admit both short selling and short borrowing). Vector θ can admit arbitrary values (both positive and negative but $\theta^T\mathcal{J} = 1$); If we are looking for nonnegative strategies we have to study (in a number of models) further part of efficient frontier (using continuity of the aim functional);
2. Function
$F_\lambda(E\{R(\theta)\}, Var(R(\theta))) = E\{R(\theta)\} - \frac{1}{2}\lambda Var(R(\theta))$

describes a constant risk aversion - the derivative of this function with respect to λ is constant. An alternative approach leads to study risk sensitive functionals of the form $F_\lambda(R(\theta)) = \frac{-1}{\lambda} \ln E\left\{\exp\left\{-\lambda R(\theta)\right\}\right\}$, the idea of which is explained in the next section.

5 Risk sensitive portfolio

5.1 Motivation

Let $g(\lambda) = \ln E\left\{\exp\left\{-\lambda X\right\}\right\}$. By Taylor expansion of this function we have
$g(\lambda) = g(0) + \lambda g'(0) + \frac{\lambda^2}{2} g''(0) + remainder(\lambda)$
where $g'(\lambda) = \frac{E\left\{-Xe^{-\lambda X}\right\}}{E\left\{e^{-\lambda X}\right\}}$ and
$g''(\lambda) = \frac{E\left\{X^2 e^{-\lambda X}\right\}E\left\{e^{-\lambda X}\right\} - \left(E\left\{-Xe^{-\lambda X}\right\}\right)^2}{\left(E\left\{e^{-\lambda X}\right\}\right)^2}$.
Hence
$\frac{-1}{\lambda} g(\lambda) = E\left\{X\right\} - \frac{1}{2}\lambda Var(X) + remainder(\lambda)$, and therefore
$F_\lambda(R(\theta)) = \frac{-1}{\lambda} \ln E\left\{\exp\left\{-\lambda R(\theta)\right\}\right\} = E\left\{R(\theta)\right\} - \frac{1}{2}\lambda Var(R(\theta)) + remainder(\lambda)$
Consequently maximizing $F_\lambda(R(\theta))$ we maximize $E\left\{R(\theta)\right\}$ and minimize (with a weight $\frac{1}{2}\lambda$) $Var(R(\theta))$ and in fact also higher moments (but with also higher powers of λ).

5.2 Risk sensitive stationary portfolio

We study further risk sensitive cost functional introduced above
$F_\lambda(R(\theta)) = \frac{-1}{\lambda} \ln E\left\{\exp\left\{-\lambda R(\theta)\right\}\right\}$
We have the following two results:
Fact 1. $\theta \mapsto F_\lambda(R(\theta))$ is (strictly) concave (Hölder inequality). Consequently there is at most one maximum point of $\theta \mapsto F_\lambda(R(\theta))$ $\theta^T \mathcal{J} = 1$.
Fact 2.(Jensen inequality) We have $F_\lambda(R(\theta)) \leq E\left\{\theta^T \zeta\right\}$.
How to solve the maximization problem of $\theta \mapsto F_\lambda(R(\theta))$ under the constraint $\theta^T \mathcal{J} = 1$.
Natural suggestion would be to use Lagrange multiplier's method introducing the function
$G(\theta, \kappa) = \frac{-1}{\lambda} \ln E\left\{\exp\left\{-\lambda R(\theta)\right\}\right\} + \kappa(\theta^T \mathcal{J} - 1)$.
Necessary condition of optimality is then :
$\frac{\partial G}{\partial \theta_i} = 0 \quad \frac{\partial G}{\partial \kappa} = 0$,
from which we obtain:
$\frac{-1}{\lambda} \frac{E\left\{\exp\left\{-\lambda\theta^T\zeta\right\}(-\lambda)\zeta^i\right\}}{E\{\exp\{-\lambda\theta^T\zeta\}\}} + \kappa = 0$ and $\mathcal{J}^T\theta = 1$
We would like to find $\theta(\kappa)$ satisfying the first equation. Is it possible? In general it is not possible. We show in the next subsection a counterexample.

5.3 Counterexample

Assume we have two assets i.e. $d = 2$. Then $\theta_2 = 1-\theta_1$. Assume furthermore that random rates of return ζ_i are independent with two values only
$\zeta_1 \in \{a_1, b_1\}$,
$\zeta_2 \in \{a_2, b_2\}$ where $-1 < a_1 < 0 < b_1$, $-1 < a_2 < 0 < b_2$, so that we have no arbitrage opportunity (profit without risk - this notion will be studied in details later). Moreover let $a_1 < a_2$, $b_1 > b_2$. Then we have
$-\lambda\theta_1(b_2 - a_1) \preceq F_\lambda(R(\theta)) \preceq \lambda\theta_1(b_1 - a_2)$ where $\preceq$ means an order of approximation. As a result we have (since there are no restriction on θ_1 it may take any positive or negative value)
Summary: $\sup_\theta F_\lambda(R(\theta)) = \infty$,
This means that the Lagrange multipliers method can not be used.
However, when we restrict ourselves to $\theta_i \geq 0$ everything is fine, although it may be difficult to find an explicit solution. Therefore we use approximate methods to find an optimal (unique) portfolio.

6 Markowitz problem with semivariance

6.1 Problem formulation

Variance of the portfolio as a measure of risk can not be considered as the best measure. It penalizes both positive positions of the portfolio (above the mean) as negative position (below the mean). Positive position are good for an investor. In fact risk corresponds only to the negative positions. This justification leads to considering semivariance ($sVar$ defined below) as a proper measure of risk. Namely we consider the risk of the form
$Risk(R(\theta)) = SVar(R(\theta)) = E\left\{\left(\sum_{i=1}^d \theta_i(\zeta_i - \mu_i)\right)^{-2}\right\}$
We would like now to minimize $E\left\{\left(\sum_{i=1}^d \theta_i(\zeta_i - \mu_i)\right)^{-2}\right\}$
under constraints $\theta^T\mu = \mu_p$ and $\theta^T\mathcal{J} = 1$.
Denote by $r_i = \zeta_i - \mu_i$. Clearly $Er_i = 0$ and $\theta_1 = 1 - \sum_{i=1}^d \theta_i$, and the problem is to minimize
$E\left\{\left(\sum_{i=1}^d \theta_i r_i\right)^{-2}\right\} = E\left\{\left(r_1 + \sum_{i=2}^d \theta_i(r_i - r_1)\right)^{-2}\right\}$
under constraints $\sum_{i=2}^d \theta_i(\mu_i - \mu_1) = \mu_p - \mu_1$.
In what follows we shall consider two cases:
Case 1. $\mu_i = \mu_1$, for each i; then $\mu = \mu_p$, and the problem leads to unconstraint minimization

$\min_{(\theta_2,\ldots,\theta_d)} E\left\{\left(r_1+\sum_{i=2}^d \theta_i(r_i-r_1)\right)^{-2}\right\}$
Case 2. $\mu_i \neq \mu_1$ for some i;
Before we continue we need to solve an auxiliary problem.

6.2 An auxiliary lemma

Consider the problem Problem (A): $\min_{x\in R^m} E\left\{(A+B^T x)^-\right\}^2$
where A and B are random variables and vectors respectively such that $B=(B_1,\ldots,B_m)^T$, $EB_i^2<\infty$, $EB_i=0$ and $EA^2<\infty$. We have

Lemma 1 *The problem (A) admits an optimal solution.*

Proof.
Assume first that $B_1,\ldots,B_m$ are linearly independent $P(\sum_{i=1}^m \alpha_i B_i=0)=1$ implies $\alpha_1=\ldots=\alpha_m=0$.
Let $S=\{(k,y)\in R^{m+1}: k\in[0,1], \|y\|=1\}$ and $c=\inf_{(k,y)\in S} E[(kA+B^T y)^-]^2$.
Clearly the mapping $(k,y)\to E[(kA+B^T y)^-]^2$ is continuous so that there is (k^*,y^*) such that $c=E[(k^*A+B^T y^*)^-]^2$. If $c=0$ we have $k^*>0$ since otherwise $E[(B^T y^*)^-]^2=0$ and $P(B^T y^*\geq 0)=1$, $EB^T y^*=0$ and finally $P(B^T y^*=0)=1$, $y^*=0$ but $\|y^*\|=1$ a contradiction.
Therefore $c=0\Rightarrow k^*>0$ and then $\frac{y^*}{k^*}$ is an optimal solution to the Problem (A).
When $c>0$ then for $\|x\|\geq 1$ we have
$E[(A+B^T x)^-]^2=\|x\|^2 E[(\frac{A}{\|x\|}+B^T\frac{x}{\|x\|})^-]^2\geq \|x\|^2 c$
so that we have coercivity of $x\to E\left\{(A+B^T x)^-\right\}^2$, and since it is a continuous function it admits a minimizer x^*.
Assume now that $\{B_1,\ldots,B_m\}$ are not independent. If $P(B=0)=1$ and $x\in R^m$ is a minimizer. If $P(B\neq 0)>0$ there is a subset D of $\{B_1,\ldots,B_m\}$ whose elements are linearly independent and every element in this set is a linear combination of D. Suppose that such subset is $\{B_1,\ldots,B_k\}$ and let $\tilde{B}=(B_1,\ldots,B_k)^T$. By the proof there is a minimizer $\tilde{x}^*$ for $x\to E\left\{(A+\tilde{B}^T x)^-\right\}^2$ and $x^*=(\tilde{x}^*,0,\ldots,0)^T$ is a minimizer. This completes the proof of lemma.

□

6.3 Semivariance main result

We are now in position to prove the main result

Theorem 3 *There is a minimizer to*
$E\left\{\left(\sum_{i=1}^d \theta_i r_i\right)^{-2}\right\} = E\left\{\left(r_1 + \sum_{i=2}^d \theta_i (r_i - r_1)\right)^{-2}\right\}$
under constraints
$\sum_{i=2}^d \theta_i(\mu_i - \mu_1) = \mu_p - \mu_1$

Proof. We continue to study the cases considered earlier: *Case 1.* $\mu_i = \mu_1$ (continuation): by Lemma 1 we have a minimizer to
$\min_{(\theta_2,\ldots,\theta_d)} E\left\{\left(r_1 + \sum_{i=2}^d \theta_i (r_i - r_1)\right)^{-2}\right\}$
Case 2. Assume that there is i such that $\mu_1 \neq \mu_i$. For simplicity let $i = 2$. We then have
$\theta_2 = \frac{\mu_p - \mu_1}{\mu_2 - \mu_1} - \sum_{i=3}^d \theta_i \frac{(\mu_i - \mu_1)}{\mu_2 - \mu_1}$
and therefore the problem is reduced to min over $(\theta_3, \ldots, \theta_d)$ of
$E\left\{\left(r_1 + \frac{\mu_p - \mu_1}{\mu_2 - \mu_1}(r_2 - r_1) + \sum_{i=2}^d \theta_i (r_i - r_1) - (r_2 - r_1)\frac{(\mu_i - \mu_1)}{\mu_2 - \mu_1}\right)^{-2}\right\}$
and by Lemma 1 again there is a minimizer.

6.4 Semivariance result generalizations

The result on semivariance can extended to the following general downside risk functions (instead of semiviarance): $f(x) = 0$ for $x \geq 0$, and $f(x) > 0$ for $x < 0$.
As an example one can consider the function $f(x) = (x^-)^p$ with $p > 0$.
It can be shown that whenever general downside risk function f is lower semicontinuous (lsc) and $f(kx) \geq g(k) f(x)$ and $\lim_{x\to\infty} g(x) = \infty$ then the problem
$E\left\{f\left(\left(\sum_{i=1}^d \theta_i(\zeta_i - \mu_i)\right)^-\right)\right\} \to \min$
with constraints $\sum_{i=2}^d \theta_i(\mu_i - \mu_1) = \mu_p - \mu_1$ has a minimal solution.
For details see Jin, H., Markowitz, H., Zhou, XY, A note on semivariance, Math. Fin. 16 (2006), 53-61.
We complete this part of lecture with the list of most important references.

6.5 References:

De Giorgi, E., – A Note on Portfolio Selection under Various Risk Measures, Working paper, University of Zurich 2002.
Elton, E. J., Gruber, M. J. – Modern Portfolio Theory and Investment Analysis, Wiley, 1981.
Jin, H., Markowitz, H., Zhou, XY, A note on semivariance, Math. Fin. 16 (2006), 53-61
Markowitz, H., – Portfolio Selection Efficient Diversification of Investments, Wiley, 1959.
Sharpe, W. F., – Portfolio Theory and Capital Markets, McGraw-Hill, 1970.

Szegö, G. P., – Portfolio Theory with Application to Bank Asset Management, Academic Press, 1980.
Whittle, P., – Risk-sensitive Optimal Control, Wiley, 1990.

7 Monetary Measures of Risk

Variance of the portfolio rate of return or semivariance were just examples of of measure of risk. In this section we start to consider a general approach to measures of risk. Our consideration will be then continued in the final section 12.

7.1 Definitions

We denote by *financial position* $X : \Omega \to R$ net worth position at maturity. and let $\mathcal{X}$ be the class of financial positions.

Definition 1 *A function* $\rho : \mathcal{X} \to R$ *is* **monetary measure of risk** *when:*
- *if* $X \leq Y$ *we have* $\rho(X) \geq \rho(Y)$ *(monotonicity)*
- *if* $m \in R$ *then* $\rho(X + m) = \rho(X) - m$ *(translation invariance; cash additivity)*

Directly from the definition we have **properties:**
$\rho(X + \rho(X)) = 0$
$|\rho(X) - \rho(Y)| \leq \|X - Y\|$
where $\|X\|$ stands for L^∞ norm. We also frequently use the following normalization $\rho(0) = 0$.

Definition 2 *Monetary measure of risk is* **convex measure** *whenever*
$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y)$ *for* $\lambda \in [0, 1]$

The meaning of convex risk measure is that diversification does not increase the risk.

Definition 3 *Convex measure of risk is called a* **coherent measure of risk** *when*
if $\lambda \geq 0$ *we have* $\rho(\lambda X) = \lambda\rho(X)$ *(positive homogeneity).*

We have the following **properties** of the coherent measures of risk:
$\rho(0) = 0$
$\rho(X + Y) \leq \rho(X) + \rho(Y)$ (subadditivity).

7.2 Acceptance sets

Given a monetary measure of risk ρ we define so called **acceptance set** as $\mathcal{A}_\rho := \{X \in \mathcal{X} | \rho(X) \leq 0\}$.
Clearly:
ρ convex iff $\mathcal{A}_\rho$ is convex,
ρ is positively homogeneous iff $\mathcal{A}_\rho$ is a cone.
$\rho(X) = \inf \{m \in R | m + X \in \mathcal{A}_\rho\}$.
A particular example of coherent risk measures is
worst case measure $\rho_{max}(X) = -essinf_\omega X(\omega)$ ($\mathcal{A}_\rho$ is the the set of all nonnegative bounded random variables).
For every monetary measure of risk ρ we have $\rho(X) \leq \rho(essinf_\omega X(\omega)) = \rho_{max}(X)$, so that ρ_{max} is a maximal monetary measure of risk.
Notice furthermore that $\rho(X) = -EX$ is also a coherent measure of risk.
Also, $\rho(X) = \sup_{Q \in \mathcal{Q}} \{E_Q \{-X\}\}$ with $\mathcal{Q}$ a family of absolutely continuous measures with respect to P, is a coherent measure of risk.

7.3 Value at Risk

The Basel Committee on Banking Supervision (BCBS) (a committee of banking supervisory authorities that was established by the central bank governors of the Group of Ten countries in 1975) recommended as a measure of risk for financial investments so called value at risk (VaR) defined as follows for a given α which is usually less that 0.05 (in fact Basel II recommendation is $\alpha = 0.01$)

$$VaR_\alpha(X) = \inf \{m \in R | P\{m + X < 0\} \leq \alpha\}.$$

Therefore $VaR_\alpha(X)$ is a minimal capital m that added to our financial position X allows us to keep the ruin probability (i.e. the probability of negative value of $X + m$ not greater than α.
In statistics we introduce so called quantiles of random variables. For a given $\alpha \in (01,)$ we have upper and lower quantiles defined as follows:
α quantiles: upper is
$q^+_\alpha(X) := \inf \{x \in R : P\{X \leq x\} > \alpha\} = \sup \{x \in R : P\{X < x\} \leq \alpha\}$
lower quantile is
$q^-_\alpha(X) := \inf \{x \in R : P\{X \leq x\} \geq \alpha\} = \sup \{x \in R : P\{X \leq x\} < \alpha\}$
Usually lower and upper quantiles coincide: this is in the case of continuous random variables (i.e. random variables X with continuous distribution function $F_X(x) = P\{X \leq x\}$.
In general $q^-_\alpha(X) \leq q^+_\alpha(X)$ and we have a sharp inequality whenever $P\{X = \alpha\} > 0$. In the last case we have an α quantiles interval $[q^-_\alpha(X), q^+_\alpha(X)]$.

Notice that
$P\{-X \leq m\} \geq 1-\alpha$ and $P\{X < -m\} \leq \alpha$.
Therefore we have the following equivalent formula for value at risk $VaR_\alpha(X) = q^-_{1-\alpha}(-X) = -q^+_\alpha(X)$. It is worth to point out sometimes in the literature we have value at risk of X defined as a minus lower quantile of X, i.e. $VaR_\alpha(X) = -q^-_\alpha(X)$ (see papers and books of H. Föllmer and A. Schied, or R. Cont).

7.4 Properties of the value at risk

In this section we consider the properties of the value at risk. Namely we have that value at risk defined as $VaR_\alpha(X) = \inf\{m \in R | P\{m + X < 0\} \leq \alpha\}$ is such that
1. if $X \geq 0$ then $VaR_\alpha(X) \leq 0$,
2. if $X \geq Y$ then $VaR_\alpha(X) \leq VaR_\alpha(Y)$
3. $VaR_\alpha(\lambda X) = \lambda VaR_\alpha(X)$ for $\lambda \geq 0$
4. $VaR_\alpha(m + X) = VaR_\alpha(X) - m$

Consequently VaR_α is a monetary measure of risk. Unfortunately VaR_α in general is not a convex measure of risk as is shown in the example considered below, which was presented by Paul Embrechts.

7.5 Important example (Embrechts)

Assume that we have 100 i.i.d. loans of 100 units with 2% coupon, that are subject to credit risk with default probability 1%.
We consider two portfolios: X_1 consisting of 100 independent loans each of 100 units, and X_2 one loan of 100×100 units.
We have $X_1 = \sum_{i=1}^{100} Y_i$ i $X_2 = 100Y_1$, where $P\{Y_i = -100\} = 0.01$ and $P\{Y_i = 2\} = 0.99$.
One can show that

$$VaR_{5\%}(X_1) > VaR_{5\%}(X_2)$$

that is $VaR_{5\%}(\sum_{i=0}^{100} Y_i) > \sum_{i=1}^{100} VaR_{5\%}(Y_i) = -200$.
In fact, $q^+_{0.05}(Y_i) = 2$, and $q^+_{0.05}(\sum_{i=0}^{100} Y_i) = -1.64 \cdot 101.4887 + 98 < 2$ using normal approximation (by central limit theorem) for sequence Y_i.

7.6 Conditional value at risk

Absence of the property that diversification of portfolio minimizes risk is an important deficiency of value at risk. This was the reason that scientists started to look for other measures of risk. The measure of risk we consider now will be defined for integrable

continuous random variables X and is called conditional VaR (denoted by CVaR) or expected shortfall, average value at risk. It is defined as follows
$CVaR_\alpha(X) = E\{-X|X + VaR_\alpha(X) \le 0\}$
and it is the expected value if minus X, given X plus $VaR_\alpha(X)$ does not exceed 0.
We have the following immediate **properties** of $CVaR_\alpha$
1. $CVaR_\alpha(\lambda X) = \lambda CVaR_\alpha(X)$ for $\lambda \ge 0$
2. $CVaR_\alpha(m + X) = CVaR_\alpha(X) - m$
For the next properties we shall need the following

Lemma 2 *If $X \in L_1$, $x \in R$ s.t. $P\{X \le x\} > 0$ then for any A s.t. $P(A) \ge P\{X \le x\}$ we have*

$$E\{X|A\} \ge E\{X|X \le x\}$$

Proof.

$$\frac{1}{P(A)} \int_A u f_X(u) du \ge \frac{1}{P\{X \le x\}} \int_{X \le x} u f_X(u) dy$$

hint: approximate X by discrete r.v.

□

We now have further properties of CVaR:
3. $X \ge Y$ implies that $CVaR_\alpha(X) \le CVaR_\alpha(Y)$
In fact, since $P(X + VaR_\alpha(X) \le 0) = \alpha = P(Y + VaR_\alpha(Y) \le 0)$ we have

$$CVaR_\alpha(Y) = -E\{Y|Y + VaR_\alpha(Y) \le 0\} \ge -E\{X|Y + VaR_\alpha(Y) \le 0\} \ge$$
$$-E\{X|X + VaR_\alpha(X) \le 0\} = CVaR_\alpha(X)$$

4. $CVaR_\alpha(X + Y) \le CVaR_\alpha(X) + CVaR_\alpha(Y)$
In fact, we have

$$CVaR_\alpha(X + Y) = E\{-X|X + Y + VaR_\alpha(X + Y) \le 0\} +$$
$$E\{-Y|X + Y + VaR_\alpha(X + Y) \le 0\} \le E\{-X|X + VaR_\alpha(X) \le 0\} +$$
$$E\{-Y|Y + VaR_\alpha(Y) \le 0\}$$

Consequently CVaR is a coherent measure of risk. Furthermore we have the following important representation (see the book of H. Föllmer and A. Schied):
5. $CVaR_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha VaR_u(X)du = -\frac{1}{\alpha}\int_0^\alpha q_u^+(X)du.$
Furthermore $CVaR_\alpha$ from 5 we see that $CVaR_\alpha(X)$ is a continuous function of the parameter α and
6. $\lim_{\alpha \to 0} CVaR_\alpha(X) = -essinf X$
$\lim_{\alpha \to 1} CVaR_\alpha(X) = -EX.$

7.7 Measures of dispersion (deviation)

We have already considered two specific measures of risk: variance and semivariance. Now we continue the topic in a more general setting. Let
$\mathcal{X} = L^2$

Definition 4 *$D : \mathcal{X} \to [0, \infty]$ is a* **measure of dispersion** *iff*
$D(X + c) = D(X)$ for any $c \in R$
$D(0) = 0$, $D(\alpha X) = \alpha D(X)$ for $\alpha > 0$
$D(X + Y) \leq D(X) + D(Y)$
$D(X) \geq 0$, and $D(X) > 0$ for $X \neq const$

We have the following examples of measure of dispersion: standard deviation, negative semi standard deviation
$\sigma_-(X) = \sqrt{E(\max(EX - X, 0))^2}$
positive semi standard deviation
$\sigma_+(X) = \sqrt{E(max(X - EX, 0))^2}$.

8 Elliptic distributions

In this section we introduce a very important family of elliptic random variables introduced by **Douglas Kelker** in 1970 from the Department of Statistics and Applied Probability, University of Alberta.

8.1 Definition

Definition 5 *We say that random vector $X = (X_1, \ldots, X_d)^T$ has elliptic law, if there is a vector μ, a positive definite symmetric matrix Ω, a nonnegative function g_d such that, $\int_0^\infty x^{\frac{d}{2}-1} g_d(x)dx < \infty$, and a norming constant c_d such that the density f_X of the vector X is of the form*

$$f_X(x) = c_d|\Omega|^{-1/2} g_d(\frac{1}{2}(x - \mu)^T \Omega^{-1}(x - \mu)),$$

where $|\Omega|$ is the determinant of Ω.

One can show that

$$c_d = \frac{\Gamma(\frac{d}{2})}{(2\pi)^{d/2}} \left(\int_0^\infty x^{d/2-1} g_d(x)dx \right)^{-1}.$$

where $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$ and for positive integer d we have $\Gamma(d) = d!$, while $\Gamma(d+\frac{1}{2}) = \frac{1 \cdot 2 \cdot 3 \cdot (2d-1)}{2^d}\sqrt{\pi}$ i $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

8.2 Properties of elliptic distributions

Characteristic function of the elliptic vector X is given by

$$\phi_X(t) = E\left(e^{itX}\right) = e^{it^T\mu}\psi(\frac{1}{2}t^T\Omega t)$$

where a function $\psi(t)$ is called characteristic generator.
We shall use the following notation: $X \sim E_d(\mu, \Omega, \psi)$ or $X \sim E_d(\mu, \Omega, g_d)$.
If $\int_0^\infty g_1(x)dx < \infty$ there exists EX and $EX = \mu$. If furthermore

$$|\psi'(0)| < \infty$$

or equivalently $\int_0^\infty \sqrt{x}g_1(x)dx < \infty$ then

$$Cov(X) := E\left\{(X - EX)(X - EX)^T\right\} = -\psi'(0)\Omega.$$

The following property is crucial in applicability of elliptic random variables:

Lemma 3 *If $X \sim E_d(\mu, \Omega, g_d)$, A is $m \times d$ matrix ($m \le d$) and b - m dimensional vector then*

$$AX + b \sim E_m(A\mu + b, A\Omega A^T, g_m)$$

i.e. **linear combination of elliptic distributions with the same generator ψ is elliptic with generator ψ.**

Therefore

Corollary 1 **Marginal law of $X \sim E_d(\mu, \Sigma, g_d)$ is elliptic**

$$X_k \sim E_1(\mu_k, \omega_k^2, g_1)$$

where ω_k^2 is the k-th element of the diagonal of Ω, the density of X_k is of the form

$$f_{X_k}(x) = \frac{c_1}{\omega_k} g_1\left(\frac{1}{2}\left(\frac{x-\mu_k}{\omega_k}\right)^2\right).$$

and

Corollary 2 **Main property**
If $X \sim E_d(\mu, \Omega, g_d)$ then for $Y = \theta_1 X_1 + \theta_2 X_2 + \ldots + \theta_d X_d = \theta^T X$ we have

$$Y \sim E_1(\theta^T\mu, \theta^T\Omega\theta, g_1).$$

8.3 Examples

We list below more important examples of elliptic random variables:
1. **Multidimensional normal** $X \sim N_d(\mu, \Sigma)$
the density is of the form: (we identify $\Omega = \Sigma$)

$$f_X(x) = \frac{c_d}{\sqrt{|\Sigma|}} \exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right\}$$

with $c_d = (2\pi)^{\frac{-d}{2}}$; the characteristic function

$$\phi_X(t) = \exp\left\{it^T\mu - \frac{1}{2}t^T\Sigma t\right\}$$

so that $g(u) = e^{-u}$ and $\psi(t) = e^{-t}$ (since $\psi'(0) = -1$ $\Sigma = Cov(X)$).
2. **Multidimensional Student**
$X \sim t_d(\mu, \Omega; p)$ with $p > \frac{d}{2}$
the density is of the form:

$$f_X(x) = \frac{c_d}{\sqrt{|\Omega|}} \left[1 + \frac{(x-\mu)^T\Omega^{-1}(x-\mu)}{2k_p}\right]^{-p}$$

where $c_d = \frac{\Gamma(p)}{\Gamma(p-\frac{d}{2})}(2\pi k_p)^{\frac{-d}{2}}$, and k_p is a constant dependent on p,
we have here $g_d(u) = (1 + \frac{u}{k_p})^{-p}$;
In particular case when $p = \frac{d+\nu}{2}$ i $k_p = \frac{\nu}{2}$ we have *multidimensional t-Student with* ν *degrees of freedom* and then $\Omega = \frac{\nu}{\nu-2}\Sigma$.
In the case when $p = \frac{d+m}{2}$ for positive integer m i d and $k_p = \frac{m}{2}$ we have

$$f_X(x) = \frac{\Gamma(\frac{d+m}{2})}{(\pi m)^{\frac{d}{2}}\Gamma(\frac{m}{2})\sqrt{|\Sigma|}} \left[1 + \frac{(x-\mu)^T\Omega^{-1}(x-\mu)}{m}\right]^{-\frac{d+m}{2}}.$$

In general case for $k_p = \frac{2p-3}{2}$ with $p > \frac{3}{2}$ we have $Cov(X) = \Omega$.
Then in particular case for $p = \frac{d+m}{2}$

$$f_X(x) = \frac{\Gamma(\frac{d+m}{2})}{(\pi(d+m-3))^{\frac{d}{2}}\Gamma(\frac{m}{2})\sqrt{|\Omega|}} \left[1 + \frac{(x-\mu)^T\Omega^{-1}(x-\mu)}{d+m-3}\right]^{-\frac{d+m}{2}}$$

with $Cov(X) = \Omega$.
Moreover, when

$$f_X(x) = \frac{c_d}{\sqrt{|\Omega|}} \left[1 + \frac{(x-\mu)^T \Omega^{-1} (x-\mu)}{2k_p} \right]^{-p}$$

and $\frac{1}{2} < p \leq \frac{3}{2}$ there are no variance (we have so called heavy tails)
when $\frac{1}{2} < p \leq 1$ we have that EX does not exist. For $m = 1$ we have a **multidimensional Cauchy distribution**

$$f_X(x) = \frac{\Gamma(\frac{d+1}{2})\pi^{-\frac{d+1}{2}}}{\sqrt{|\Omega|}} \left[1 + (x-\mu)^T \Omega^{-1} (x-\mu) \right]^{-\frac{d+1}{2}}$$

3. **multidimensional logistic** h$g(u) = \frac{e^{-u}}{(1+e^{-u})^2}$,
4. **multidimensional exponential** when $g(u) = e^{-ru^s}$.

9 Portfolio analysis with elliptic rate of return

Now we shall continue the portfolio analysis using elliptic rate of return of asset prices.

9.1 Assumption and preliminary results

Our principal assumption in this section is:
Assumption: **random rate of return ζ is** $E_d(\mu, \Omega, \psi)$
The portfolio rate of return $R(\theta))$ (for a strategy $\theta = (\theta_1, \dots, \theta_d)$) is elliptic with the law

$$E_1(\theta^T \mu, \theta^T \Omega \theta, \psi).$$

Furthermore $Var(R(\theta)) = -\psi'(0)\omega^2$, where $\omega^2 = \theta^T \Omega \theta$, and

$$\frac{R(\theta) - \theta^T \mu}{\omega} \sim E_1(0, 1, \psi)$$

This procedure allows **standardization** of the elliptic random variables.

9.2 Risk measures for elliptic rate of returns

We assume that ζ is $E_d(\mu, \Omega, \psi)$ and consequently $R(\theta)$ is $E_1(\theta^T \mu, \theta^T \Omega \theta, \psi)$.
Consider first a measure of risk called **probability of the shortfall**

$$Risk(R(\theta)) = P\{R(\theta) \leq q\}.$$

Using standardization we obtain

$$Risk(R(\theta)) = F_Y(\frac{q - \theta^T\mu}{\omega}),$$

where F_Y is the distribution of $E_1(0,1,\psi)$.
Consequently this leads to the following lower bound for the expected portfolio rate of return

$$\theta^T\mu + \kappa_\alpha\omega \geq q,$$

where κ_α is an α quantile of $E_1(0,1,\psi)$.

9.3 Value at Risk - VaR_α

We shall now calculate value at risk for the portfolio with elliptic rates of return **Value at Risk (VaR_α)**:

$$VaR_\alpha(R(\theta)) = \inf\{x : P\{R(\theta) + x \leq 0\} \leq \alpha\}$$

We recall that it is the minimal value added to the portfolio rate of return which guarantees nonpositive rate of return with probability at most α. We have

$$VaR_\alpha(R(\theta)) = -\kappa_\alpha\omega - \theta^T\mu$$

where κ_α is α quantile of $E_1(0,1,\psi)$.

9.4 Conditional VaR_α or $CVaR_\alpha$

For *conditional* VaR_α ($CVaR_\alpha$), called also shortfall (expected shortfall) $CVaR_\alpha(R(\theta)) = E\{-R(\theta)|R(\theta) + VaR_\alpha \leq 0\}$, which is the expected value of $-R(\theta)$ given nonpositive $R(\theta) + VaR_\alpha$ we use the following representation form (see the book of Föllmer Schied)

$$CVaR_\alpha(R(\theta)) = \frac{1}{\alpha}\int_0^\alpha VaR_\beta d\beta = -\omega\frac{1}{\alpha}\int_0^\alpha \kappa_\beta d\beta - \theta^T\mu.$$

Remark 1 *We see that if we had no ω, where $\omega^2 = \theta^T\Omega\theta$, for a given α both $VaR_\alpha(R(\theta))$ and $CVaR_\alpha(R(\theta))$ would be a linear function of the investment strategy θ or in other words they would depend on the expected portfolio rate of return $\theta^T\mu$ only. Unfortunately this is not the case.*

9.5 Risk functions for elliptic rates of return

Since $CVaR_\alpha$ is a coherent measure of risk it is reasonable consider the following optimization problem: maximize

$$F_\lambda(E(R(\theta), CVaR_\alpha(R(\theta)) = E(R(\theta)) - \frac{1}{2}\lambda CVaR_\alpha(R(\theta)).$$

Notice that

$$F_\lambda(E(R(\theta), CVaR_\alpha(R(\theta))) = (1 + \frac{1}{2}\lambda)\theta^T\mu + \frac{1}{2}\lambda\sqrt{\theta^T\Omega\theta}\frac{1}{\alpha}\int_0^\alpha \kappa_\beta d\beta.$$

The second term is negative since for small α (usually below 0.05) the value of $\int_0^\alpha \kappa_\beta d\beta$ is negative. We are not able to solve the problem explicitly. One can find the maximum of

$$F_\lambda(E(R(\theta), CVaR_\alpha(R(\theta)))$$

using approximate methods
In fact, consider Lagrange multiplier's method to the function

$$F_\lambda(E(R(\theta), CVaR_\alpha(R(\theta))).$$

We form

$$G(\theta, \kappa) = (1 + \frac{1}{2}\lambda)\theta^T\mu + \frac{1}{2}\lambda\sqrt{\theta^T\Omega\theta}\frac{1}{\alpha}\int_0^\alpha \kappa_\beta d\beta + \kappa(\theta^T\mathcal{J} - 1).$$

Necessary condition for optimality is
$\frac{\partial G}{\partial \theta_i} = 0 \quad \frac{\partial G}{\partial \kappa} = 0$
Hence

$$(1 + \frac{1}{2}\lambda)\mu + \frac{1}{2}\lambda\frac{1}{\sqrt{\theta^T\Omega\theta}}2\Omega\theta z(\alpha) + \mathcal{J}\kappa = 0$$

with $\mathcal{J}^T\theta = 1$.
We are not able to solve θ from the first equation, which earlier together with the second equation gives us κ.
The difficulties come because of the existence of the term with square root. Since $\sqrt{x} \leq x$ for $x \geq 1$ one can optimize the *modified risk function* for elliptic rates of return

$$F_\lambda^m(\theta) = (1 + \frac{1}{2}\lambda)\theta^T\mu + \frac{1}{2}\lambda\theta^T\Omega\theta\frac{1}{\alpha}\int_0^\alpha \kappa_\beta d\beta,$$

which is diminished the term with $\theta^T\Omega\theta$.

We use to this new problem the Lagrange multiplier's method and have

$$G(\theta,\kappa) = (1+\frac{1}{2}\lambda)\theta^T\mu + \frac{1}{2}\lambda\theta^T\Omega\theta\frac{1}{\alpha}\int_0^\alpha \kappa_\beta d\beta + \kappa(\theta^T\mathcal{J} - 1)$$

Necessary condition for optimality is then of the form
$\frac{\partial G}{\partial \theta_i} = 0 \quad \frac{\partial G}{\partial \kappa} = 0$
from which we obtain

$$(1+\frac{1}{2}\lambda)\mu + \frac{1}{2}\lambda 2\Omega\theta z(\alpha) + \mathcal{J}\kappa = 0$$

with $\mathcal{J}^T\theta = 1$.
From the first equation

$$\theta = \frac{1}{\lambda z(\alpha)}\Omega^{-1}\left(-(1+\frac{1}{2}\lambda)\mu - \mathcal{J}\kappa\right),$$

substituting this to the second equation we obtain

$$\mathcal{J}^T\frac{1}{\lambda z(\alpha)}\Omega^{-1}\left(-(1+\frac{1}{2}\lambda)\mu - \mathcal{J}\kappa\right) = 1$$

or

$$\frac{1}{\lambda z(\alpha)}\left(-(1+\frac{1}{2}\lambda)\mathcal{J}^T\Omega^{-1}\mu - \mathcal{J}^T\Omega^{-1}\mathcal{J}\kappa\right) = 1$$

we obtain

$$\kappa = \left(\left(-(1+\frac{1}{2}\lambda)\mathcal{J}^T\Omega^{-1}\mu\right) - \lambda z(\alpha)\right)\frac{1}{\mathcal{J}^T\Omega^{-1}\mathcal{J}}.$$

Therefore

$$\kappa = \frac{1}{\mathcal{J}^T\Omega^{-1}\mathcal{J}}\left(-(1+\frac{1}{2}\lambda)\mathcal{J}^T\Omega^{-1}\mu\right) - \frac{\lambda z(\alpha)}{\mathcal{J}^T\Omega^{-1}\mathcal{J}}$$

and finally the optimal strategy is of the form

$$\theta = \frac{-(1+\frac{1}{2}\lambda)}{\lambda z(\alpha)}\Omega^{-1}\mu + \frac{(1+\frac{1}{2}\lambda)\mu^T\Omega^{-1}\mathcal{J}}{\lambda z(\alpha)\mathcal{J}^T\Omega^{-1}\mathcal{J}}\Omega^{-1}\mathcal{J} + \frac{1}{\mathcal{J}^T\Omega^{-1}\mathcal{J}}\Omega^{-1}\mathcal{J}$$

with $z(\alpha) = \frac{1}{\alpha}\int_0^\alpha \kappa_\beta d\beta$.

9.6 Further alternatives

One can consider also a function

$$Risk(R(\theta)) = (CVaR_\alpha(R(\theta)) + ER(\theta))^2.$$

which corresponds to the square of the CVaR (an analogy to the variance considered as a square of the standard deviation).
Then

$$\tilde{F}_\lambda = \theta^T \mu - \frac{1}{2}\lambda\theta^T\Omega\theta(z(\alpha))^2$$

the only difference is that $z(\alpha) = \frac{1}{\alpha}\int_0^\alpha \kappa_\beta d\beta$ in the aim function has been replaced by $(z(\alpha))^2$. The optimal strategy is of the form

$$\tilde{\theta} = \frac{-(1+\frac{1}{2}\lambda)}{\lambda(z(\alpha))^2}\Omega^{-1}\mu + \frac{(1+\frac{1}{2}\lambda)}{\lambda(z(\alpha))^2}\frac{\mu^T\Omega^{-1}\mathcal{J}}{\mathcal{J}^T\Omega^{-1}\mathcal{J}}\Omega^{-1}\mathcal{J} + \frac{1}{\mathcal{J}^T\Omega^{-1}\mathcal{J}}\Omega^{-1}\mathcal{J}.$$

We complete this section with references.

9.7 Further references

De Giorgi, E., – A Note on Portfolio Selection under Various Risk Measures, Working paper, University of Zurich 2002.
Fang, K.T., Kotz, S., Ng, K.W., Symmetric Multivariate and Related Distributions, London, Chapman & Hall 1987.
Föllmer, H., Schied, A. – Stochastic Finance. An Introduction in Discrete Time, W. de Gruyter, 2002.
Kelker, D., – Distribution Theory of Spherical Distributions and Location-Scale Parameter Generalization, Sankhya 32 (1970), 419-430
Landsman, Z., Valdez, E. A. – Tail Conditional Expectations for Elliptical Distributions, North American Actuarial Journal 7 (2003), 55-71,

10 Risk of the insurance company

This section is devoted to aspect of risk in insurance companies. Probability of ruin (negative value of surplus) was considered as a measure of risk. We shall continue to study this measure considering also two typical problems of actuary sciences: optimal dividend and reinsurance problems.

10.1 Definition of risk

We consider first ruin probability for Cramer Lundberg model

$$\psi(x) = P\left\{X_t < 0 \quad for \quad some \quad t \geq 0\right\},$$

with $X_t = x + ct - \sum_{i=1}^{N_t} Y_i$.
One can show that when $c \leq \lambda EY_1$ we have $\psi(x) = 1$. Therefore our typical assumption would be so called *"safety loading":* which means that $c > \lambda EY_1$.
We calculate ψ and obtain

$$\psi(x) = \int_0^\infty \lambda e^{-\lambda t}\left(\int_0^{x+ct} \psi(x+ct-y)G(dy) + (1 - G(x+ct))\right) dt.$$

Letting $z = x + ct$ we have

$$\psi(x) = \frac{1}{c} e^{\lambda \frac{x}{c}} \int_x^\infty \lambda e^{-\lambda \frac{z}{c}} \left(\int_0^z \psi(z-y)G(dy) + (1 - G(z))\right) dz$$

and

$$\psi'(x) = (\frac{\lambda}{c})^2 e^{\lambda \frac{x}{c}} \int_x^\infty e^{-\lambda \frac{z}{c}} \left(\int_0^z \psi(z-y)G(dy) + (1 - G(z))\right) dz$$
$$+\frac{\lambda}{c} e^{\lambda \frac{x}{c}} (-1) e^{-\lambda \frac{x}{c}} \left(\int_0^x \psi(x-y)G(dy) + (1 - G(x))\right).$$

Consequently

$$c\psi'(x) = \lambda\psi(x) - \lambda\left(\int_0^x \psi(x-y)G(dy) + (1 - G(x))\right)$$

and the inifnitesimal generator of the function ψ defined as

$$A\psi(x) = \lim_{t\to 0} E\left\{\frac{\psi(X_{t\wedge\tau}) - \psi(x)}{t}\right\}$$

exists and is of the form

$$A\psi(x) = \lambda\left(E\left\{\psi(x - Y_1)\right\} - \psi(x)\right) + c\psi'(x) = 0.$$

10.2 Optimal dividend problem

We start first with the main problem of insurance company which is optimal distribution of dividends. The surplus is of the form

$$X_{n+1} = X_n - U_n + Y_{n+1},$$

where X_n is the surplus before dividend at time n,
U_n is the dividend at time n, which should not exceed the value of surplus i.e. $0 \leq U_n \leq X_n$, and
Y_n is a value of premia minus payout of claims at time n, which is a sequence of (i.i.d.) random variables with distribution G.
Our strategy is a sequence $U = (U_0, U_1, \ldots, U_n, \ldots)$ of dividends paid at times $0, 1, \ldots, n$.
Denote by $\tau = \inf\{n : X_n < 0\}$ the ruin moment.
Our optimization problem is to maximize

$$V^U(x) = E_x \left\{ \sum_{n=0}^{\tau-1} e^{-\delta n} U_n \right\}$$

the discounted (δ is a discount rate) dividend till the ruin problem. Optimal solution to such problem is of the form threshold strategies (with one or more thresholds depending on the form of distribution G). In the case of one threshold we pay dividend to reach with surplus a certain threshold and below we pay nothing.

10.3 Optimal reinsurance problem

For an insurance company one possibility to minimize risk is to insure part of claims in another company. Then the surplus is of the form

$$X_{n+1} = X_n + c(b_n) - s(Y_{n+1}, b_n)$$

with $Y_n \geq 0$ standing for claims (i.i.d. with distribution G), $c(b)$ is premium left if reinsurance b is chosen, and $0 \leq s(y, b) \leq y$ is the expense of the first insurer (cedent) (self-insurance function), under claim y and reinsurance b. The*reinsurance function is fixed* however we have a possibility to choose the reinsurance level b The reinsurance function has the properties:
$s(y, b) \geq s(y, b')$ for $y \geq 0$ implies that $c(b) \geq c(b')$; and also $c(b_{max}) < 0$ $(s(y, b_{max}) = y)$
We also assume that $y \mapsto c(y, b)$ is increasing and continuous, $b \mapsto c(b) \leq c$ continuous, where c corresponds to the absence of reinsurance and $P\{s(Y, b) > c(b)\} > 0$ i $E\{c(b) - s(Y, b)\} > 0$ for certain b.
Our purpose is to choose the strategy $B = (b_0, b_1, \ldots, b_n, \ldots)$ so as to minimize functional

$$V^B(x) = -P_x\{\tau < \infty\}.$$

10.4 Forms of self-insurance function s

We can consider various form of self-insurance function s:
- proportional reinsurance: $s(y,b) = yb$, where $b \in (0,1)$ (retention level); reinsurer pays $(1-b)y$, insurer by
- excess of loss reinsurance: $s(y,b) = \min\{y,b\}$; reinsurer pays $(y-b)^+$, insurer $\min\{y,b\}$
- first risk deductible: $s(y,b) = (y-b)^+$, reinsurer pays $\min\{y,b\}$,
- proportional reinsurance in a layer:
$s(y) = \min\{y,a\} + (y-a-\gamma)^+ + b\min\{(y-a)^+,\gamma\}$, dla $a,\gamma > 0$, $b \in (0,1)$
We complete this section with references concerning the insurance problems we formulated above.

10.5 References

A. Martin - Löf, "Lectures on the use of control theory in insurance", Scand. Acturial J. 1994, 1-25,
C. Hipp, "Stochastic control with Application in Insurance", Stochastic Methods in Finance, LNM 1856, Springer 2004,
H. Schmidli, "Stochastic Control in Insurance", Springer 2008,
T. Rolski, H. Schmidli, V. Schmidt, J. Teugels, "Stochastic Processes for Insurance and Finance", Wiley 1998,
P. Brockett, X. Xia, "Operations research in insurance: a review", Trans. Act. Soc. XLVII (1995), 7-80,

11 Pricing of financial derivatives

11.1 Options (contingent claims)

We consider first two financial instruments used frequently to get protection against risk of the changes of asset prices. They depend on the fact whether we intend to buy an asset at time T called maturity and have a certain guarantee concerning the price we shall pay for it at maturity or we want to sell assets at time T and again want to have a guarantee concerning the price for which we can sell them. We have therefore two kinds of options:
European call option:
The buyer of this contract has the option to buy (exercise the option), at time T one share of the stock at the specified (at time $t=0$) price K called exercise (striking) price. Possible profit of the buyer at time T is of the form

$$(S_1(T) - K)^+$$

European put option
The buyer of this contract has the option to sell (exercise the option), at time T one share of the stock at the specified (at time $t = 0$) price K called exercise (striking) price. Possible profit of the buyer at time T is of the form

$$(K - S_1(T))^+$$

In this option maturity T is fixed. An alternative is so called an American option win which the buyer is allowed to choose the time of exercise the option. We have
American call option - situation is similar as in the case of European call option with the only difference that the buyer can choose the exercise time τ. The profit then is of the form

$$(S_1(\tau) - K)^+$$

American put option - similar situation to the European put option again with possibility to choose the exercise time τ

$$(K - S_1(\tau))^+ .$$

The exercise time τ is in fact a stopping time: $\{\tau \leq t\} \in F_t$ for $t \geq 0$ i.e. a random time which is chosen basing on available information.
We have therefore the following time horizon
0 T (τ)
In options the contract terminates at time T or τ and then we get a possible payoff. The question (and our problem) is: how much the option should cost at time 0? It appears that there is a certain ideology behind pricing of financial instruments which is explained in the next section.

11.2 Wealth process

Given an initial capital v we consider the wealth process V defined as follows
$V(0) = v$, $V(t)$
Our **investment strategy** consists of the number N_i of i-th assets held in our portfolio, or portions π_i of our capital V invested in the i-th assets
$(N_0(t), N_1(t), \ldots, N_d(t))$ or $(\pi_0(t), \pi_1(t), \ldots, \pi_d(t))$
Then we have

$$V(t) = \sum_{j=0}^{d} N_j(t) S_j(t) = V(t-1) + \sum_{j=0}^{d} N_j(t-1)(S_j(t) - S_j(t-1)) =$$

$$V(t-1)+N(t-1)\cdot\Delta S(t)=V(t-1)\sum_{j=0}^{d}\pi_j(t-1)\zeta_j(z(t),\xi(t))$$

with $\Delta S_j(t)=S_j(t)-S_j(t-1)$.
Such property usually is called *selffinancing* property of the portfolio or wealth process $V^{v,\pi}(t)$
With the use of selffinancing portfolios we define option prices. For a given contingent claim H we define the seler price $p_s(H)$ as follows:
seller:

$$p_s(H)=\inf\left\{v:\exists_\pi,V^{v,\pi}(T)\geq H\right\},$$

which means that it is a minimal initial capital that allows us to hedge the claim H at time T (to get $V^{v,\pi}(T)\geq H$ with probability 1).
We also have the buyer price $p_b(H)$ of the contingent claim H
buyer:

$$p_b(H)=\sup\left\{v:\exists_\pi,V^{-v,\pi}(T)+H\geq 0\right\}$$

which means that it is a maximal value of the initial capital v such that with a debt v we are able to construct a portfolio which together with income H gives us at time T nonnegative position with probability 1 (we shall have no debts at time T)
We clearly have that

$$p_b(H)\leq p_s(H).$$

Definition 6 Absence of arbitrage (AA) *We say that we have an absence of arbitrage (AA) if there are no portfolio π, such that $V^{\pi,0}(t)\geq 0$ for $t\in[0,T]$ and $P\left\{V^{0,\pi}(T)>0\right\}>0$*

The interval $[p_b(H),p_s(H)]$ is called **(AA) interval**.

11.3 Martingale measures

On a given (Ω,F,P) probability space we define so called martingale measure. Denote by $B(t)$ the value at time t of the banking account consisting of an investment at time 0 equal to $B(0)=1$. We call it discount factor.

Definition 7 *A probability measure Q is called a martingale measure, if $Q\sim P$ (equivalent) and $(\frac{S_i(t)}{B_t})$ is a Q martingale $i=1,2,\ldots,d$ that is it is Q integrable ($E^Q\left\{\frac{S_i(t+1)}{B_{t+1}}\right\}<\infty$) and*

$$E^Q\left\{\frac{S_i(t+1)}{B_{t+1}}|F_t\right\}=\frac{S_i(t)}{B_t}$$

The notion of martingale measures has been introduced in the papers M. Harrison, D. Kreps (1979), M. Harrison - S. Pliska (1981), in which the absence of arbitrage, which is a fundamental economical requirement for any economical consideration was explained in terms of martingales measures (methodology coming from stochastic processes theory !). These results were first formulated for finite markets (the case when Ω was finite) and their generalization for a long time (till 1990) was an open problem.

11.4 Fundamental theorem of mathematical finance - discrete time

The above mentioned relationship between absence of arbitrage and martingale measures in a general discrete setting can be formulated as follows

Theorem 4 *(Dalang Morton Willinger 1990) (AA) $\equiv$ existence of a martingale measure Q ($\mathcal{Q} \neq \emptyset$)*

We denote by $\mathcal{Q}$ - family of all martingale measures
In what folows we shall need a notion of generalized martingale and generalized supermatingale.

Definition 8 *We say that (Z_t) is a generalized (F_t) martingale with respect to measure P whenever*

$$E^P\left[Z_{t+1}^+|F_t\right] < \infty, \quad E^P\left[Z_{t+1}^-|F_t\right] < \infty$$

and

$$E^P\left[Z_{t+1}|F_t\right] := E^P\left[Z_{t+1}^+|F_t\right] - E^P\left[Z_{t+1}^-|F_t\right] = Z_t$$

If instead of the last condition we have $E^P\left[Z_{t+1}^+|F_t\right] - E^P\left[Z_{t+1}^-|F_t\right] \leq Z_t$, then (Z_t) is a generalized (F_t) supermartingale with respect to measure P.

Proof. We prove only sufficiency of the existence of martingale measure.
We start with the following

Lemma 4 *(Jacod Shiryaev 1998) Generalized martingale $(X_t)_{t=0,1,\dots,T}$, such that $E^P X_T^+ < \infty$ or $E^P X_T^- < \infty$ is a martingale. Whenever sequence (ξ_n) is F_n-adapted, and (X_n) is a generalized martingale then martingale transformation*

$$\sum_{n=1}^{t} \xi_{n-1}(X_n - X_{n-1})$$

is also a generalized martingale.

Under investment strategy $\pi = (N_0(t), N_1(t), \ldots, N_d(t))$ we have

$$V_t^{0,\pi} = \sum_{n=0}^{t-1} N_n \cdot \Delta S_n$$

where $\Delta S_n = S_{n+1} - S_n$, is a generalized martingale with respect to measure Q. Therefore if $V_T^{0,\pi} \geq 0$ then by Lemma 4, $(V_t^{0,\pi})$ is a martingale with respect to the measure Q, that is

$$E^Q[V_T^{0,\pi}] = V_0 = 0,$$

which means that $V_T^{0,\pi} = 0$, so that there are no arbitrage.

□

Next theorem is fundamental in pricing of European contingent claims (claims with fixed maturity T). We have

Theorem 5 *Under (AA) if $E^Q H^- < \infty$ then $p_s(H) = \sup_{Q \in \mathcal{Q}} E^Q H$ and if $EQH^+ < \infty$ then $p_b(H) = \inf_{Q \in \mathcal{Q}} E^Q H$.*

Proof. Steps:
1. The family $\left\{E^Q(H|F_t), Q \in \mathcal{Q}\right\}$ is directed upward i.e. for $Q_1, Q_2 \in \mathcal{Q}$ there is $Q_3 \in \mathcal{Q}$, such that

$$E^{Q_3}[H|F_t] \geq \max\left\{E^{Q_1}[H|F_t], E^{Q_2}[H|F_t]\right\}.$$

Hence there is a process $Y_t := esssup_{Q \in \mathcal{Q}} E^Q[H|F_t]$ is attained by an increasing sequence of random variables.
2. If $Q_1, Q_2 \in \mathcal{Q}$, then there is $Q_3 \in \mathcal{Q}$, for which
$E^{Q_3}[H|F_t] = E^{Q_1}[E^{Q_2}[H|F_{t+1}]|F_t]$
3. We prove here the following

Lemma 5 *The process $Y_t := esssup_{Q \in \mathcal{Q}} E^Q[H|F_t]$ is a generalized Q-supermartingale for each $Q \in \mathcal{Q}$, i.e. $E^Q[Y_{t+1}|F_t] \leq Y_t$, dla $t = 0, 1, \ldots, T-1$.*

Proof. From the step 1 we know that there is a sequence of measures $Q_n \in \mathcal{Q}$, such that $E^{Q_n}[H|F_{t+1}] \uparrow Y_{t+1}$. By step 2 for a given measure $Q \in \mathcal{Q}$ there is a sequence $Q_n' \in \mathcal{Q}$ such that

$$E^{Q_n}[H|F_t] \geq E^Q[E^{Q_n}[H|F_{t+1}]|F_t].$$

Hence

$$Y_n \geq E^Q[E^{Q_n}[H|F_{t+1}]|F_t]$$

and by Lebesgue monotonical convergence theorem we obtain $Y_n \geq E^Q[Y_{t+1}|F_t]$, which completes the proof.

□

To complete the proof of theorem we need another important result so called optional decomposition, which we formulate in the next section.

11.5 Optional decomposition theorem and its consequences

4. We continue the proof recalling first the following result:

Theorem 6 *(Föllmer i Kabanov (1998)) If $\mathcal{Q} \neq \emptyset$ and Y_t is a generalized Q-supermartingale for each measure martingale measure Q, then there is a strategy $\pi = (N_t)_{t\geq 0}$ and increasing F_t adapted process d_t with $d_0 = 0$ such that*
$Y_0 + \sum_{t=0}^{T-1} N_t \cdot \Delta S_t - d_T = Y_T$

Proof of pricing theorem cont. Since $Y_T = H$, by the definition of $p_s(H)$ and the above theorem we have $Y_0 = \sup_{Q\in\mathcal{Q}} E^Q H \geq p_s(H)$. It remains to show the inverse inequality. Since

$$x + \sum_{t=0}^{T-1} N_t \cdot \Delta S_t \geq H$$

from Jacod Shiryaev's Lemma 4 it follows that $(\sum_{t=0}^{n-1} N_t \cdot \Delta S_t)$ is a martingale with respect to any $Q \in \mathcal{Q}$. Hence we obtain $x \geq E^Q H$ for any $Q \in \mathcal{Q}$, which completes the proof.

□

11.6 Remarks on the seller price

In the case when there are not transaction costs we have the following simple relation between the buyer and seller prices

$$p_b(H) := -p_s(-H)$$

or equivalently

$$p_b(H) = \sup\left\{x : \exists_\pi V_T^{-x,\pi} + H \geq 0, Pa.e. \right\},$$

which is the way we defined buyer price.
Since $V^{-x,\pi} = -V^{x,-\pi}$ we have

$$p_b(H) = \sup\left\{x : \exists_\pi V_T^{x,\pi} \leq H, Pa.e.\right\}.$$

Furthermore

$$p_s(H) - p_b(H) = \inf\left\{x : \exists_\pi V_T^{x,\pi} \geq H\right\} + \inf\left\{\bar{x} : \exists_{\bar{\pi}} V_T^{\bar{x},\bar{\pi}} \geq -H\right\}$$
$$\geq \inf\left\{x + \bar{x} : \exists_\pi \exists_{\bar{\pi}} V_T^{x,\pi} \geq H, X_T^{\bar{x},\bar{\pi}} \geq -H\right\} \geq \inf\left\{x : \exists_\pi V_T^{x,\pi} \geq 0\right\} \geq 0$$

which means that $p_s(H) \geq p_b(H)$. Such result is intuitively clear and follows immediately from Theorem 5.

11.7 Complete markets

Financial market is complete if for any F_T measurable bounded contingent claim H there is a strategy π and initial capital v such that $V_T^{v,\pi} = H$, P a.e.. Such strategy is called *replicating* strategy. We have

Theorem 7 *(Jacod Shiryaev 1998) Market is complete iff $\mathcal{Q}$ is a singleton.*

There are not too many markets which are complete. In general markets are incomplete. However we have the following example
Example of a complete market: Cox Ross Rubinstein model (1979) We consider the case $d = 1$ with the random rate of return $\zeta_t = \frac{S_t - S_{t-1}}{S_{t-1}}$ and (ζ_t) being a sequence of i.i.d. random variables. Furthermore we assume that
$P\{\zeta_t = a\} > 0$ and $P\{\zeta_t = b\} = 1 - P\{\zeta_t = a\} > 0$ where a i b, $b > 0 > a > -1$, so that there are only two possible values of $\zeta(t)$: the random rate of return is either below zero (is equal to a and the price of the asset diminishes) or is above zero (is equal to b and the price of the asset is increasing).
For such evolution of asset prices we can characterize martingale measures Q. By the very definition we should have

$$E^Q\{\zeta_{t+1}|F_t\} = aQ\{\zeta_{t+1} = a|F_t\} + b(1 - Q\{\zeta_{t+1} = a|F_t\}) = 0,$$

Therefore $Q\{\zeta_{t+1} = a|F_t\} = \frac{b}{b-a}$, and $Q\{\zeta_{t+1} = b|F_t\} = \frac{-a}{b-a}$. It can be further shown that under Q, (ζ_t) are also of i.i.d. random variables.

11.8 Pricing under Cox Ross Rubinstein model

We consider the case when the contingent claim is of the European form $H = h(S_T)$.

Theorem 8 *We have*

$$p_s(H) = p_b(H) = \\ \sum_{i=0}^{T} \frac{T!}{(T-i)!i!} \left(\frac{b}{b-a}\right)^i \left(\frac{-a}{b-a}\right)^{T-i} h(S_0(1+a)^{T-i}(1+b)^i),$$

Proof. It Follows form Bernoulli scheme under measure Q or from direct replication argument.

□

11.9 Continuous time models

In continuous time case the modeling and results are more difficult. Our standard assumption is that the asset prices $(S(t))$ are in general semimartingales (roughly speaking they are the sums of local martingales and processes with locally finite variation - for simplicity we shall have in mind everywhere here $(S(t))$ of the form (3) defined in section 3.4). Then the wealth value of our investments $(\pi(t))$, where we assume that we are allowed to change the portfolio instantaneously is of the form

$$V^{v,\pi}(t) = v + \int_0^t \pi(u) V^{v,\pi}(u) dS(u)$$

with an integral with respect to the semimartingale $(S(t))$ considered as a limit of integral with simple (piecewise constant) integrands. For the purpose of this lecture we skip here mathematical precision which would lead us far away to the stochastic integration theory. In general continuous time setting the only admissible strategies are so called tame strategies i.e. such (predictable) strategies π for which there is a deterministic constant $K < 0$ such that for each t we have that $V^{v,\pi}(t) \geq K$. P a.e.. The use of other strategies may lead to arbitrage.
The absence of arbitrage is defined now by analogy to discrete time case

Definition 9 Absence of arbitrage (AA) *We say that we have an absence of arbitrage (AA) if there are no admissible strategy π, such that $V^{\pi,0}(t) \geq 0$ for $t \in [0, T]$ and $P\left\{V^{0,\pi}(T) > 0\right\} > 0$.*

11.10 Two important examples

We consider now two examples which show that without restricting to tame strategy even using simple strategies we are able to create an arbitrage.
Example 1.
Let $(B_t)_{t\geq 0}$ be a Brownian motion starting from 1, i.e. $B(0) = 1$. Denote by $\tau = \inf\{t > 0 : B_t = 0\}$ first entry time to 0. We define the asset price S_t as follows

$$S_t = B_{\tan(\frac{t\pi}{2})\wedge\tau}$$

for $t \leq 1$, and $S_1 = B_\tau$ for $t = 1$.
Consider the fixed strategy $N(t) = -1$
Then we have that

$$(N \cdot S)_1 = \int_0^1 N(t) dS_t = 1$$

P a.s. and $(N \cdot S)_0 = 0$, so that we have an arbitrage for a simple strategy based on shortselling of one asset.
This strategy is not tame since before reaching 0 Brownian motion can take arbitrarily large values.
Example 2.
Assume now that we are given the discrete time process $(x(n))$ defined as follows $x(0) = 1$ and $P\{x(1) = 1\} = \frac{1}{2}$, $P\{x(1) = -1\} = \frac{1}{2}$, ...,
$P\{x(n) = 2^{n-1}\} = \frac{1}{2}$, $P\{x(n) = -2^{n-1}\} = \frac{1}{2}$.
We construct the continuous time process $(z(t))$
$z(t) = \sum_{i=0}^{kn} x(i)$ for $kn \le t < (k+1)n$
and let
$S_t = z(\tan(\frac{t\pi}{2}) \wedge \tau)$ for $t \le 1$, and $S_1 = z(\tau)$ for $t = 1$, with $\tau = \inf\{t > 0 : z_t = 0\}$
Then for $N(t) = -1$, $(N \cdot S)_1 = \int_0^1 N(t) dS_t = 1$ a.s. and $(N \cdot S)_0 = 0$
so that again we have an arbitrage.

11.11 No free lunch with vanishing risk (NFLVR)

In the case of continuous time we don't have a version of Theorem 4. We have to introduce new notion of so called no free lunch vanishing risk condition defined as follows (NFLVR): there are no sequence of strategies $(\pi^k(t))$ such that:
$\exists_{\alpha_k}$, $P\left\{V^{\pi^k,0}(t) \ge \alpha_k, t \in [0,T]\right\} = 1$
$\forall_k, \exists_{\delta_1,\delta_2 > 0}$, $P\left\{V^{\pi^k,0}(T) > \delta_1\right\} > \delta_2$, and
$V^{\pi^k,0}(T) \ge -\frac{1}{k}$
We have then the following theorem

Theorem 9 *(F. Delbaen-W. Schachermayer 1994)*

$$(AA) \Longleftarrow (NFLVR) \Longleftrightarrow \mathcal{Q} \ne \emptyset$$

where $\mathcal{Q}$ the family of equivalent measures, such that $(\frac{S(t)}{B_t})$ is a local martingale i.e. there is a sequence of stopping times (τ_n) such that $\forall_n$ $\left(\frac{S(t\wedge\tau_n)}{B_{t\wedge\tau_n}}\right)$ is a Q - martingale.

As we see in continuous time case we don't have equivalence we have only implication.

11.12 Black and Scholes formula

The most known (and applicable) result in continuous time mathematical finance is the so called Black Scholes formula established in the paper Black, Fischer; Myron Scholes (1973). "The Pricing of Options and Corporate Liabilities". Journal of Political Economy

81 (3) 637-654, by Fischer Black, 11.01.1938 30.08.1995, PhD Applied Math. Harward 1964, and Myron Scholes, 1.07.1941, PhD Economics, University of Chicago 1969. M. Scholes received Noble price in 1997 for this formula (Black died two years earlier). The formula concern pricing of European call option with maturity T and striking price K on the market consisting of the banking account with constant interest rates $r_t = r$, which means that $B(t) = e^{rt}$ and one asset with price given by the formula

$$S(t) = S(0)e^{\int_0^t ads + \int_0^t \sigma dW(s)} \tag{22}$$

Consequently we want to price $H = (S(T) - K)^+$. This market is another (continuous time) example of complete market which means that the seller and buyer prices are the same. We have

$$p_s(H) = p_b(H) = S(0)N(d_1(S(0), T)) - Ke^{rT}N(d_2(S(0), T))$$

where $d_1(s,t) = \frac{\ln \frac{s}{K} + (r + 0.5\sigma^2)t}{\sigma\sqrt{t}}$, $d_2(s,t) = d_1(s,t) - \sigma\sqrt{t}$. Notice that we have an explicit formula for the price, which is rather simple (can be easily computed) and does not depend on a. The explicit form of the price was the main factor that stimulated frequent use of the formula although the market considered here is too simple to model real market.

11.13 References

R.N. Dana, M. Jeanblanc, Financial Markets in Continuous Time, Springer 2003
H. Föllmer, A. Schied, Stochastic Finance. An Introduction in Discrete Time, W. de Gruyter, 2002,
J. Jakubowski, A. Palczewski, M. Rutkowski, Ł. Stettner, Matematyka Finansowa, Instrumenty pochodne, WNT 2003 (in Polish)
I. Karatzas, Lecture on the Mathematics of Finance, AMS 1997,
A. Shiryaev, Essentiels of Stochastic Finance, World Scientific 1999,
S. Shreve, Stochastic Calculus for Finance I. The Binomial Asset Pricing Model. Springer 2003,
S. Shreve, Stochastic Calculus for Finace II. Continuous-Time Models. Springer 2003,
A. Weron, R. Weron, Inzynieria Finansowa, WNT 1998,

12 Coherent risk measures

In the last part of the lectures we come back to coherent measures of risk. We present here some most recent result stimulated by the necessity to measure risk in the real world problems, partly stimulated by the works of Basel II and III arrangements.

12.1 Representation results

(Ω, F, P) a given probability space we consider here the class of financial positions $\mathcal{X} = L^\infty$. We have the following fundamental result

Theorem 10 *(Artzner Ph., Delbaen F., Eber J.M, Heath D. (1999)) Let ρ be a coherent risk measure. The following conditions are equivalent:*
1. there is a closed convex set of probability measures $\mathcal{Q}$ such that any $q \in \mathcal{Q}$ is absolutely continuous with respect to P and for $X \in L^\infty$
$\rho(X) = \sup \left\{ E^Q [-X] ; Q \in \mathcal{Q} \right\}$
2. ρ satisfies the Fatou property, i.e. if $(X_n) \subset L^\infty$ be a sequence of uniformly bounded random variables converging in probability to X, then
$\rho(X) \leq \lim inf_{n\to\infty} \rho(X_n)$.
3. If (X_n) is a uniformly bounded sequence that decreases to X then $\rho(X_n)$ converges to $\rho(X)$.

□

Let
$$\rho_m(X) = -\int_0^1 q_u^+(X) m(du)$$
where m is a probability measure on $(0, 1)$. Such a measure is monotone, cash additive and positive homogeneous.
We have the following examples of such measures:
1. $VaR_\alpha(X) = -q_\alpha^+(X)$, for $m = \delta_\alpha$,
2. expected shortfall (conditional VaR): $ES_\alpha(X) = \frac{1}{\alpha} \int_0^\alpha VaR_u(X) du$, for m uniform distribution over $(0, \alpha)$,
3. spectral risk measures (weighted VaRs): $\rho_\phi(X) = \int_0^1 VaR_u(X)\phi(u)du$, with $m(du) = \phi(u)du$, where $\phi : [0, 1] \to [0, \infty)$ is a density on $[0, 1]$ and $u \mapsto \phi(u)$ is decreasing,
Furthermore we have

Theorem 11 *ρ_m is subadditive iff it is a spectral risk measure.*

which means that the set of coherent measures is equivalent to the set of spectral risk measures

12.2 Law invariant measures of risk

In majority of applications the risk measure depends on the distribution (law) of random financial position only.

Definition 10 *We say that a map $\rho : L^\infty \to R$ is law invariant if $\rho(X) = \rho(Y)$, whenever X, Y have the same probability law.*

In this section we shall use the following assumption
Assumption:(A) (Ω, F, P) a fixed probability space with measure P non-atomic; $L^\infty(\Omega, F, P)$
We have

Theorem 12 *(Kusuoka 2001) Under (A) the following conditions are equivalent:*
1. there is a (compact, convex) set $\mathcal{M}_0$ of probability measures on $[0,1]$ such that $\rho(X) = \sup\left\{\int_0^1 ES_\alpha m(d\alpha) : m \in \mathcal{M}_0\right\}$, for $X \in L^\infty$
2. ρ is a law invariant coherent risk measure with the Fatou property

12.3 Comonotone measures of risk

We start with the following definition

Definition 11 *We say that a pair X and Y of random variables is comonotone if $(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geq 0$, $P(d\omega) \times P(d\omega')$ a.s.*
We say that a map $\rho : L^\infty \to R$ is comonotone if $\rho(X+Y) = \rho(X) + \rho(Y)$ for any comonotone pair $X, Y \in L^\infty$.

As an example of a comonotone measure of risk may serve expected shortfall ES_α.
In fact we have the following identification result for the class of comonotone measures of risk

Theorem 13 *(Kusuoka 2001) Under (A) the following conditions are equivalent*
1. there is a probability measure m on $[0,1]$ such that for $X \in L^\infty$
$\rho(X) = \int_0^1 ES_\alpha(X) m(d\alpha)$
2. ρ is a law invariant and comonotone coherent risk measure with the Fatou property.

12.4 Further properties of coherent measures of risk

We present here further results on measures of risk. We have

Theorem 14 *(Kusuoka 2001) Under (A) if ρ is law invariant coherent measure of risk such that*
$\rho(X) \geq VaR_\alpha(X)$, *for $X \in L^\infty$*
then we have $\rho(X) \geq ES_\alpha(X)$ for $X \in L^\infty$.

12.5 Historical risk estimators

Since law invariant measures of risk depend on the distribution of random variable only instead of $\rho(X)$ we may write $\rho(F_X)$.
Then we can estimate risk measures using empirical distributions:

$$F_X^{emp}(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{x \geq x_i}$$

where $X = (x_1, x_2, \dots, x_n)$ is a random sample, we denote such estimation as

$$\hat{\rho}^h(X) = \rho(F_X^{emp}).$$

Whenever $\rho_m(X) = -\int_0^1 q_u^+(X) m(du)$ we have
$\hat{\rho}^h(X) = -\sum_{i=1}^n x_{(i)} m((\frac{i-1}{n}, \frac{i}{n}])$ where $x_{(i)}$ is the i-th element of $X = (x_1, x_2, \dots, x_n)$.
Consider now the examples of estimators:
1. $V\hat{a}R_\alpha^h(X) = -x_{([n\alpha]+1)}$
2. $\hat{ES}_\alpha^h(X) = -\frac{1}{n\alpha}\left(\sum_{i=1}^{[n\alpha]} x_{(i)} + x_{([n\alpha]+1)}(n\alpha - [n\alpha]\right)$
3. $\hat{\rho}_\phi^h(X) = -\sum_{i=1}^n x_{(i)} \int_{\frac{(i-1)}{n}}^{\frac{i}{n}} \phi(u) du$

12.6 Consistency of risk estimators

Risk estimator is **consistent** for distribution F whenever

$$\hat{\rho}(X_1, X_2, \dots, X_n) \to \rho(F),$$

as $n \to \infty$ a.s. for and i.i.d. sequence X_i with distribution F
In particular,
$V\hat{a}R_\alpha^h(F)$ is consistent for F such that $q_\alpha^+(F) = q_\alpha^-(F)$,
$\hat{ES}_\alpha^h(F)$ and $\hat{\rho}_\phi^h(X)$ are consistent as linear functions of order statistics (van Zwet 1980)
For further consideration it will be important for us to introduce so called Levy distance between distribution functions:

$$d(F, G) = \inf\{\epsilon > 0 : F(x - \epsilon) - \epsilon \leq G(x) \leq F(x + \epsilon) + \epsilon, \quad for \quad x \in R\}$$

It is clear that it is a metric which is consistent with the weak topology of laws (generated by random variables with distributions F and G.
Denote by
$\mathcal{L}_n(\hat{\rho}, G)$ the distribution of the estimator of ρ based on the sample $(X_1, X_2 \dots, X_n)$ of i.i.d. random variables with distribution G.

12.7 Robustness of risk estimators

Consider a family $\mathcal{C}$ of distribution functions.

Definition 12 *A risk estimator $\hat{\rho}$ is $\mathcal{C}$ robust at F if for any $\epsilon > 0$ there exists $\delta > 0$ nad $N_0 \geq 1$ such that for all $G \in \mathcal{C}$ we have*
$d(F, G) \leq \delta$ implies $d(\mathcal{L}_n(\hat{\rho}, G), \mathcal{L}_n(\hat{\rho}, F)) \leq \epsilon$ for $n \geq n_0$.

The notion of robustness is one of principal requirements in estimation theory. We have

Theorem 15 *(Cont, Deguest, Scandolo 2010) Let ρ be a risk measure and $F \in \mathcal{C}$. If $\hat{\rho}^h$ - the historical estimator of ρ is consistent with ρ at every $G \in \mathcal{C}$, then the following are equivalent*
1. the restriction of ρ to $\mathcal{C}$ is continuous at F (with respect to Levy distance)
2. $\hat{\rho}^h$ is $\mathcal{C}$ robust at F.

Corollary 3 *If ρ is continuous in $\mathcal{C}$ then $\hat{\rho}^h$ is $\mathcal{C}$ robust at any $F \in \mathcal{C}$.*

Proof. (We use Glivienko-Cantelli theorem which says that empirical distributions converge uniformly to approximated distribution with probability 1 and in the case of continuous distributions such convergence does not depend on the form of (continuous) distribution)

12.8 Continuity of risk measures

We consider again risk measure $\rho_m(X) = -\int_0^1 q_u^+(X) m(du)$, and let $\mathcal{D}_\rho$ be the space of distributions for which ρ_m is defined, and

$$\mathcal{A}_m = \{\alpha \in [0,1] : m\{\alpha\} > 0\}$$

We have the following result

Theorem 16 *(Huber 1981). If the support of m does not contain 0 nor 1 then ρ_m is continuous for any $F \in \mathcal{D}_\rho$ such that $q_\alpha^+(F) = q_\alpha^-(F)$ for any $\alpha \in \mathcal{A}_m$. Otherwise ρ_m is not continuous at any $F \in \mathcal{D}_\rho$.*

Directly from this theorem we obtain

Corollary 4 *$\hat{VaR}_\alpha^h(F)$ is $\mathcal{C}_\alpha$ robust at any $F \in \mathcal{C}_\alpha$ where*
$\mathcal{C}_\alpha = \{F \in \mathcal{D} : q_\alpha^+(F) = q_\alpha^-(F)\}$

Let $m(du) = \phi(u)du$ and $\phi \in L^q$. Then ρ_ϕ is defined in $\mathcal{D}^p$ - the set of distributions with finite left tail p moment ($\frac{1}{p} + \frac{1}{q} = 1$). We have

Corollary 5 *For any $F \in \mathcal{D}^p$ the historical estimator of ρ_ϕ is $\mathcal{D}^p$ robust at F if and only if for some $\epsilon > 0$ $\phi(u) = 0$ for all $u \in (0, \epsilon) \cup (1 - \epsilon, 1)$ i.e. ϕ vanishes in a neighborhood of 0 and 1.*

Therefore

Corollary 6 *The historical estimator of any spectral risk measure ρ_ϕ defined on $\mathcal{D}^p$ is not $\mathcal{D}^p$ - robust at any $F \in \mathcal{D}^p$. In particular, the historical estimator ES_α is not $\mathcal{D}^1$ - robust at any $F \in \mathcal{D}^1$.*

This is an important deficiency of spectral risk measures and consequently of coherent measures of risk. On the other hand VaR is not coherent but its historical estimator is robust. There is a long discussion among risk specialist which measure is good in practice. It seems that it should be a measure which is between coherent risk measures and value at risk. In particular the historical estimator of the risk measure

$$\frac{1}{\alpha_2 - \alpha_1} \int_{\alpha_1}^{\alpha_2} VaR_u(F) du$$

is a $\mathcal{D}^1$ robust and such measure is not far away from coherent measures of risk.

12.9 References

Acerbi, C., Spectral measures of risk: a coherent representation of subjective risk aversion. J. Bank. Finance, 2002, 26, 1505-1518.

Artzner, P., Delbaen, F., Eber, J. and Heath, D., Coherent measures of risk. Math. Finance, 1999, 9, 203-228.

Cont, R., Deguest, R., and Scandolo, G., Robustness and sensitivity analysis of risk meausurement procedures. Quantitative Finance, 10(6):593–606, 2010.

Föllmer, H. and Schied, A., Convex measures of risk and trading constraints. Finance Stochast., 2002, 6, 429-447.

Föllmer, H. and Schied, A., Stochastic Finance: An Introduction in Discrete Time, 2004 (Walter de Gruyter: Berlin).

Huber, P., Robust Statistics, 1981 (Wiley: New York).

Kusuoka, S., On law invariant coherent risk measures. Adv. Math. Econ., 2001, 3, 83-95.

Van Zwet, W., A strong law for linear functions of order statistics. Ann. Probab., 1980, 8, 986-990.