Large deviations in finance

Huyên PHAM *

Third SMAI European Summer School in Financial Mathematics

Paris, August 23-27, 2010

*University Paris Diderot and Institut Universitaire de France

Outline

Lecture I. **An overview of large deviations theory: methods and results**

Lecture II. **Large deviations in option pricing**

Lecture III. **Large deviations in risk management**

Lecture I. An overview of large deviations theory: methods and results

0. Introduction

1. Laplace transform and change of probability measures

2. Cramer's theorem and large deviations of level 1

3. Large deviations and Laplace principles

4. Relative entropy and Donsker-Varadhan formula

5. Sanov's theorem and large deviations of level 2

6. Freidlin-Wentzell theory and large deviations of level 3

0. Introduction

• Large deviations theory: **asymptotic estimates of probabilities of rare events** A^ε $\leftrightarrow$ **random processes** X^ε

$$\mathbb{P}[A^\varepsilon] \;=\; C_\varepsilon \,\exp\Big(-\frac{I}{\varepsilon}\Big) \;=\; \exp\Big(-\frac{I}{\varepsilon} + o(1/\varepsilon)\Big)$$

for some $I > 0$, and (C_ε) sequence converging at a subexponential rate, i.e. $\varepsilon \ln C_\varepsilon \to 0$, as ε goes to zero;

I is the leading order term on logarithm scale in large deviations: rate function

C_ε is the correction term.

- Large deviations results $\leftrightarrow$ change of probability measures under which the event A^ε (rare under $\mathbb{P}$) is non longer rare under $\mathbb{P}^\varepsilon$.

- Typically, the Radon-Nikodym $\frac{d\mathbb{P}}{d\mathbb{P}^\varepsilon}$ has an exponential form

- One needs to determine the dominant contribution to the exponent (when ε is small)

- Large deviations results $\leftrightarrow$ change of probability measures under which the event A^ε (rare under $\mathbb{P}$) is non longer rare under $\mathbb{P}^\varepsilon$.

  - Typically, the Radon-Nikodym $\frac{d\mathbb{P}}{d\mathbb{P}^\varepsilon}$ has an exponential form

  - One needs to determine the dominant contribution to the exponent (when ε is small)

▶ Change of probability measures useful in **simulating rare events**

  - If $p_\varepsilon = \mathbb{P}[A^\varepsilon]$ is small, then sampling according to $\mathbb{P}$ is unlikely to produce $A^\varepsilon \to$ high relative error in estimating p_ε

  - Quick simulation samples according to $\mathbb{P}^\varepsilon$, which gives more weight to the rare but important outcomes of A^ε: importance sampling method

• Large deviation rate function $\leftrightarrow$ entropy.

Illustration through an elementary example:

Throw a (fair) dice n times and set f_i: the frequency of number $i = 1, \ldots, 6$

Denote by $p_n(f)$: the probability that the numbers $1, \ldots, 6$ appear with frequencies $f = (f_1 = n_1/n, \ldots, f_6 = n_6/n)$ in the n throws of dices:

$$p_n(f) = \frac{1}{6^n} \frac{n!}{n_1! \ldots n_6!}$$

• Large deviation rate function $\leftrightarrow$ entropy.

Illustration through an elementary example:

Throw a (fair) dice n times and set f_i: the frequency of number $i = 1, \dots, 6$

Denote by $p_n(f)$: the probability that the numbers $1, \dots, 6$ appear with frequencies $f = (f_1 = n_1/n, \dots, f_6 = n_6/n)$ in the n throws of dices:

$$p_n(f) \;=\; \frac{1}{6^n}\frac{n!}{n_1! \dots n_6!}$$

Using Stirling formula: $k! \simeq k^k e^{-k}\sqrt{2\pi k}$, we get when n is large:

$$\frac{1}{n}\ln p_n(f) \simeq \;-I(f) := -\sum_{i=1}^{6} f_i \ln \frac{f_i}{\frac{1}{6}}$$

$I(f) \geq 0$ is the **relative entropy** of the *a posteriori* probability $f = (f_i)$ with respect to the *a priori* probability $r = (\frac{1}{6})$.

• Hence $p_n(f) = \exp(-nI(f) + o(n))$

▶ This means that when n is large, $p_n(f)$ is **concentrated where** $I(f)$ **is minimal**.

• Hence $p_n(f) = \exp(-nI(f) + o(n))$

▶ This means that when n is large, $p_n(f)$ is **concentrated where** $I(f)$ **is minimal**.

▶ The minimizing point is attained for $f^* = (1/6, \ldots, 1/6)$, and $I(f^*) = 0$: this is the ordinary law of large numbers!

▶ For $f \neq f^*$, $I(f) > 0$, and $p_n(f)$ tends to zero exponentially small!

▶ These ideas, concepts and computations in large deviations (**concentration phenomenon, entropy functional minimization**, etc ...) still hold in general random contexts, including diffusion processes,

$\rightarrow$ but need more sophisticated mathematical treatments.

1. Laplace transform and exponential change of measures

Let X be a real-valued random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with probability distribution $\mu(dx)$.

The logarithm **Laplace** (or **moment generating**) function of X is:

$$\Gamma(\theta) \;=\; \ln \mathbb{E}[e^{\theta X}] \;=\; \ln \int e^{\theta x} \mu(dx) \;\in\; (-\infty, \infty], \quad \theta \in \mathbb{R}.$$

• $\Gamma(0) = 0$, Γ convex (Hölder inequality).

• For any $\theta \in \mathcal{D}(\Gamma) = \{\theta \in \mathbb{R} : \Gamma(\theta) < \infty\}$, we define a probability measure μ_θ on $\mathbb{R}$ by:

$$\mu_\theta(dx) \;=\; \exp(\theta x - \Gamma(\theta))\mu(dx).$$

• Let $X_1, \ldots, X_n$ i.i.d. $\rightsquigarrow \mu$, and consider the probability measure $\mathbb{P}_\theta$ on $(\Omega, \mathcal{F})$:

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}}(X_1, \ldots, X_n) \;=\; \prod_{i=1}^n \frac{d\mu_\theta}{d\mu}(X_i) \;=\; \exp\Big(\theta \sum_{i=1}^n X_i - n\Gamma(\theta)\Big).$$

$\Longleftrightarrow$ (Bayes formula)

$$\mathbb{E}\Big[f(X_1, \ldots, X_n)\Big] \;=\; \mathbb{E}_\theta\Big[f(X_1, \ldots, X_n) \exp\Big(-\theta \sum_{i=1}^n X_i + n\Gamma(\theta)\Big)\Big],$$

• For any θ in the interior of $\mathcal{D}(\Gamma)$:

$$\mathbb{E}_\theta[X] \;=\; \Gamma'(\theta), \qquad \mathrm{Var}_\theta[X] \;=\; \Gamma''(\theta).$$

In particular if $0 \in \mathrm{int}(\mathcal{D}(\Gamma))$, then $\mathbb{E}[X] = \Gamma'(0)$, $\mathrm{Var}[X] = \Gamma''(0)$.

Examples

Bernoulli distribution: let $\mu \rightsquigarrow \mathcal{B}(p)$. Then

$$\Gamma(\theta) \;=\; \ln(1-p+pe^{\theta}), \quad \text{and} \;\; \mu_\theta \rightsquigarrow \mathcal{B}(p_\theta),\; p_\theta = \frac{pe^{\theta}}{1-p+pe^{\theta}}.$$

Poisson distribution: let $\mu \rightsquigarrow \mathcal{P}(\lambda)$. Then

$$\Gamma(\theta) \;=\; \lambda(e^{\theta}-1), \quad \text{and} \;\; \mu_\theta \rightsquigarrow \mathcal{P}(\lambda e^{\theta})$$

Normal distribution: let $\mu \rightsquigarrow \mathcal{N}(0,\sigma^2)$. Then

$$\Gamma(\theta) \;=\; \frac{\theta^2\sigma^2}{2}, \quad \text{and} \;\; \mu_\theta \rightsquigarrow \mathcal{N}(\theta\sigma^2,\sigma^2)$$

Exponential distribution: let $\mu \rightsquigarrow \mathcal{E}(\lambda)$. Then

$$\Gamma(\theta) \;=\; \begin{cases} \ln\left(\frac{\lambda}{\lambda-\theta}\right), & \theta < \lambda \\ \infty, & \theta \geq \lambda \end{cases} \quad \text{and} \;\; \mu_\theta \rightsquigarrow \mathcal{E}(\lambda-\theta)$$

2. Cramer's theorem

Large deviations of level 1: concern random variables valued in a finite-dimensional space.

Let (X_i) be an i.i.d. sequence of real random variables with probability distribution μ and finite mean $\bar{x} = \mathbb{E}[X_1] = \int x\mu(dx) < \infty$, and consider the empirical mean:

$$\bar{S}_n = \frac{S_n}{n}, \quad S_n = \sum_{i=1}^{n} X_i$$

▶ By the law of large numbers, $\bar{S}_n$ converges in probability to $\bar{x}$.

▶ Cramer's theorem focus on the asymptotics for probabilities of rare events, e.g. $\mathbb{P}[\bar{S}_n \geq x]$, for $x > \bar{x}$, and states that

$$\frac{1}{n} \ln \mathbb{P}\left[\bar{S}_n \geq x\right] \rightarrow -\gamma < 0.$$

The rate of convergence is determined by the **Fenchel-Legendre** transform of the Log-Laplace function Γ of X_i:

$$\Gamma^*(x) \;=\; \sup_{\theta\in\mathbb{R}} \Big[\theta x - \Gamma(\theta)\Big] \;\in\; [0,\infty], \quad x \in \mathbb{R}.$$

▶ Γ^* is convex, $\Gamma^*(\bar{x}) = 0$, $\Gamma^*(x) = \sup_{\theta\geq 0}\Big[\theta x - \Gamma(\theta)\Big]$, for $x \geq \bar{x}$, and so Γ^* is nondecreasing on $[\bar{x},\infty)$

▶ Given $x \in \mathbb{R}$, if $\theta = \theta(x)$ is solution to the saddle-point equation: $x = \Gamma'(\theta)$, then $\Gamma^*(x) = \theta x - \Gamma(\theta)$, and

$$\mathbb{E}_\theta[X_i] \;=\; x.$$

Cramer's theorem. For any $x \geq \bar{x}$, we have

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\left[\bar{S}_n \geq x\right] \;=\; -\Gamma^*(x) \;=\; -\inf_{y\geq x} \Gamma^*(y).$$

Cramer's theorem. For any $x \geq \bar{x}$, we have

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\Big[\bar{S}_n \geq x\Big] \;=\; -\Gamma^*(x) \;=\; -\inf_{y\geq x} \Gamma^*(y).$$

Proof (Sketch). *Upper bound.* The main step in the upper bound $\leq$ is based on Chebichev inequality combined with the i.i.d. assumption on the X_i :

$$\mathbb{P}\Big[\bar{S}_n \geq x\Big] \;=\; \mathbb{E}\Big[1_{\frac{S_n}{n}\geq x}\Big] \;\leq\; \mathbb{E}\Big[e^{\theta(S_n - nx)}\Big] \;=\; \exp\Big(n\Gamma(\theta) - \theta n x\Big), \quad \forall\theta \geq 0.$$

Cramer's theorem. For any $x \geq \bar{x}$, we have

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\Big[\bar{S}_n \geq x\Big] \;=\; -\Gamma^*(x) \;=\; -\inf_{y\geq x} \Gamma^*(y).$$

Proof (Sketch). *Upper bound.* The main step in the upper bound $\leq$ is based on Chebichev inequality combined with the i.i.d. assumption on the X_i :

$$\mathbb{P}\Big[\bar{S}_n \geq x\Big] \;=\; \mathbb{E}\Big[1_{\frac{S_n}{n}\geq x}\Big] \;\leq\; \mathbb{E}\Big[e^{\theta(S_n - nx)}\Big] \;=\; \exp\Big(n\Gamma(\theta) - \theta n x\Big), \quad \forall \theta \geq 0.$$

By taking the infimum over $\theta \geq 0$, and by definition of Γ^*, we get

$$\mathbb{P}\Big[\bar{S}_n \geq x\Big] \;\leq\; \exp\Big(-n\Gamma^*(x)\Big),$$

and we conclude by taking logarithm.

Lower bound. For simplicity, assume that there exists a solution $\theta = \theta(x) > 0$ to the saddle-point equation: $\Gamma'(\theta) = x$, i.e. attaining the supremum in $\Gamma^*(x) = \theta(x)x - \Gamma(\theta(x))$. The key step is now to introduce the new probability distribution μ_θ and $\mathbb{P}_\theta$ the corresponding probability measure on $(\Omega, \mathcal{F})$ with likelihood ratio:

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} \;=\; \prod_{i=1}^{n} \frac{d\mu_\theta}{d\mu}(X_i) \;=\; \exp\Big(\theta S_n - n\Gamma(\theta)\Big),$$

so that $\mathbb{E}_\theta[X_i] = x$ (the event $\{\bar{S}_n \geq x\}$ is no longer rare under $\mathbb{P}_\theta$).

Lower bound. For simplicity, assume that there exists a solution $\theta = \theta(x) > 0$ to the saddle-point equation: $\Gamma'(\theta) = x$, i.e. attaining the supremum in $\Gamma^*(x) = \theta(x)x - \Gamma(\theta(x))$. The key step is now to introduce the new probability distribution μ_θ and $\mathbb{P}_\theta$ the corresponding probability measure on $(\Omega, \mathcal{F})$ with likelihood ratio:

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = \prod_{i=1}^{n} \frac{d\mu_\theta}{d\mu}(X_i) = \exp\Big(\theta S_n - n\Gamma(\theta)\Big),$$

so that $\mathbb{E}_\theta[X_i] = x$ (the event $\{\bar{S}_n \geq x\}$ is no longer rare under $\mathbb{P}_\theta$). Then, we have for all $\varepsilon > 0$:

$$\begin{aligned}
\mathbb{P}\Big[\bar{S}_n \in [x, x+\varepsilon)\Big] &= \mathbb{E}_\theta\Big[\exp\Big(-\theta S_n + n\Gamma(\theta)\Big) 1_{\frac{S_n}{n} \in [x, x+\varepsilon)}\Big] \\
&= e^{-n(\theta x - \Gamma(\theta))} \mathbb{E}_\theta\Big[\exp\Big(-n\theta(\frac{S_n}{n} - x)\Big) 1_{\frac{S_n}{n} \in [x, x+\varepsilon)}\Big] \\
&\geq e^{-n(\theta x - \Gamma(\theta))} e^{-n|\theta|\varepsilon} \mathbb{P}_\theta\Big[\bar{S}_n \in [x, x+\varepsilon)\Big],
\end{aligned}$$

Taking logarithm:

$$
\begin{aligned}
\frac{1}{n}\ln\mathbb{P}\Big[\bar{S}_n \in [x, x+\varepsilon)\Big] &\geq -[\theta x - \Gamma(\theta)] - |\theta|\varepsilon + \frac{1}{n}\ln\mathbb{P}_\theta\Big[\bar{S}_n \in [x, x+\varepsilon)\Big] \\
&= -\Gamma^*(x) - |\theta|\varepsilon + \frac{1}{n}\ln\mathbb{P}_\theta\Big[\bar{S}_n \in [x, x+\varepsilon)\Big]
\end{aligned}
$$

Now, since $\mathbb{E}_\theta[X_i] = x$, we have: $\lim_n \mathbb{P}_\theta\Big[\bar{S}_n \in [x, x+\varepsilon)\Big] = 1/2\ (> 0)$. Thus,

$$
\begin{aligned}
\liminf_{n\to\infty}\frac{1}{n}\ln\mathbb{P}\Big[\bar{S}_n \geq x\Big] &\geq \lim_{\varepsilon\to 0}\liminf_{n\to\infty}\frac{1}{n}\ln\mathbb{P}\Big[\bar{S}_n \in [x, x+\varepsilon)\Big] \\
&\geq -\Gamma^*(x).
\end{aligned}
$$

$\square$

Examples

Bernoulli distribution: let $X_i \rightsquigarrow \mathcal{B}(p)$. Then

$$\Gamma^*(x) \;=\; x \ln\left(\frac{x}{p}\right) + (1-x)\ln\left(\frac{1-x}{1-p}\right) \text{ for } x \in [0,1] \text{ and } \infty \text{ otherwise.}$$

Poisson distribution: let $X_i \rightsquigarrow \mathcal{P}(\lambda)$. Then

$$\Gamma^*(x) \;=\; x \ln\left(\frac{x}{\lambda}\right) + \lambda - x, \quad \text{for } \; x \geq 0, \text{ and } \infty \text{ otherwise.}$$

Normal distribution: let $X_i \rightsquigarrow \mathcal{N}(0,\sigma^2)$. Then

$$\Gamma^*(x) \;=\; \frac{x^2}{2\sigma^2}, \quad x \in \mathbb{R}.$$

Exponential distribution: let $X_i \rightsquigarrow \mathcal{E}(\lambda)$. Then

$$\Gamma^*(x) \;=\; \lambda x - 1 - \ln(\lambda x) \quad \text{for } x > 0 \text{ and } \infty \text{ otherwise.}$$

Remark 1.

Cramer's theorem possesses a multivariate counterpart dealing with the large deviations of the empirical means of i.i.d. random vectors in $\mathbb{R}^d$.

Remark 2.

The independence of the random variables X_i in the large deviations result for the empirical mean $\bar{S}_n = \sum_{i=1}^n X_i/n$ can be relaxed with the **Gärtner-Ellis theorem**, once we get the existence of the limit:

$$\Gamma(\theta) \;:=\; \lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}\Big[e^{n\theta.\bar{S}_n}\Big], \;\; \theta \in \mathbb{R}^d.$$

The rate of convergence of the large deviation principle is then given by the Fenchel-Legendre transform of Γ:

$$\Gamma^*(x) \;=\; \sup_{\theta\in\mathbb{R}^d} [\theta.x - \Gamma(\theta)], \quad x \in \mathbb{R}^d,$$

under the condition that Γ is **steep**, i.e. $\Gamma'(\theta_n) \to \infty$ for any sequence (θ_n) converging to a boundary point of the domain of Γ (this ensures the existence of a saddle-point for any $x \in \mathbb{R}^d$).

Remark 3: Relation with importance sampling

Fix n and let us consider the estimation of $p_n = \mathbb{P}[\bar{S}_n \geq x]$. A standard estimator for p_n is the average with N independent copies of $X = 1_{\bar{S}_n \geq x} \rightarrow$

$$\text{relative error} = \frac{\text{standard deviation}}{\text{mean}} = \frac{\sqrt{p_n(1-p_n)}}{p_n\sqrt{N}}.$$

Since p_n is extremely small, we see that a large sample size N is required for the estimator to achieve a reasonable relative error bound.

Remark 3: Relation with importance sampling

Fix n and let us consider the estimation of $p_n = \mathbb{P}[\bar{S}_n \geq x]$. A standard estimator for p_n is the average with N independent copies of $X = 1_{\bar{S}_n \geq x} \rightarrow$

$$\text{relative error} = \frac{\text{standard deviation}}{\text{mean}} = \frac{\sqrt{p_n(1-p_n)}}{p_n\sqrt{N}}.$$

Since p_n is extremely small, we see that a large sample size N is required for the estimator to achieve a reasonable relative error bound.

▶ By using an exponential change of measure $\mathbb{P}_\theta$ with likelihood ratio

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = \exp\Big(\theta S_n - n\Gamma(\theta)\Big),$$

so that

$$p_n = \mathbb{E}_\theta\Big[\exp\Big(-\theta S_n + n\Gamma(\theta)\Big)1_{\bar{S}_n \geq x}\Big],$$

we have an importance sampling (IS) (unbiased) estimator of p_n, by taking the average of independent replications (under $\mathbb{P}_\theta$) of

$$\exp\Big(-\theta S_n + n\Gamma(\theta)\Big)1_{\bar{S}_n \geq x}.$$

The parameter θ is chosen in order to **minimize the variance** of this estimator, or equivalently **its second moment**:

$$M_n^2(\theta, x) = \mathbb{E}_\theta\Big[\exp\Big(-2\theta S_n + 2n\Gamma(\theta)\Big)1_{\bar{S}_n \geq x}\Big]$$
$$\leq \exp\Big(-2n(\theta x - \Gamma(\theta))\Big)$$

The parameter θ is chosen in order to **minimize the variance** of this estimator, or equivalently **its second moment**:

$$
\begin{aligned}
M_n^2(\theta, x) &= \mathbb{E}_\theta\Big[\exp\Big(-2\theta S_n + 2n\Gamma(\theta)\Big)1_{\bar{S}_n \geq x}\Big] \\
&\leq \exp\Big(-2n(\theta x - \Gamma(\theta))\Big) \qquad (1)
\end{aligned}
$$

By noting from Cauchy-Schwarz's inequality that $M_n^2(\theta, x) \geq p_n^2 = \mathbb{P}[\bar{S}_n \geq x] \simeq Ce^{-2n\Gamma^*(x)}$ as n goes to infinity, from Cramer's theorem, we see that the fastest possible exponential rate of decay of $M_n^2(\theta, x)$ is twice the rate of the probability itself, i.e. $2\Gamma^*(x)$. Hence, from (1), and with the choice of $\theta = \theta(x)$ s.t. $\Gamma^*(x) = \theta(x)x - \Gamma(\theta(x))$, we get an asymptotic optimal IS estimator in the sense that :

$$
\lim_{n\to\infty} \frac{1}{n} \ln M_n^2(\theta_x, x) = 2 \lim_{n\to\infty} \frac{1}{n} \ln p_n.
$$

This parameter $\theta(x)$ is such that $\mathbb{E}_{\theta(x)}[\bar{S}_n] = x$ so that the event $\{\bar{S}_n \geq x\}$ is no more rare under $\mathbb{P}_{\theta(x)}$, and is precisely the parameter used in the derivation of the large deviations result in Cramer's theorem.

3. Large deviations and Laplace principles

General definition of a large deviation principle (LDP)

Consider a sequence $\{X^\varepsilon\}_\varepsilon$ on $(\Omega, \mathcal{F}, \mathbb{P})$ valued in some topological space $\mathcal{X}$.

The LDP characterizes the limiting behaviour as $\varepsilon \to 0$ of the family of probability measures $\{\mathbb{P}[X^\varepsilon \in dx]\}_\varepsilon$ on $\mathcal{X}$ in terms of a **rate function**.

A *rate function* I is a lower semicontinuous function mapping $I : \mathcal{X} \to [0, \infty]$ such that the level sets $\{x \in \mathcal{X} \;: I(x) \leq M\}$ are compact for all $M < \infty$.

The sequence $\{X^\varepsilon\}_\varepsilon$ satisfies a LDP on $\mathcal{X}$ with rate function I (and speed ε) if :

(i) **Upper bound** : for any closed subset F of $\mathcal{X}$

$$\limsup_{\varepsilon\to 0} \varepsilon \ln \mathbb{P}[X^\varepsilon \in F] \;\leq\; -\inf_{x\in F} I(x).$$

(ii) **Lower bound** : for any open subset G of $\mathcal{X}$

$$\liminf_{\varepsilon\to 0} \varepsilon \ln \mathbb{P}[X^\varepsilon \in G] \;\geq\; -\inf_{x\in G} I(x).$$

If F is a subset of $\mathcal{X}$ s.t. $\inf_{x\in F^o} I(x) = \inf_{x\in \bar{F}} I(x) := I_F$, then

$$\lim_{\varepsilon\to 0} \varepsilon \ln \mathbb{P}[X^\varepsilon \in F] \;=\; -I_F,$$

which means that $\mathbb{P}[X^\varepsilon \in F] = \exp\Big(-\frac{I_F}{\varepsilon} + o(1/\varepsilon)\Big)$. The classical Cramer's theorem considered the case of the empirical mean $X^\varepsilon = S_n/n$ of i.i.d. random variables in $\mathbb{R}^d$, with $\varepsilon = 1/n$.

Transformation of LDP by contraction principle

The LDP is preserved under continuous mappings.

Suppose that $\{X^\varepsilon\}_\varepsilon$ satisfies a LDP on $\mathcal{X}$ with rate function I, and let f be a continuous mapping from $\mathcal{X}$ to $\mathcal{Y}$.

Then $\{f(X^\varepsilon)\}_\varepsilon$ satisfies a LDP on $\mathcal{Y}$ with rate function:

$$J(y) \;=\; \inf\{I(x) : x \in \mathcal{X}, y = f(x)\}.$$

Remark

If f is a continuous bijection, then $J(.) = I(f^{-1}(.))$.

Laplace method gives an equivalent formulation of LDP, relying on **Varadhan's formula**.

Theorem (Varadhan)

Suppose that $\{X^\varepsilon\}_\varepsilon$ satisfies a LDP on $\mathcal{X}$ with good rate function I. Then, $\{X^\varepsilon\}_\varepsilon$ satisfies the **Laplace principle**: for any bounded continuous function $\varphi : \mathcal{X} \to \mathbb{R}$, we have

$$\lim_{\varepsilon\to 0} \varepsilon \ln \mathbb{E}\Big[e^{\varphi(X^\varepsilon)/\varepsilon}\Big] \;=\; \sup_{x\in\mathcal{X}} \Big[\varphi(x) - I(x)\Big].$$

Remark. This can be viewed as a stochastic extension of the (deterministic) Laplace integral's formula:

$$\lim_{\varepsilon\to 0} \varepsilon \ln \int_0^1 e^{\varphi(x)/\varepsilon} dx \;=\; \sup_{x\in[0,1]} \varphi(x).$$

Laplace method gives an equivalent formulation of LDP, relying on **Varadhan's formula**.

Theorem (Varadhan)

Suppose that $\{X^\varepsilon\}_\varepsilon$ satisfies a LDP on $\mathcal{X}$ with good rate function I. Then, $\{X^\varepsilon\}_\varepsilon$ satisfies the **Laplace principle**: for any bounded continuous function $\varphi : \mathcal{X} \to \mathbb{R}$, we have

$$\lim_{\varepsilon\to 0} \varepsilon \ln \mathbb{E}\Big[e^{\varphi(X^\varepsilon)/\varepsilon}\Big] \;=\; \sup_{x\in\mathcal{X}} \Big[\varphi(x) - I(x)\Big].$$

Interpretation. By writing formally the LDP for (X^ε) with rate function I as $\mathbb{P}[X^\varepsilon \in dx] \simeq e^{-I(x)/\varepsilon}dx$, we can write

$$\begin{aligned}\mathbb{E}\Big[e^{\varphi(X^\varepsilon)/\varepsilon}\Big] \;=\; \int e^{\varphi(x)/\varepsilon}\mathbb{P}[X^\varepsilon \in dx] \;&\simeq\; \int e^{(\varphi(x)-I(x))/\varepsilon}dx \\ &\simeq\; C\exp\left(\frac{\sup_{x\in\mathcal{X}}(\varphi(x)-I(x))}{\varepsilon}\right).\end{aligned}$$

As in Laplace's method for integrals, Varadhan's formula states that to exponential order, the main contribution to the integral is due to the largest value of the exponent.

Theorem

The Laplace principle implies the large deviation principle with the same good rate function. More precisely, if I is rate function on $\mathcal{X}$ and the limit

$$\lim_{\varepsilon \to 0} \varepsilon \ln \mathbb{E}\left[e^{\varphi(X^\varepsilon)/\varepsilon}\right] = \sup_{x \in \mathcal{X}} \left[\varphi(x) - I(x)\right]$$

is valid for all bounded continuous functions φ, then (X^ε) satisfies a LDP on $\mathcal{X}$ with rate function I.

Theorem

The Laplace principle implies the large deviation principle with the same good rate function. More precisely, if I is rate function on $\mathcal{X}$ and the limit

$$\lim_{\varepsilon\to 0} \varepsilon \ln \mathbb{E}\Big[e^{\varphi(X^\varepsilon)/\varepsilon}\Big] = \sup_{x\in\mathcal{X}} \Big[\varphi(x) - I(x)\Big]$$

is valid for all bounded continuous functions φ, then (X^ε) satisfies a LDP on $\mathcal{X}$ with rate function I.

Formal proof. Given $F \subset \mathcal{X}$, consider: $\psi(x) = 0$ if $x \in F$, and ∞ otherwise.

$$\begin{aligned} \varepsilon \ln \mathbb{P}[X^\varepsilon \in F] &= \varepsilon \ln \mathbb{E}[\exp(-\psi(X^\varepsilon)/\varepsilon)] \\ &\to \sup_{x\in\mathcal{X}}[-\psi(x) - I(x)] = -\inf_{x\in F} I(x). \end{aligned}$$

4. Relative entropy and Donsker-Varadhan formula

We are given a topological space $\mathcal{S}$, and we denote by $\mathcal{P}(\mathcal{S})$ the set of probability measures on $\mathcal{S}$ equipped with its Borel σ field.

For $\nu \in \mathcal{P}(\mathcal{S})$, the **relative entropy** $R(.|\nu)$ is a mapping from $\mathcal{P}(\mathcal{S})$ into $\bar{\mathbb{R}}$, defined by

$$R(\mu|\nu) = \begin{cases} \int_{\mathcal{S}} \left(\ln \frac{d\mu}{d\nu} \right) d\mu = \int_{\mathcal{S}} \frac{d\mu}{d\nu} \left(\ln \frac{d\mu}{d\nu} \right) d\nu, & \text{if } \mu \ll \nu \\ \infty, & \text{otherwise} \end{cases}$$

By observing that $s \ln s \geq s - 1$ with equality if and only if $s = 1$, we see that $R(\mu|\nu) \geq 0$, and $R(\mu|\nu) = 0$ if and only if $\mu = \nu$.

Proposition (Log-Laplace and relative entropy)

Let φ be a bounded measurable function on $\mathcal{S}$, i.e. $\varphi \in \mathcal{B}(\mathcal{S})$, and $\nu \in \mathcal{P}(\mathcal{S})$. Then,

$$\ln \int_{\mathcal{S}} e^{\varphi} d\nu \; = \; \sup_{\mu \in \mathcal{P}(\mathcal{S})} \left[\int_{\mathcal{S}} \varphi d\mu - R(\mu|\nu) \right],$$

and the supremum is attained uniquely by the probability measure μ_0:

$$\frac{d\mu_0}{d\nu} \; = \; \frac{e^{\varphi}}{\int_{\mathcal{S}} e^{\varphi} d\nu}.$$

Proposition (Log-Laplace and relative entropy)

Let φ be a bounded measurable function on $\mathcal{S}$, i.e. $\varphi \in \mathcal{B}(\mathcal{S})$, and $\nu \in \mathcal{P}(\mathcal{S})$. Then,

$$\ln \int_{\mathcal{S}} e^{\varphi} d\nu \;=\; \sup_{\mu \in \mathcal{P}(\mathcal{S})} \Big[\int_{\mathcal{S}} \varphi d\mu - R(\mu|\nu) \Big],$$

and the supremum is attained uniquely by the probability measure μ_0:

$$\frac{d\mu_0}{d\nu} \;=\; \frac{e^{\varphi}}{\int_{\mathcal{S}} e^{\varphi} d\nu}.$$

Proof. For any $\mu \ll \nu$, we have:

$$\begin{aligned}
\int_{\mathcal{S}} \varphi d\mu - R(\mu|\nu) &= \int_{\mathcal{S}} \varphi d\mu - \int_{\mathcal{S}} \Big(\ln \frac{d\mu}{d\nu} \Big) d\mu \\
&= \int_{\mathcal{S}} \varphi d\mu - \int_{\mathcal{S}} \Big(\ln \frac{d\mu}{d\mu_0} \Big) d\mu - \int_{\mathcal{S}} \Big(\ln \frac{d\mu_0}{d\nu} \Big) d\mu \\
&= \ln \int_{\mathcal{S}} e^{\varphi} d\nu - R(\mu|\mu_0).
\end{aligned}$$

We conclude by using the fact that $R(\mu|\mu_0) \geq 0$ and $R(\mu|\mu_0) = 0$ if and only if $\mu = \mu_0$. $\square$

Dual Proposition (Donsker-Varadhan variational formula)

For all μ, $\nu \in \mathcal{P}(\mathcal{S})$, we have

$$R(\mu|\nu) \;=\; \sup_{\varphi \in \mathcal{B}(\mathcal{S})} \left[\int_{\mathcal{S}} \varphi d\mu - \ln \int_{\mathcal{S}} e^{\varphi} d\nu \right]$$

Dual Proposition (Donsker-Varadhan variational formula)

For all μ, $\nu \in \mathcal{P}(\mathcal{S})$, we have

$$R(\mu|\nu) \;=\; \sup_{\varphi\in\mathcal{B}(\mathcal{S})}\left[\int_{\mathcal{S}}\varphi d\mu - \ln\int_{\mathcal{S}} e^{\varphi} d\nu\right]$$

Proof (Sketch). From the previous proposition, we have

$$\ln\int_{\mathcal{S}} e^{\varphi} d\nu \;\geq\; \int_{\mathcal{S}}\varphi d\mu - R(\mu|\nu)$$

Since this holds true for all $\varphi \in \mathcal{B}(\mathcal{S})$, we get

$$R(\mu|\nu) \;\geq\; \sup_{\varphi\in\mathcal{B}(\mathcal{S})}\left[\int_{\mathcal{S}}\varphi d\mu - \ln\int_{\mathcal{S}} e^{\varphi} d\nu\right] =: H(\mu,\nu).$$

Proof. (Ctd).

Conversely, let μ s.t. $H(\mu,\nu) < \infty$. Then $\mu \ll \nu$. Indeed, if $\nu(A) = 0$, then by considering $\varphi_n = n1_A$, we have for all n:

$$\infty \; > \; H(\mu,\nu) \; \geq \; \int_{\mathcal{S}} \varphi_n d\mu - \ln \int_{\mathcal{S}} e^{\varphi_n} d\nu \; = \; n\mu(A) \quad \text{and so } \mu(A) = 0.$$

Proof. (Ctd).

Conversely, let μ s.t. $H(\mu,\nu) < \infty$. Then $\mu \ll \nu$. Indeed, if $\nu(A) = 0$, then by considering $\varphi_n = n1_A$, we have for all n:

$$\infty \; > \; H(\mu,\nu) \;\; \geq \;\; \int_{\mathcal{S}} \varphi_n d\mu - \ln \int_{\mathcal{S}} e^{\varphi_n} d\nu \; = \; n\mu(A) \quad \text{and so } \mu(A) = 0.$$

Set $f = d\mu/d\nu$ and assume for simplicity that f is bounded and uniformly positive so that $\varphi = \ln f \in \mathcal{B}(\mathcal{S})$:

$$H(\mu,\nu) \;\; \geq \;\; \int_{\mathcal{S}} \varphi d\mu - \ln \int_{\mathcal{S}} e^{\varphi} d\nu \; = \; \int_{\mathcal{S}} \ln \frac{d\mu}{d\nu} d\mu \; = \; R(\mu|\nu).$$

$\square$

5. Sanov's theorem

Large deviations of level 2: concern random measures.

Let (X_i) be an i.i.d. sequence of random variables valued in $\mathcal{S}$ with probability distribution ρ, and consider the empirical measure valued in $\mathcal{P}(\mathcal{S})$:

$$L^n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$$

- By the law of large numbers, L^n converges weakly to ρ.

Theorem (Sanov) The sequence of empirical measures $(L^n)_n$ satisfies a LDP with rate function the relative entropy $R(.|\rho)$.

Idea of proof by Laplace method (Dupuis-Ellis)

Study the asymptotic behavior of

$$V^n \ := \ \frac{1}{n} \ln \mathbb{E}[\exp(n\varphi(L^n))],$$

where φ is any bounded continuous function mapping $\mathcal{P}(\mathcal{S})$ into $\mathbb{R}$.

Idea of proof by Laplace method (Dupuis-Ellis)

Study the asymptotic behavior of

$$V^n := \frac{1}{n} \ln \mathbb{E}[\exp(n\varphi(L^n))],$$

where φ is any bounded continuous function mapping $\mathcal{P}(\mathcal{S})$ into $\mathbb{R}$.

Corresponding dynamic problem. We introduce a sequence of random subprobability measures related to the empirical measures as follows. For $t \in [0,1]$, we denote $\mathcal{M}_t(\mathcal{S})$ the set of measures on $\mathcal{S}$ with total mass equal to t. Fix $n \in \mathbb{N}^*$, and for $i = 0, \ldots, n-1$, we define $L_0^n = 0$, and

$$L_{i+1}^n = L_i^n + \frac{1}{n}\delta_{X_{i+1}},$$

so that L_n^n equals the empirical measure L^n, and L_i^n is valued in $\mathcal{M}_{i/n}(\mathcal{S})$. We also introduce, for each $i = 0, \ldots, n$, and $\mu \in \mathcal{M}_{i/n}(\mathcal{S})$, the function

$$V^n(i,\mu) = \frac{1}{n} \ln \mathbb{E}_{i,\mu}[\exp(n\varphi(L_n^n))],$$

where $\mathbb{E}_{i,\mu}$ denotes the expectation conditioned on $L_i^n = \mu$. Thus, $V^n(0,0)$ $= V^n$, and $V^n(n,\mu) = \varphi(\mu)$.

Since X_i are i.i.d. $\rightsquigarrow \rho$, we see that the random measures $\{L_i^n, i = 0, \ldots, n\}$ form a Markov chain on state spaces $\{\mathcal{M}_{i/n}(\mathcal{S}), i = 0, \ldots, n\}$ with probability transition:

$$\mathbb{P}[L_{i+1}^n \in A | L_i^n = \mu] \ = \ \mathbb{P}[\mu + \frac{1}{n}\delta_{X_i} \in A] \ = \ \int_{\mathcal{S}} 1_A(\mu + \frac{1}{n}\delta_y)\rho(dy).$$

Since X_i are i.i.d. $\rightsquigarrow \rho$, we see that the random measures $\{L_i^n, i = 0, \ldots, n\}$ form a Markov chain on state spaces $\{\mathcal{M}_{i/n}(\mathcal{S}), i = 0, \ldots, n\}$ with probability transition:

$$\mathbb{P}[L_{i+1}^n \in A | L_i^n = \mu] \;=\; \mathbb{P}[\mu + \frac{1}{n}\delta_{X_i} \in A] \;=\; \int_{\mathcal{S}} 1_A(\mu + \frac{1}{n}\delta_y)\rho(dy).$$

By the law of iterated conditional expectations and Markov property:

$$\begin{aligned}
V^n(i,\mu) \;&=\; \frac{1}{n}\ln \mathbb{E}_{i,\mu}\Big[\mathbb{E}_{i+1,L_{i+1}^n}[\exp(n\varphi(L_n^n))]\Big] \\
&=\; \frac{1}{n}\ln \mathbb{E}_{i,\mu}\Big[\exp(nV^n(i+1, L_{i+1}^n))\Big] \\
&=\; \frac{1}{n}\ln \int_{\mathcal{S}} \exp\Big[nV^n(i+1, \mu + \frac{1}{n}\delta_y)\Big]\rho(dy).
\end{aligned}$$

From the variational formula relating Log-Laplace and relative entropy:

$$V^n(i,\mu) \;=\; \sup_{\nu\in\mathcal{P}(\mathcal{S})}\left[\int_{\mathcal{S}} V^n(i+1,\mu+\frac{1}{n}\delta_y)\nu(dy) - \frac{1}{n}R(\nu|\rho)\right]$$

From the variational formula relating Log-Laplace and relative entropy:

$$V^n(i,\mu) \;=\; \sup_{\nu\in\mathcal{P}(\mathcal{S})} \Big[\int_{\mathcal{S}} V^n(i+1,\mu+\frac{1}{n}\delta_y)\nu(dy) - \frac{1}{n}R(\nu|\rho)\Big]$$

$\rightarrow$ **dynamic programming equation** for the following stochastic control problem. The **controlled process** is a Markov chain $\{\bar{L}_i^n, i=0,\ldots,n\}$ starting from $\bar{L}_0^n = 0$, with controlled probability transitions:

$$\mathbb{P}[\bar{L}_{i+1}^n \in A|\bar{L}_i^n = \mu] \;=\; \int_{\mathcal{S}} 1_A(\mu + \frac{1}{n}\delta_y)\nu_i(dy),$$

where $\{\nu_i, i=0,\ldots,n\}$ is the control process valued in $\mathcal{P}(\mathcal{S})$, in feedback type, i.e. for each i, the decision ν_i depends on $\bar{L}_i^n$. The running gain is $-1/nR(\nu|\rho)$, and the terminal gain is φ. We deduce the **stochastic control representation formula**:

$$V^n \;=\; V^n(0,0) \;=\; \sup_{(\nu_i)} \mathbb{E}\Big[\varphi(\bar{L}_n^n) - \frac{1}{n}\sum_{i=0}^{n-1} R(\nu_i|\rho)\Big].$$

Asymptotic behavior of V^n

Fix some arbitrary $\nu \in \mathcal{P}(\mathcal{S})$, and consider the constant control $\nu_i = \nu$. With this choice, $\bar{L}_n^n$ is the empirical measure of i.i.d. r.v. $\rightsquigarrow \nu$, and the control representation for V^n yields

$$V^n \;\geq\; \mathbb{E}[\varphi(\bar{L}_n^n) - R(\nu|\rho)].$$

Since $\bar{L}_n^n$ converges weakly to ν, we have:

$$\lim_{n\to\infty} \mathbb{E}[\varphi(\bar{L}_n^n)] \;=\; \varphi(\nu).$$

Since ν is arbitrary in $\mathcal{P}(\mathcal{S})$, we deduce that

$$\liminf_{n\to\infty} V^n \;\geq\; \sup_{\nu\in\mathcal{P}(\mathcal{S})} [\varphi(\nu) - R(\nu|\rho)].$$

The corresponding upper-bound requires more technical details, and we get finally the **Laplace principle with rate function as relative entropy**:

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[\exp(n\varphi(L^n))] \;=\; \lim_{n\to\infty} V^n \;=\; \sup_{\nu\in\mathcal{P}(\mathcal{S})} [\varphi(\nu) - R(\nu|\rho)].$$

Remark

There are extensions of Sanov's theorem on LDP for empirical measure of Markov chains and occupation times of continuous-time Markov processes.

$\rightarrow$ Main references are the works by **Donsker and Varadhan**: Consider an ergodic Feller-Markov process X. Under some conditions, the occupation measure $L_t = \frac{1}{t}\int_0^t \delta_{X_s} ds$ satisfies a LDP with rate function I, and we have the Laplace principle:

$$\lim_{t\to\infty} \frac{1}{t} \ln \mathbb{E}\Big[\exp\Big(\int_0^t \phi(X_s)ds\Big)\Big] \;=\; \sup_{\mu\in\mathcal{P}(\mathcal{S})}\Big[\int \varphi d\mu - I(\mu)\Big].$$

for any bounded continuous function ϕ on $\mathcal{S}$,

6. Freidlin-Wentzell theory

Large deviations of level 3: concern random processes $\to$ sample path large deviations results

Key result: Schilder's theorem

$X^\varepsilon = \sqrt{\varepsilon} W$, with $W = (W_t)_{t \in [0,T]}$ Brownian motion in $\mathbb{R}^d$, valued in $C([0,T])$ the space of continuous functions on $[0,T]$.

(X^ε) **satisfies a LDP on** $C([0,T])$ **with rate function** (action functional):

$$I(h) = \begin{cases} \frac{1}{2}\int_0^T |\dot{h}(t)|^2 dt, & \text{if } h \in H_0([0,T]), \\ \infty, & \text{otherwise} \end{cases}$$

where $H_0([0,T]) = \{h \in H([0,T]) : h(0) = 0\}$, and $H([0,T])$ is the Cameron-Martin space consisting of absolutely continuous functions h, with square-integrable derivative $\dot{h}$.

Sketch of proof. (Lower bound)

Let $G \neq \emptyset$ be an open set of $C([0,T])$, $h \in G$, and $\delta > 0$ s.t. $B(h,\delta) \subset G$.

We want to prove that

$$\liminf_{\varepsilon \to 0} \varepsilon \ln \mathbb{P}[\sqrt{\varepsilon} W \in B(h,\delta)] \;\; \geq \;\; -I(h).$$

Sketch of proof. (Lower bound)

Let $G \neq \emptyset$ be an open set of $C([0,T])$, $h \in G$, and $\delta > 0$ s.t. $B(h,\delta) \subset G$.

We want to prove that

$$\liminf_{\varepsilon \to 0} \varepsilon \ln \mathbb{P}[\sqrt{\varepsilon} W \in B(h,\delta)] \;\; \geq \;\; -I(h).$$

For $h \notin H_0([0,T])$, this inequality is trivial since $I(h) = \infty$. Suppose now $h \in H_0([0,T])$, and consider the probability measure:

$$\frac{d\mathbb{Q}_h}{d\mathbb{P}} \;\; = \;\; \exp\Big(\int_0^T \frac{\dot{h}(t)}{\sqrt{\varepsilon}} dW_t - \frac{1}{2\varepsilon} \int_0^T |\dot{h}(t)|^2 dt \Big),$$

$\Longrightarrow$

$$W^h \;\; = \;\; W - \frac{h}{\sqrt{\varepsilon}} \qquad \text{is a Brownian motion under } \mathbb{Q}^h.$$

Sketch of proof (Ctd) (Lower bound)

$$
\begin{aligned}
\mathbb{P}[\sqrt{\varepsilon}W \in B(h,\delta)] &= \mathbb{P}[|W^h| < \frac{\delta}{\sqrt{\varepsilon}}] \\
&= \mathbb{E}^{\mathbb{Q}^h}\left[\exp\left(-\int_0^T \frac{\dot h(t)}{\sqrt{\varepsilon}} dW_t^h - \frac{1}{2\varepsilon}\int_0^T |\dot h(t)|^2 dt\right) 1_{|W^h| < \frac{\delta}{\sqrt{\varepsilon}}}\right] \\
(W^h\ \mathbb{Q}^h\text{-BM }) &= \mathbb{E}\left[\exp\left(-\int_0^T \frac{\dot h(t)}{\sqrt{\varepsilon}} dW_t - \frac{1}{2\varepsilon}\int_0^T |\dot h(t)|^2 dt\right) 1_{|W| < \frac{\delta}{\sqrt{\varepsilon}}}\right] \\
(W \sim -W) &= \mathbb{E}\left[\exp\left(+\int_0^T \frac{\dot h(t)}{\sqrt{\varepsilon}} dW_t - \frac{1}{2\varepsilon}\int_0^T |\dot h(t)|^2 dt\right) 1_{|W| < \frac{\delta}{\sqrt{\varepsilon}}}\right] \\
&= \mathbb{E}\left[\exp(-\frac{1}{2\varepsilon}\int_0^T |\dot h(t)|^2 dt) \cosh(\int_0^T \frac{\dot h(t)}{\sqrt{\varepsilon}} dW_t) 1_{|W| < \frac{\delta}{\sqrt{\varepsilon}}}\right] \\
&\geq \exp(-\frac{1}{2\varepsilon}\int_0^T |\dot h(t)|^2 dt)\ \mathbb{P}[|W| < \frac{\delta}{\sqrt{\varepsilon}}]
\end{aligned}
$$

$\Longrightarrow$

$$
\varepsilon \ln \mathbb{P}[\sqrt{\varepsilon}W \in B(h,\delta)] \geq -I(h) + \varepsilon \ln \mathbb{P}[|W| < \frac{\delta}{\sqrt{\varepsilon}}]
$$

$\square$

Corollary 1: diffusion with small noise parameter

$$dX_s^\varepsilon \;=\; b_\varepsilon(s, X_s^\varepsilon)ds + \sqrt{\varepsilon}\sigma(s, X_s^\varepsilon)dW_s, \quad t \le s \le T, \quad X_t^\varepsilon = x$$

with $\lim_{\varepsilon \to 0} b_\varepsilon = b$.

$\{X^{\varepsilon,x,t}, t \le s \le T\}$ **satisfies on** $C([t,T])$ **a LDP with rate function**:

$$I(h) \;=\; \begin{cases} \frac{1}{2}\int_t^T |\dot{h} - b(s,h)|^2_{(\sigma\sigma'(s,h))^{-1}} ds, & \text{if } h \in H([t,T]),\ h(t) = x, \\ \infty, & \text{otherwise} \end{cases}$$

This result can be derived from Schilder's theorem by contraction principle

Corollary 2: exit probability from a domain

Let $\{X^{\varepsilon,x,t}, t \leq s \leq T\}$ be the diffusion with small noise parameter, and consider the exit time from an open set Γ:

$$\tau(X^{\varepsilon,t,x}) \;:=\; \inf\Big\{s \geq t \;: X_s^{\varepsilon,t,x} \notin \Gamma\Big\},$$

Then,

$$\begin{aligned}\lim_{\varepsilon\to 0} \varepsilon \ln \mathbb{P}[\tau(X^{\varepsilon,t,x}) \leq T] &= -\inf\{I(h) : h \in H([t,T]), h(t) = x, \tau(h) \leq T\} \\ &=: -V_0(t,x).\end{aligned}$$

$\rightarrow$ $\sqrt{2V_0(t,x)}$ can be interpreted as a distance between x and $\partial\Gamma$ in the Riemannian metric defined by $(\sigma\sigma')^{-1}$.

$\rightarrow$ Sharp large deviations and asymptotic expansions by removing the log-estimate (see Baldi, Fleming-James):

$$\mathbb{P}[\tau(X^{\varepsilon,t,x}) \leq T] \;=\; e^{-V_0(t,x)/\varepsilon}(w(t,x) + \varepsilon w_1(t,x) + \ldots),$$

Sketch of proof by stochastic control method (Fleming)

We consider the exit probability large deviations problem:

$$v_\varepsilon(t,x) \;=\; \mathbb{P}[\tau(X^{\varepsilon,t,x}) \leq T], \quad (t,x) \in [0,T] \times \mathbb{R}^d.$$

It is well-known that the function v_ε satisfies the linear PDE

$$\frac{\partial v_\varepsilon}{\partial t} + b_\varepsilon(t,x).D_x v_\varepsilon + \frac{\varepsilon}{2}\mathrm{tr}(\sigma\sigma'(t,x)D_x^2 v_\varepsilon) \;=\; 0, \quad (t,x) \in [0,T) \times \Gamma$$

together with the boundary conditions

$$\begin{aligned} v_\varepsilon(t,x) &= 1, \quad (t,x) \in [0,T) \times \partial\Gamma \\ v_\varepsilon(T,x) &= 0, \quad x \in \Gamma. \end{aligned}$$

We make the **logarithm transformation**:

$$V_\varepsilon \;=\; -\varepsilon \ln v_\varepsilon.$$

Then, V^ε satisfies the nonlinear PDE

$$\begin{aligned}
-\frac{\partial V_\varepsilon}{\partial t} - b_\varepsilon(t,x).D_x V_\varepsilon - \frac{\varepsilon}{2}\mathrm{tr}(\sigma\sigma'(t,x)D_x^2 V_\varepsilon) & \\
+ \frac{1}{2}(D_x V_\varepsilon)'\sigma\sigma'(t,x)D_x V_\varepsilon \;&=\; 0, \quad (t,x) \in [0,T) \times \Gamma,
\end{aligned}$$

together with boundary data:

$$\begin{aligned}
V_\varepsilon(t,x) \;&=\; 0, \quad (t,x) \in [0,T) \times \partial\Gamma \\
V_\varepsilon(T,x) \;&=\; \infty, \quad x \in \Gamma.
\end{aligned}$$

At the limit $\varepsilon \to 0$, the above PDE becomes a first-order PDE:

$$-\frac{\partial V_0}{\partial t} - b(t,x).D_x V_0 + \frac{1}{2}(D_x V_0)'\sigma\sigma'(t,x)D_x V_0 \;=\; 0, \quad (t,x) \in [0,T) \times \Gamma,$$

with the same boundary data:

$$\begin{aligned} V_0(t,x) &= 0, \quad (t,x) \in [0,T) \times \partial\Gamma \\ V_0(T,x) &= \infty, \quad x \in \Gamma. \end{aligned}$$

By PDE methods and viscosity solutions, one can prove that $V^\varepsilon \to V_0$, solution to the above PDE.

This PDE for V_0 can be rewritten as an **Hamilton-Jacobi equation**:

Consider the Hamiltonian function

$$\mathcal{H}(t,x,p) \;=\; -b(t,x).p + \frac{1}{2}p'\sigma\sigma'(t,x)p, \quad (t,x,p) \in [0,T] \times \Gamma \times \mathbb{R}^d,$$

which is quadratic and in particular convex in p, so that:

$$-\frac{\partial V_0}{\partial t} + \mathcal{H}(t,x,D_x V_0) \;=\; 0.$$

This PDF for V_0 can be rewritten as an **Hamilton-Jacobi equation**:

Consider the Hamiltonian function

$$\mathcal{H}(t,x,p) \;=\; -b(t,x).p + \frac{1}{2}p'\sigma\sigma'(t,x)p, \quad (t,x,p) \in [0,T]\times\Gamma\times\mathbb{R}^d,$$

which is quadratic and in particular convex in p, so that:

$$-\frac{\partial V_0}{\partial t} + \mathcal{H}(t,x,D_xV_0) \;=\; 0.$$

Then, using the Legendre transform, we may rewrite

$$\mathcal{H}(t,x,p) \;=\; \sup_{q\in\mathbb{R}^d}\Big[-q.p - \mathcal{H}^*(t,x,q)\Big] \;=\; -\inf_{q\in\mathbb{R}^d}\Big[q.p + \mathcal{H}^*(t,x,q)\Big],$$

where

$$\begin{aligned}\mathcal{H}^*(t,x,q) &= \sup_{p\in\mathbb{R}^d}\Big[-p.q - \mathcal{H}(t,x,p)\Big] \\ &= \frac{1}{2}|q-b(t,x))|^2_{(\sigma\sigma'(t,x))^{-1}}, \;\; (t,x,q)\in[0,T]\times\Gamma\times\mathbb{R}^d.\end{aligned}$$

Hence, the PDE for V_0 is rewritten as an **Hamilton-Jacobi equation**:

$$\frac{\partial V_0}{\partial t} + \inf_{q \in \mathbb{R}^d} \Big[q.D_x V_0 + \mathcal{H}^*(t,x,q) \Big] \;=\; 0, \quad (t,x) \in [0,T) \times \Gamma,$$

which, together with the boundary data, is associated to the value function for the following calculus of variations problem:

$$\begin{aligned} V_0(t,x) &= \inf_{h \in H_x([t,T])} \int_t^T \mathcal{H}^*(s, h(s), \dot{h}(s)) ds, \\ &= \inf_{h \in H_x([t,T])} \int_t^T \frac{1}{2} |\dot{h} - b(s,h)|^2_{(\sigma\sigma'(s,h))^{-1}} du \end{aligned}$$

where

$$H_x([t,T]) \;=\; \{h \in H([t,T]) : \; h(t) = x \;\text{ and }\; \tau(h) \leq T\}.$$

Corollary 3: Diffusion processes densities in small time

Let X_t be the diffusion:

$$dX_t = \sigma(X_t)dW_t,$$

and denote $p(t,x,y)$ the transition probability density i.e.:

$$p(t,x,y)dy = \mathbb{P}[X_t \in dy | X_0 = x].$$

also called *Green function*, and satisfying the (backward) Kolmogorov equation: for fixed y,

$$\frac{\partial p(t,x,y)}{\partial t} = \frac{1}{2}\mathrm{tr}(\sigma\sigma'(x)D_x^2 p(t,x,y)),$$

and the (forward) Kolmogorov equation: for fixed x,

$$\frac{\partial p(t,x,y)}{\partial t} = \frac{1}{2}\mathrm{tr}(D_y^2(\sigma\sigma'(x)p(t,x,y))).$$

Varadhan's result gives a large deviation estimate of p for small time asymptotics:

$$
\begin{aligned}
\lim_{t \to 0} t \ln p(t,x,y) &= -\inf \left\{ \frac{1}{2} \int_0^1 |\dot{h}|^2_{(\sigma\sigma'(h))^{-1}} : h \in H([0,1]), h(0) = x, h(1) = y \right\} \\
&= -\frac{1}{2} d^2(x,y).
\end{aligned}
$$

$d(x,y)$ is the distance between x and y in the Riemannian metric associated with the inverse of the diffusion coefficient.

• Varadhan's result can be derived from Freidlin-Wentzell theory by time scaling.

• Improvement by Molchanov, Kiefer, Kanai, who derived sharp large deviations estimates and asymptotic expansions of $p(t,x,y)$ as t goes to zero (**Heat kernel expansion**)

• The practical use of these sample path large deviations results require the computation of the distance for the Riemannian metric associated to the inverse of the diffusion matrix: $(\sigma\sigma')^{-1}$

▶ Solve the **problem of calculus of variations** (e.g. by Euler-Lagrange equation) defining the Riemannian distance, and find the associated geodesics (critical path attaining the infimum).

▶ Solve the **Eikonal equation** (in geometric optics) satisfied by the Riemannian distance $d(x, y)$:

$$\begin{aligned} |\nabla_x d|_{\sigma\sigma'(x)} &= 1 \\ d(x, x) &= 0. \end{aligned}$$

References:

• Dembo A. and O. Zeitouni (1998): *Large deviations techniques and applications*, 2nd ed, Springer.

• Deuschel J.D and D. Stroock (1989): *Large deviations*, Academic Press.

• Dupuis P. and R. Ellis (1997): *A weak convergence approach to the theory of large deviations*, Wiley Series in Prob. and Stat.

• Fleming W. (1992): "Stochastic control and large deviations", Lect. Notes in Comp. Science, **653**, 291-300,

• Varadhan S. (2008): "Large deviations and applications", *Annals of Probability*, **36**, 397-419.

Lecture II. Large deviations in option pricing

1. Optimal importance sampling via large deviations approximation

2. Asymptotics in stochastic volatility models

1. Optimal importance sampling via large deviations approximation

• Option pricing problem: computation of

$$I_g = \mathbb{E}\Big[g(S_t, 0 \leq t \leq T)\Big],$$

Standard Monte-Carlo approximation:

$$I_g^N = \frac{1}{N}\sum_{i=1}^{N} g(S^i).$$

Consistency of this estimator by law of large numbers

Error approximation (CLT) measured by the variance

▶ **Importance sampling**: variance reduction method by changing probability measure from which paths are generated

Importance sampling for diffusions: Girsanov's theorem

• Consider a diffusion in $\mathbb{R}^d$ on $(\Omega, \mathcal{F}, \mathbb{P})$:

$$dX_s \;=\; b(X_s)ds + \Sigma(X_s)dW_s,$$

and define the (option price) function:

$$v(t,x) \;=\; \mathbb{E}\Big[g(X_s^{t,x}, t \le s \le T)\Big], \quad (t,x) \in [0,T] \times \mathbb{R}^d.$$

• Introduce a probability measure $\mathbb{Q} \sim \mathbb{P}$ by its Radon-Nikodym density:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \;=\; M_T \;=\; \exp\Big(- \int_0^T \phi'_u dW_u - \frac{1}{2}\int_0^T |\phi_t|^2 dt \Big),$$

for some $\mathbb{R}^d$-valued adapted process $\phi = (\phi_t)_{0 \le t \le T}$ s.t. $\mathbb{E}[M_T] = 1$.

• By Girsanov's theorem, the dynamics of X under $\mathbb{Q}$ is:

$$dX_s \;=\; \Big(b(X_s) - \Sigma(X_s)\phi_s\Big)ds + \Sigma(X_s)d\hat{W}_s,$$

where $\hat{W}$ is a brownian motion under $\mathbb{Q}$, and from Bayes formula, the (option price) function is:

$$v(t,x) \;=\; \mathbb{E}^{\mathbb{Q}}\Big[g(X_s^{t,x}, t \leq s \leq T)L_T\Big],$$

where L is the $\mathbb{Q}$-martingale density of $\mathbb{P}$ w.r.t. $\mathbb{Q}$:

$$L_t \;=\; \frac{1}{M_t} \;=\; \exp\Big(\int_0^t \phi_u' d\hat{W}_u - \frac{1}{2}\int_0^t |\phi_u|^2 du\Big), \quad 0 \leq t \leq T.$$

• By Girsanov's theorem, the dynamics of X under $\mathbb{Q}$ is:

$$dX_s = \Big(b(X_s) - \Sigma(X_s)\phi_s\Big)ds + \Sigma(X_s)d\hat{W}_s,$$

where $\hat{W}$ is a brownian motion under $\mathbb{Q}$, and from Bayes formula, the (option price) function is:

$$v(t,x) = \mathbb{E}^{\mathbb{Q}}\Big[g(X_s^{t,x}, t \leq s \leq T)L_T\Big],$$

where L is the $\mathbb{Q}$-martingale density of $\mathbb{P}$ w.r.t. $\mathbb{Q}$:

$$L_t = \frac{1}{M_t} = \exp\Big(\int_0^t \phi_u' d\hat{W}_u - \frac{1}{2}\int_0^t |\phi_u|^2 du\Big), \quad 0 \leq t \leq T.$$

▶ Alternative Monte-Carlo estimator for $v(t,x)$:

$$I_{g,\phi}^N(t,x) = \frac{1}{N}\sum_{i=1}^N g(X^{i,t,x})L_T^i,$$

by simulation of X and L under $\mathbb{Q}$.

▶ **Variance reduction technique**: choice of ϕ inducing a **smaller variance for** $I_{g,\phi}^N$.

▶ Two approaches for the construction of such ϕ (called **accelerator**), both relying on asymptotic results from large deviations:

A. Approximation of the option price via Freidlin-Wentzell results $\rightarrow$ stochastic accelerator ϕ

• Fournié E., Lasry J.M. and P.L. Lions (1997): "Some nonlinear methods to study far-from-the-money contingent claims", *Numerical Methods in Finance*, L.C.G. Rogers et D. Talay, eds, Cambridge University Press.

• Fournié E., Lasry J.M. and N. Touzi (1997): "Monte-Carlo methods for stochastic volatility models", *Numerical Methods in Finance*, L.C.G. Rogers et D. Talay, eds, Cambridge University Press.

B. Optimal deterministic accelerator ϕ via Laplace principle

• Glasserman P., Heidelberger P. and P. Shahabuddin (1999): "Asymptotically optimal importance sampling and stratification for pricing path-dependent options", *Mathematical finance*, **9**, 117-152.

• Guasoni P. and S. Robertson (2008): "Optimal importance sampling with explicit formulas in continuous-time", *Finance and Stochastics*, **12**, 1-19.

• Robertson S. (2010): "Sample path large deviations and optimal importance sampling for stochastic volatility models", *Stochastic Processes and their Applications*, **120**, 66-83.

A. Option pricing approximation via Freidlin-Wentzell theory

Suppose that the payoff g depends only on the terminal value X_T, and apply Itô's formula to the $\mathbb{Q}$-martingale $v(s, X_s^{t,x})L_s$ between $s = t$ and $s = T$:

$$g(X_T^{t,x})L_T \;=\; v(t,x)L_t + \int_t^T L_s\Big(D_x v(s, X_s^{t,x})'\Sigma(X_s^{t,x}) + v(x, X_s^{t,x})\phi_s'\Big)d\widehat{W}_s.$$

A. Option pricing approximation via Freidlin-Wentzell theory

Suppose that the payoff g depends only on the terminal value X_T, and apply Itô's formula to the $\mathbb{Q}$-martingale $v(s, X_s^{t,x})L_s$ between $s = t$ and $s = T$:

$$g(X_T^{t,x})L_T \;=\; v(t,x)L_t + \int_t^T L_s\Big(D_x v(s, X_s^{t,x})'\Sigma(X_s^{t,x}) + v(x, X_s^{t,x})\phi_s'\Big)d\widehat{W}_s.$$

Hence, the variance of $I_{g,\phi}^N(t,x)$ is given by

$$Var_{\mathbb{Q}}(I_{g,\phi}^N(t,x)) \;=\; \frac{1}{N}\mathbb{E}^{\mathbb{Q}}\Big[\int_t^T L_s^2\Big|D_x v(s, X_s^{t,x})'\Sigma(X_s^{t,x}) + v(x, X_s^{t,x})\phi_s'\Big|^2 ds\Big].$$

If the function v **were known**, then one could **vanish the variance** by choosing an accelerator:

$$\phi_s \;=\; \phi_s^* \;=\; -\frac{1}{v(s, X_s^{t,x})}\Sigma'(X_s^{t,x})D_x v(s, X_s^{t,x}), \;\; t \le s \le T.$$

▶ This suggests to use an accelerator ϕ from the above formula with an approximation of the function v.

▶ We may then reasonably hope to reduce the variance, and also to use such a method for more general payoff functions, possibly path-dependent.

▶ We shall use a large deviations approximation for the function v.

Basic idea:

Many derivatives contracts are designed to offer a payout in some exercice domain, and otherwise expire with no value.

Then, a large proportion of simulated paths may end up out of the exercice domain (for example deep out the money option), giving no contribution to the Monte-Carlo estimator, but increasing the variance

▶ However, by considering the large deviations of the process of interest around the deterministic system, then the proportion of simulated paths, which end up in the exercice domain, is increased significantly, reducing therefore the variance.

▶ Freidlin-Wentzell theory: asymptotics in small diffusion term, or equivalently by time-scaling, Varadhan theory: asymptotics in small time.

Illustration

Within a stochastic volatility model:

$$dX_t = d\begin{pmatrix} S_t \\ Y_t \end{pmatrix} = \begin{pmatrix} 0 \\ \eta(Y_t) \end{pmatrix} dt + \begin{pmatrix} \sigma(Y_t)S_t & 0 \\ \rho\gamma(Y_t) & \sqrt{1-\rho^2}\gamma(Y_t) \end{pmatrix} \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix}$$

consider an up-and-in bond of price:

$$\begin{aligned} v(t,x) &= \mathbb{E}\Big[1_{\max_{t\le u\le T} S_u^{t,x} \ge K}\Big] \quad t \in [0,T],\ x = (s,y) \in (0,\infty)\times\mathbb{R}, \\ &= \mathbb{P}[\tau(X^{t,x}) \le T], \end{aligned}$$

where

$$\tau(X^{t,x}) = \inf\Big\{u \ge t \ : X_u^{t,x} \notin \Gamma\Big\}, \quad \Gamma = (0,K)\times\mathbb{R}.$$

The payout event $\{\max_{t \leq u \leq T} S_u^{t,x} \geq K\} = \{\tau(X^{t,x}) \leq T\}$ is rare when $x = (s,y) \in \Gamma$, i.e. $s < K$ (out the money option) and the time to maturity $T-t$ is small.

▶ The large deviations asymptotics for the option price $v(t,x)$ in small time to maturity $T-t$ is provided by the **Freidlin-Wentzell and Varadhan theories**:

By **time scaling**, we have $v(t,x) = w_{T-t}(x)$ where

$$w_\varepsilon(x) \;=\; \mathbb{P}[\tau(X^{\varepsilon,x}) \leq 1],$$

and $X^{\varepsilon,x}$ is the solution to

$$dX_s^\varepsilon \;=\; \varepsilon b(X_s^\varepsilon)ds + \sqrt{\varepsilon}\Sigma(X_s^\varepsilon)dW_s, \quad X_0^\varepsilon \;=\; x.$$

and $\tau(X^{\varepsilon,x}) = \inf\left\{s \geq 0 : X_s^{\varepsilon,x} \notin \Gamma\right\}$.

From the sample paths large deviations result, we then have:

$$\lim_{t \nearrow T} (T-t) \ln v(t,x) \;=\; -V_0(x)$$

with

$$V_0(x) \;=\; \inf\left\{ \frac{1}{2}\int_0^1 |\dot{h}(t)|^2_{(\Sigma\Sigma'(h))^{-1}} dt : h \in H([0,1]), h(0) = x, \tau(h) \le 1 \right\}$$

From the sample paths large deviations result, we then have:

$$\lim_{t \nearrow T} (T-t) \ln v(t,x) \;=\; -V_0(x)$$

with

$$V_0(x) \;=\; \inf \left\{ \frac{1}{2} \int_0^1 |\dot{h}(t)|^2_{(\Sigma\Sigma'(h))^{-1}} dt : h \in H([0,1]), h(0) = x, \tau(h) \leq 1 \right\}$$

▶ Another expression of V_0 in terms of Riemanian distance associated to the metric $(\Sigma\Sigma)^{-1}$: $L_0(x) = \sqrt{2V_0(x)}$ is solution to the **eikonal equation**:

$$\begin{aligned} (D_x L_0)' \Sigma\Sigma'(x) D_x L_0 \;&=\; 1, \quad x \in \Gamma \\ L_0(x) \;&=\; 0, \quad x \in \partial\Gamma \end{aligned}$$

that may be numerically solved by finite difference methods. It is also represented as

$$L_0(x) \;=\; \inf_{z \in \partial\Gamma} L_0(x,z), \quad x \in \Gamma,$$

where $L_0(x,z)$ is the distance from x to z for the metric $\Sigma\Sigma^{-1}$:

$$L_0(x,z) \;=\; \inf \left\{ \int_0^1 |\dot{h}(t)|_{(\Sigma\Sigma'(h))^{-1}} dt : h \in H([0,1]), h(0) = x, h(1) = z \right\}.$$

This leads to the choice of an accelerator:

$$\phi(t,x) \;=\; \frac{L_0(x)}{T-t}\Sigma'(x)D_x L_0(x).$$

Such an accelerator ϕ may also be used for computing any option whose exercice domain looks similar to the up and in bond, e.g. deep out the money options.

B. Choice of accelerator via Laplace principle

This approach does not require knowledge (approximation) of the option price, and restricts to **deterministic** accelerators ϕ.

We identify the **option payoff with a nonnegative functional** $G(W)$ of the Brownian motion $W = (W_t)_{0 \le t \le T}$ on the set $C([0,T])$ of continuous functions on $[0,T]$, and we define $F = \ln G$ valued in $\mathbb{R} \cup \{-\infty\}$.

For example, in the case of the Black-Scholes model for the stock price S, with interest rate r and volatility σ, the payoff of an arithmetic Asian option is $(\frac{1}{T}\int_0^T S_t dt - K)_+$, corresponds to a functional:

$$G(w) \;=\; \Big(\frac{1}{T}\int_0^T S_0 \exp\big(\sigma w_t + (r - \sigma^2/2)t\big) - K\Big)_+.$$

Deterministic change of drifts via Girsanov's theorem

For any $h \in H_0([0,T])$ (Cameron-Martin space), we define the probability measure $\mathbb{Q}_h$:

$$\frac{d\mathbb{Q}_h}{d\mathbb{P}} = \exp\Big(\int_0^T \dot{h}(t)dW_t - \frac{1}{2}\int_0^T |\dot{h}(t)|^2 dt\Big),$$

$\rightarrow$ Monte-Carlo estimator of $\mathbb{E}[G(W)]$ by simulating under $\mathbb{Q}_h$ the payoff:

$$G(W)\frac{d\mathbb{P}}{d\mathbb{Q}_h}$$

Deterministic change of drifts via Girsanov's theorem

For any $h \in H_0([0,T])$ (Cameron-Martin space), we define the probability measure $\mathbb{Q}_h$:

$$\frac{d\mathbb{Q}_h}{d\mathbb{P}} = \exp\Big(\int_0^T \dot{h}(t)dW_t - \frac{1}{2}\int_0^T |\dot{h}(t)|^2 dt\Big),$$

$\rightarrow$ Monte-Carlo estimator of $\mathbb{E}[G(W)]$ by simulating under $\mathbb{Q}_h$ the payoff:

$$G(W)\frac{d\mathbb{P}}{d\mathbb{Q}_h}$$

$\rightarrow$ Objective: minimize over h the variance or equivalently the second moment of this estimator:

$$\begin{aligned} M^2(h) &= \mathbb{E}^{\mathbb{Q}_h}\Big[\Big(G(W)\frac{d\mathbb{P}}{d\mathbb{Q}_h}\Big)^2\Big] = \mathbb{E}\Big[G(W)^2\frac{d\mathbb{P}}{d\mathbb{Q}_h}\Big] \\ &= \mathbb{E}\Big[\exp\Big(2F(W) - \int_0^T \dot{h}(t)dW_t + \frac{1}{2}\int_0^T |\dot{h}(t)|^2 dt\Big)\Big]. \end{aligned}$$

Approximation method by small noise asymptotics

$$M_{\varepsilon}{}^{2}(h) \;=\; \mathbb{E}\Big[\exp\Big\{\frac{1}{\varepsilon}\Big(2F(\sqrt{\varepsilon}W) - \int_0^T \sqrt{\varepsilon}\dot{h}(t)dW_t + \frac{1}{2}\int_0^T |\dot{h}(t)|^2 dt\Big)\Big\}\Big].$$

Approximation method by small noise asymptotics

$$M_\varepsilon{}^2(h) \;=\; \mathbb{E}\Big[\exp\Big\{\frac{1}{\varepsilon}\Big(2F(\sqrt{\varepsilon}W) - \int_0^T \sqrt{\varepsilon}\dot h(t)dW_t + \frac{1}{2}\int_0^T |\dot h(t)|^2 dt\Big)\Big\}\Big].$$

▶ Schilder's theorem (LDP for $(X^\varepsilon = \sqrt{\varepsilon}W)_\varepsilon$) + Varadhan's integral formula yields:

$$\lim_{\varepsilon\to 0} \varepsilon \ln M_\varepsilon^2(\mu) \;=\; \sup_{z\in H_0([0,T])} \Big[2F(z) + \frac{1}{2}\int_0^T |\dot z(t) - \dot h(t)|^2 dt - \int_0^T |\dot z(t)|^2 dt\Big].$$

$\to$ We then say that $\widehat{h} \in H_0([0,T])$ is an **asymptotic optimal accelerator** if it is solution to the problem:

$$\inf_{h\in H_0([0,T])} \sup_{z\in H_0([0,T])} \Big[2F(z) + \frac{1}{2}\int_0^T |\dot z(t) - \dot h(t)|^2 dt - \int_0^T |\dot z(t)|^2 dt\Big].$$

Approximation method by small noise asymptotics

$$M_\varepsilon{}^2(h) \; = \; \mathbb{E}\Big[\exp\Big\{\frac{1}{\varepsilon}\Big(2F(\sqrt{\varepsilon}W) - \int_0^T \sqrt{\varepsilon}\dot{h}(t)dW_t + \frac{1}{2}\int_0^T |\dot{h}(t)|^2 dt\Big)\Big\}\Big].$$

▶ Schilder's theorem (LDP for $(X^\varepsilon = \sqrt{\varepsilon}W)_\varepsilon$) + Varadhan's integral formula yields:

$$\lim_{\varepsilon\to 0} \varepsilon \ln M_\varepsilon^2(\mu) \; = \; \sup_{z\in H_0([0,T])} \Big[2F(z) + \frac{1}{2}\int_0^T |\dot{z}(t) - \dot{h}(t)|^2 dt - \int_0^T |\dot{z}(t)|^2 dt\Big].$$

$\rightarrow$ We then say that $\widehat{h} \in H_0([0,T])$ is an **asymptotic optimal accelerator** if it is solution to the problem:

$$\inf_{h\in H_0([0,T])} \sup_{z\in H_0([0,T])} \Big[2F(z) + \frac{1}{2}\int_0^T |\dot{z}(t) - \dot{h}(t)|^2 dt - \int_0^T |\dot{z}(t)|^2 dt\Big].$$

▶ Swapping the order of optimization, this min-max problem is reduced to:

$$\sup_{h\in H_0([0,T])} \Big[2F(h) - \int_0^T |\dot{h}(t)|^2 dt\Big].$$

$\rightarrow$ Problem of calculus of variations: may be solved by Euler-Lagrange equation.

Example

Geometric Asian option: $G(w) = \left(S_0 e^{(r-\frac{\sigma^2}{2})T} e^{\frac{1}{T}\int_0^T \sigma w_t dt} - K\right)_+$

$\rightarrow$

$$\sup_{h \in H_0([0,T])} \left[2 \ln \left(e^{a \int_0^T h(t)dt} - c\right) - \int_0^T |\dot{h}(t)|^2 dt\right],$$

where $a = \sigma/T$, $c = \frac{K}{S_0} \exp\left(-(r - \frac{\sigma^2}{2})\frac{T}{2}\right)$.

- Euler-Lagrange equation:

$$\ddot{h} = -\alpha, \quad \text{with} \quad \alpha = a \frac{\exp\left(\int_0^T h(t)dt\right)}{\exp\left(\int_0^T h(t)dt\right) - c}, \tag{2}$$

$\rightarrow$

$$h(t) \;=\; -\frac{\alpha}{2}t^2 + \gamma t. \qquad (3)$$

The parameter γ is found by substituting (3) into α in (2), which yields

$$\gamma(\alpha) \;=\; \frac{aT^3\alpha - 6\ln\left(\frac{\alpha-a}{c\alpha}\right)}{3aT^2}.$$

Then, for this value of $\gamma = \gamma(\alpha)$, the problem (2) is solved by maximizing over $\alpha > a$. The optimal $\widehat{\alpha}$ is unique by strict concavity, and found implicitly via the first-order equation

$$a\widehat{\alpha}T^3 + 3\ln\left(\frac{\widehat{\alpha}-a}{c\widehat{\alpha}}\right) \;=\; 0.$$

This $\widehat{\alpha}$ satisfies $\gamma(\widehat{\alpha}) = \widehat{\alpha}T$, and thus the optimal drift is

$$\widehat{h}(t) \;=\; \frac{\widehat{\alpha}}{2}t^2 + \widehat{\alpha}Tt.$$

Numerical results (I)

Monte-Carlo simulations without applying variance reduction, and by applying the above importance sampling method. Parameter values are $T = 1$, $r = 3\%$, $\sigma = 30\%$, $S_0 = 100$, $K = 145$.

Number of simulations	without	IS
20000 Standard deviation/mean	13.9905	1.07
10000 Standard deviation/mean	13.7428	1.065

Numerical results (II)

Performance, in terms of variance ratios between the risk-neutral sample and the sample with the optimal accelerator for an Asian option in a Black-Scholes model.

Parameter values are $T = 1$, $r = 5\%$, $\sigma = 20\%$, $S_0 = 50$, and strikes are varying. 10^6 simulations.

Strike	Price	Variance ratios
50	304.0	7.59
60	28.00	26.5
70	1.063	310

2. Asymptotics in stochastic volatility (SV) models

• Demand from practioners for closed-form or quasi-closed form pricing of options and the need to calibrate models to data in a robust way

▶ Recent years, increasing interest for asymptotic and expansion methods in option pricing and **implied volatility** for SV models $\rightarrow$ considerable literature dealing with

- various asymptotics: small time or large time to maturity, extreme strike, fast and slow time scales.

▶ Some of these methods are related to large deviations, heat kernel expansion, or singular perturbation methods.

Short review of the literature

• Large deviations approach: small time asymptotics

Avellaneda, Boyer-Olsen, Busca, Friz (2003), Berestycki, Busca and Florent (2004): PDE and viscosity solutions methods

Series of paper by Forde and Jacquier (2009, 2010), thesis of A. Jacquier supervised by A. Mijatovic: probabilistic methods

• Heat kernel and geometric approach: small time asymptotics

Hagan, Lesniewski (2002), Henry-Labordère (2005), Bourgade, Croissant (2005), Lewis (2007), Gatheral, Laurence et al. (2009), ...

- Large time to maturity asymptotics: Tehranchi (2009), Forde and Jacquier (2009)

- Singular perturbation methods for fast-mean reverting asymptotics: see lectures of J.P. Fouque

- Extreme strike asymptotics: see lectures of R. Lee

• Large time to maturity asymptotics: Tehranchi (2009), Forde and Jacquier (2009)

• Singular perturbation methods for fast-mean reverting asymptotics: see lectures of J.P. Fouque

• Extreme strike asymptotics: see lectures of R. Lee

▶ In this lecture, we focus on **small-time asymptotics** in SV models by **large deviations** methods.

Stochastic Volatility Model

Log stock price $X_t = \ln S_t$ (and zero interest rate):

$$\begin{aligned} dX_t &= -\frac{1}{2}\sigma^2(Y_t)dt + \sigma(Y_t)dW_t^1 \\ dY_t &= \eta(Y_t)dt + \gamma(Y_t)dW_t^2, \end{aligned}$$

with $X_0 = x_0$, $Y_0 = y_0$, (W^1, W^2) Brownian motion (eventually correlated) on $(\Omega, \mathcal{F}, \mathbb{P})$.

▶ Compute approximation of call option price and implied volatility when time to maturity is small.

Large deviations for the log-stock price

$$\lim_{t \to 0} t \ln \mathbb{P}[X_t - x_0 \geq k] \; = \; -I(k), \quad k \geq 0,$$

• For general SV models, this LDP is derived from Freidlin-Wentzell theory and Varadhan sample path LD, and the rate function $I(k)$ is determined by the distance-minimizing geodesic from $(0, y_0)$ to the line $\{x = k\}$ on $\mathbb{R}^2$ for the Riemannian metric associated to the inverse of diffusion coefficient of (X, Y): $I(k) = \frac{1}{2} d(k)^2$.

$\to$ Differential geometry problem, but no explicit solution in general!

Large deviations for the log-stock price

$$\lim_{t \to 0} t \ln \mathbb{P}[X_t - x_0 \geq k] \;=\; -I(k), \quad k \geq 0,$$

• For general SV models, this LDP is derived from Freidlin-Wentzell theory and Varadhan sample path LD, and the rate function $I(k)$ is determined by the distance-minimizing geodesic from $(0, y_0)$ to the line $\{x = k\}$ on $\mathbb{R}^2$ for the Riemannian metric associated to the diffusion coefficient of (X, Y): $I(k) = \frac{1}{2}d(k)^2$.

$\rightarrow$ Differential geometry problem, but no explicit solution in general!

• For the Heston model and more generally for affine SV models, the LDP can be derived directly from explicit computation of the moment generating function and Ellis-Gartner theorem (see details later).

Corollary 1: **Pricing** for out-the-money call options of small maturity:

$$\lim_{t \to 0} t \ln \mathbb{E}[(S_t - K)_+] \;=\; -I(x) \;=\; \lim_{t \to 0} t \ln \mathbb{P}[S_t \geq K].$$

where

$$x \;=\; \ln(K/S_0) \;>\; 0$$

is the **log-moneyness**.

Similar result for out-of-the money put options.

Proof of lower bound.

For any $\varepsilon > 0$, we have

$$\mathbb{E}[(S_t - K)_+] \;\geq\; \mathbb{E}[(S_t - K)_+ 1_{S_t - K \geq \varepsilon}] \;\geq\; \varepsilon \mathbb{P}[S_t \geq K + \varepsilon].$$

Proof of lower bound.

For any $\varepsilon > 0$, we have

$$\mathbb{E}[(S_t - K)_+] \;\geq\; \mathbb{E}[(S_t - K)_+ 1_{S_t - K \geq \varepsilon}] \;\geq\; \varepsilon \mathbb{P}[S_t \geq K + \varepsilon].$$

By using the LDP for $X_t - x_0 = \ln(S_t/S_0)$, we then get

$$\begin{aligned} t \ln \mathbb{E}[(S_t - K)_+] &\geq t \ln \mathbb{P}[S_t \geq K + \varepsilon] \\ &= t \ln \mathbb{P}\Big[X_t - x_0 \geq \ln\Big(\frac{K+\varepsilon}{S_0}\Big)\Big] \\ &\to -I\Big(\ln\Big(\frac{K+\varepsilon}{S_0}\Big)\Big) \quad \text{as } t \to 0. \end{aligned}$$

By sending ε to zero, and from the continuity of I, we obtain the desired lower bound.

Proof of upper bound.

Apply Hölder inequality for any $p, q > 1$, $1/p + 1/q = 1$:

$$\begin{aligned}\mathbb{E}[(S_t - K)_+] = \mathbb{E}[(S_t - K)_+ 1_{S_t \geq K}] &\leq \left(\mathbb{E}[(S_t - K)_+^p]\right)^{\frac{1}{p}} \left(\mathbb{E}[1_{S_t \geq K}]\right)^{\frac{1}{q}} \\ &\leq \left(\mathbb{E}[S_t^p]\right)^{\frac{1}{p}} \left(\mathbb{P}[S_t \geq K]\right)^{\frac{1}{q}}.\end{aligned}$$

Proof of upper bound.

Apply Hölder inequality for any $p, q > 1$, $1/p + 1/q = 1$:

$$
\begin{aligned}
\mathbb{E}[(S_t - K)_+] = \mathbb{E}[(S_t - K)_+ 1_{S_t \geq K}] &\leq \left(\mathbb{E}[(S_t - K)_+^p]\right)^{\frac{1}{p}} \left(\mathbb{E}[1_{S_t \geq K}]\right)^{\frac{1}{q}} \\
&\leq \left(\mathbb{E}[S_t^p]\right)^{\frac{1}{p}} \left(\mathbb{P}[S_t \geq K]\right)^{\frac{1}{q}}.
\end{aligned}
$$

Taking ln and multiplying by t, this implies

$$
t \ln \mathbb{E}[(S_t - K)_+] \leq \frac{t}{p} \ln \mathbb{E}[S_t^p] + \left(1 - \frac{1}{p}\right) t \ln \mathbb{P}[S_t \geq K]
$$

Fom the LDP for $X_t - x_0 = \ln(S_t/S_0)$, it follows that

$$
\limsup_{t \to 0} t \ln \mathbb{E}[(S_t - K)_+] \leq -\left(1 - \frac{1}{p}\right) I(x).
$$

By sending p to infinity, we obtain the required upper-bound and so finally the desired result.

Implied volatility

Recall that the implied volatility $\sigma_t^{imp} = \sigma_t^{imp}(x)$ of a call option on S_t with strike $K = S_0 e^x$, and time to maturity t is determined from the implicit relation:

$$\begin{aligned}\mathbb{E}[(S_t - K)_+] &= C^{BS}(t, S_0, x, \sigma_t^{imp}) = \mathbb{E}[(S_t^{\sigma_t^{imp}} - K)_+] \\ &= S_0 \Phi(d_1(t, x, \sigma_t^{imp})) - S_0 e^x \Phi(d_2(t, x, \sigma_t^{imp})),\end{aligned}$$

where

$$d_1(t, x, \sigma) = \frac{-x + \frac{1}{2}\sigma^2 t}{\sigma\sqrt{t}}, \qquad d_2(t, x, \sigma) = d_1(t, x, \sigma) - \sigma\sqrt{t},$$

and $\Phi(d) = \int_{-\infty}^{d} \varphi(x) dx$ is the cdf of the normal law $\mathcal{N}(0, 1)$.

Corollary 2:

$$\lim_{t \to 0} \sigma_t^{imp}(x) = \frac{|x|}{\sqrt{2I(x)}}, \qquad x \neq 0.$$

Sketch of Proof.

• Standard estimate on the cdf $\Phi(d) = \int_{-\infty}^{d} \varphi(x)dx$ of the normal law:

$$\Phi(-d) \;=\; 1 - \Phi(d) \;\sim\; \frac{\varphi(d)}{d}, \quad \text{as } d \to \infty.$$

• Call option price $\mathbb{E}[(S_t - K)_+]$ goes to zero as t goes to zero (out-the-money) $\Longrightarrow$

$$\sigma_t^{imp}\sqrt{t} \;\to\; 0, \quad \text{and so} \quad d_1 = d_1(t, x, \sigma_t^{imp}),\; d_2 = d_2(t, x, \sigma_t^{imp}) \;\to\; -\infty,$$

as t goes to zero.

Proof of lower bound

From the large deviation estimate for the call option pricing, and the relation defining the implied volatility, we have for any $\varepsilon > 0$, and t small enough:

$$\exp\left(-\frac{I(x)+\varepsilon}{t}\right) \;\leq\; \mathbb{E}[(S_t - K)_+] \;\leq\; S_0\Phi(d_1) \;\sim\; \frac{S_0}{-d_1}\varphi(-d_1)$$

Proof of lower bound

From the large deviation estimate for the call option pricing, and the relation defining the implied volatility, we have for any $\varepsilon > 0$, and t small enough:

$$\exp\Big(-\frac{I(x)+\varepsilon}{t}\Big) \;\leq\; \mathbb{E}[(S_t - K)_+] \;\leq\; S_0\Phi(d_1) \;\sim\; \frac{S_0}{-d_1}\varphi(-d_1)$$

Now, since $-d_1 \sim \frac{x}{\sigma_t^{imp}\sqrt{t}}$ and $\varphi(d) = e^{-d^2/2}/\sqrt{2\pi}$, we get by taking ln, and sending t to zero:

$$-(I(x)+\varepsilon) \;\leq\; -\frac{x^2}{2\liminf_{t\to 0}|\sigma_t^{imp}|^2}.$$

We then send ε to zero, and get the lower bound:

$$\liminf_{t\to 0}|\sigma_t^{imp}|^2 \;\geq\; \frac{x^2}{2I(x)}.$$

Proof of upper bound

For all $\varepsilon > 0$, and t small enough,

$$\begin{aligned}
\exp\left(-\frac{I(x)-\varepsilon}{t}\right) &\geq \mathbb{E}[(S_t - K)_+] = \mathbb{E}[(S_t^{\sigma_t^{imp}} - K)_+] \\
&\geq \varepsilon \mathbb{P}[S_t^{\sigma_t^{imp}} \geq K + \varepsilon] = \varepsilon \Phi(d_{2,\varepsilon}) \sim \frac{\varepsilon}{-d_{2,\varepsilon}} \varphi(-d_{2,\varepsilon})
\end{aligned}$$

where

$$d_{2,\varepsilon} = -\frac{\ln\left(\frac{K+\varepsilon}{S_0}\right) + \frac{1}{2}|\sigma_t^{imp}|^2 t}{\sigma_t^{imp}\sqrt{t}} \sim -\frac{\ln\left(\frac{K+\varepsilon}{S_0}\right)}{\sigma_t^{imp}\sqrt{t}} \to -\infty,$$

as t goes to zero.

Proof of upper bound

For all $\varepsilon > 0$, and t small enough,

$$\begin{aligned}
\exp\left(-\frac{I(x)-\varepsilon}{t}\right) &\geq \mathbb{E}[(S_t - K)_+] = \mathbb{E}[(S_t^{\sigma_t^{imp}} - K)_+] \\
&\geq \varepsilon \mathbb{P}[S_t^{\sigma_t^{imp}} \geq K + \varepsilon] = \varepsilon \Phi(d_{2,\varepsilon}) \sim \frac{\varepsilon}{-d_{2,\varepsilon}} \varphi(-d_{2,\varepsilon})
\end{aligned}$$

where

$$d_{2,\varepsilon} = -\frac{\ln\left(\frac{K+\varepsilon}{S_0}\right) + \frac{1}{2}|\sigma_t^{imp}|^2 t}{\sigma_t^{imp}\sqrt{t}} \sim -\frac{\ln\left(\frac{K+\varepsilon}{S_0}\right)}{\sigma_t^{imp}\sqrt{t}} \to -\infty,$$

as t goes to zero. Taking ln, sending t to zero, and then ε to zero, we get the upper bound:

$$\limsup_{t \to 0} |\sigma_t^{imp}|^2 \leq \frac{x^2}{2I(x)}.$$

Heston model (CIR process for Y):

$$\begin{aligned} dX_t &= -\frac{1}{2} Y_t dt + \sqrt{Y_t}(\sqrt{1-\rho^2} dW_t^1 + \rho dW_t^2) \\ dY_t &= \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t} dW_t^2, \end{aligned}$$

with $X_0 = x_0 \in \mathbb{R}$, $Y_0 = y_0 > 0$, $\rho \in (-1, 1)$, κ, θ, $\sigma > 0$ and $2\kappa\theta > \sigma^2$.

▶ In this case, the LDP for $X_t - x_0$ can be derived directly from explicit calculation of the moment generating function (literature on affine processes: Filipovic, Teichmann, Keller-Ressel, Andersen and Piterbarg, etc ...), and Ellis-Gärtner-theorem.

Logarithm moment generating function

$$\Gamma_t(p) \;\; := \;\; \ln \mathbb{E}\Big[\exp\big(p(X_t - x_0)\big)\Big], \quad p \in \mathbb{R}.$$

Logarithm moment generating function

$$\Gamma_t(p) \;:=\; \ln \mathbb{E}\Big[\exp\big(p(X_t - x_0)\big)\Big], \quad p \in \mathbb{R}.$$

We rewrite as:

$$\begin{aligned}
\Gamma_t(p) &= \ln \mathbb{E}\Big[\exp\Big(-\frac{p}{2}\int_0^t Y_s ds + p\rho\int_0^t \sqrt{Y_s}dW_s^2 + p\sqrt{1-\rho^2}\int_0^t \sqrt{Y_s}dW_s^1\Big)\Big] \\
&= \ln \mathbb{E}\Big\{\exp\Big(-\frac{p}{2}\int_0^t Y_s ds + p\rho\int_0^t \sqrt{Y_s}dW_s^2\Big) \\
&\qquad \mathbb{E}\Big[\exp\Big(p\sqrt{1-\rho^2}\int_0^t \sqrt{Y_s}dW_s^1\Big)\Big|(W_s^2)_{s\le t}\Big]\Big\} \\
&= \ln \mathbb{E}\Big[\exp\Big(-\frac{p}{2}\int_0^t Y_s ds + p\rho\int_0^t \sqrt{Y_s}dW_s^2 + \frac{p^2(1-\rho^2)}{2}\int_0^t Y_s ds\Big)\Big] \\
&= \ln \mathbb{E}\Big[\exp\Big(p\rho\int_0^t \sqrt{Y_s}dW_s^2 - \frac{p^2\rho^2}{2}\int_0^t Y_s ds\Big)\exp\Big(\frac{p(p-1)}{2}\int_0^t Y_s ds\Big)\Big],
\end{aligned}$$

where we used the law of iterated conditional expectation in the second equality, and the fact that Y_t is measurable with respect to W^2.

By Girsanov's theorem, we then get

$$\Gamma_t(p) \;=\; \ln \mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(\frac{p(p-1)}{2}\int_0^t Y_s ds\Big)\Big],$$

where under $\mathbb{Q}$, the process Y satisfies the sde

$$dY_t \;=\; (\kappa\theta - (\kappa - \rho\sigma p)Y_t)dt + \sigma\sqrt{Y_t}dW_t^{2,\mathbb{Q}},$$

with $W^{2,\mathbb{Q}}$ a Brownian motion.

$\rightarrow$ exponential of functionals of CIR process

By Girsanov's theorem, we then get

$$\Gamma_t(p) \;=\; \ln \mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(\frac{p(p-1)}{2}\int_0^t Y_s ds\Big)\Big],$$

where under $\mathbb{Q}$, the process Y satisfies the sde

$$dY_t \;=\; (\kappa\theta - (\kappa - \rho\sigma p)Y_t)dt + \sigma\sqrt{Y_t}dW_t^{2,\mathbb{Q}},$$

with $W^{2,\mathbb{Q}}$ a Brownian motion.

$\rightarrow$ exponential of functionals of CIR process

$\rightarrow$ $\Gamma_t(p) = \phi(t,p) + y_0\psi(t,p)$, with $\phi(.,p)$, $\psi(.,p)$ solutions to the **Riccati system**:

$$\begin{aligned}
\frac{\partial\psi}{\partial t} &= \frac{p(p-1)}{2} - (\kappa-\rho\sigma p)\psi + \frac{\sigma^2}{2}\psi^2, & \psi(0,p) &= 0\\
\frac{\partial\phi}{\partial t} &= \kappa\theta\psi, & \phi(0,p) &= 0.
\end{aligned}$$

The Riccati equation is solved under the condition:

$$\delta \;=\; \delta(p) \;:=\; (\kappa - \rho\sigma p)^2 - \sigma^2 p(p-1) \;\geq\; 0$$

and the solution is given by:

$$\psi(t,p) \;=\; p(p-1)\frac{\sinh\left(\frac{\sqrt{\delta}}{2}t\right)}{(\kappa-\rho\sigma p)\sinh\left(\frac{\sqrt{\delta}}{2}t\right)+\sqrt{\delta}\cosh\left(\frac{\sqrt{\delta}}{2}t\right)},$$

$$\phi(t,p) \;=\; \frac{\kappa\theta}{\sigma^2}\Bigg[(\kappa-\rho\sigma p-\sqrt{\delta})t + 2\ln\left(\frac{\sqrt{\delta}e^{\frac{\sqrt{\delta}}{2}t}}{(\kappa-\rho\sigma p)\sinh\left(\frac{\sqrt{\delta}}{2}t\right)+\sqrt{\delta}\cosh\left(\frac{\sqrt{\delta}}{2}t\right)}\right)\Bigg],$$

which are defined for $t \in [0, T^*)$ until the **moment explosion time**

$$T^* \;=\; T^*(p) \;=\; \begin{cases} \infty, & \text{if } \kappa - \rho\sigma p \geq 0, \\ \frac{1}{\sqrt{\delta}}\ln\left(\frac{\kappa-\rho\sigma p-\sqrt{\delta}}{\kappa-\rho\sigma p+\sqrt{\delta}}\right), & \text{if } \kappa - \rho\sigma p < 0. \end{cases}$$

In the case $\delta(p) < 0$, the functions ϕ and ψ are extended by analytic continuation by substituting $\sqrt{\delta}$ by $i\sqrt{-\delta}$, which yields:

$$
\begin{aligned}
\psi(t,p) &= p(p-1)\frac{\sin\left(\frac{\sqrt{-\delta}}{2}t\right)}{(\kappa-\rho\sigma p)\sin\left(\frac{\sqrt{-\delta}}{2}t\right)+\sqrt{-\delta}\cos\left(\frac{\sqrt{-\delta}}{2}t\right)},\\
\phi(t,p) &= \frac{\kappa\theta}{\sigma^2}\Big[(\kappa-\rho\sigma p-i\sqrt{-\delta})t\\
&\qquad + 2\ln\left(\frac{\sqrt{-\delta}e^{\frac{i\sqrt{-\delta}}{2}t}}{(\kappa-\rho\sigma p)\sin\left(\frac{\sqrt{-\delta}}{2}t\right)+\sqrt{-\delta}\cos\left(\frac{\sqrt{-\delta}}{2}t\right)}\right)\Big],
\end{aligned}
$$

and this analytic continuation holds for $t \in [0, T^*)$ until the **moment explosion time**

$$
T^* = T^*(p) = \frac{2}{\sqrt{-\delta}}\Big[\pi 1_{\kappa-\rho\sigma p>0} + \arctan\left(\frac{\sqrt{-\delta}}{\rho\sigma p-\kappa}\right)\Big].
$$

Recalling that a moment generating function is analytic in the interior of its convex domain (when its is not empty), we deduce that Γ_t is explicitly given by

$$\Gamma_t(p) \;=\; \begin{cases} \phi(t,p) + y_0\psi(t,p), & t < T^*(p),\ p \in \mathbb{R} \\ \infty, & t \geq T^*(p),\ p \in \mathbb{R}. \end{cases}$$

Recalling that a moment generating function is analytic in the interior of its convex domain (when its is not empty), we deduce that Γ_t is explicitly given by

$$\Gamma_t(p) \;=\; \begin{cases} \phi(t,p) + y_0\psi(t,p), & t < T^*(p),\ p \in \mathbb{R} \\ \infty, & t \geq T^*(p),\ p \in \mathbb{R}. \end{cases}$$

Now, in view of deriving a LDP for $X_t - x_0$ (when t goes to zero) by means of Ellis-Gärtner theorem, we need to determine the *limiting logarithm moment generating function*:

$$\Gamma(p) \;:=\; \lim_{t\to 0} t\Gamma_t(p/t).$$

$\rightarrow$ Substitute p by p/t and send t to zero in the above calculations.

$$\Gamma(p) \;=\; \begin{cases} \frac{p}{\sigma}\frac{y_0}{\sqrt{1-\rho^2}\cot\left(\frac{\sigma p\sqrt{1-\rho^2}}{2}\right)-\rho}, & \text{for } p \in (p_-, p_+) \\ \infty, & \text{otherwise.} \end{cases}$$

where $p_- < 0$ (resp. p_+) is defined by

$$p_- \text{ (resp. } p_+) \;=\; \begin{cases} \dfrac{2\arctan\left(\frac{\sqrt{1-\rho^2}}{\rho}\right)}{\sigma\sqrt{1-\rho^2}}, & \text{if } \rho < 0 \text{ (resp. } > 0) \\ -\frac{\pi}{\sigma}, & \text{if } \rho = 0 \\ \dfrac{2\arctan\left(\frac{\sqrt{1-\rho^2}}{\rho}\right) \pm 2\pi}{\sigma\sqrt{1-\rho^2}}, & \text{if } \rho > 0 \text{ (resp. } < 0) \end{cases}$$

$$\Gamma(p) \;=\; \begin{cases} \frac{p}{\sigma}\dfrac{y_0}{\sqrt{1-\rho^2}\cot\left(\frac{\sigma p\sqrt{1-\rho^2}}{2}\right)-\rho}, & \text{for } p \in (p_-, p_+) \\ \infty, & \text{otherwise.} \end{cases}$$

where $p_- < 0$ (resp. p_+) is defined by

$$p_- \text{ (resp. } p_+) \;=\; \begin{cases} \dfrac{2\arctan\left(\frac{\sqrt{1-\rho^2}}{\rho}\right)}{\sigma\sqrt{1-\rho^2}}, & \text{if } \rho < 0 \text{ (resp. } > 0) \\ -\frac{\pi}{\sigma}, & \text{if } \rho = 0 \\ \dfrac{2\arctan\left(\frac{\sqrt{1-\rho^2}}{\rho}\right) \pm 2\pi}{\sigma\sqrt{1-\rho^2}}, & \text{if } \rho > 0 \text{ (resp. } < 0) \end{cases}$$

▶ One checks that Γ is steep, so that by **Ellis-Gärtner theorem**, $X_t - x_0$ satisfies a LDP with rate function:

$$I(x) \;=\; \sup_{p\in(p_-,p_+)} [px - \Gamma(p)], \;\; x \in \mathbb{R}.$$

Level, slope and curvature of the small-time implied volatility at-the-money in the Heston model

By Taylor expansion of the small-time implied volatility formula:

$$\sigma_0^{imp}(x) \;=\; \frac{|x|}{\sqrt{2I(x)}},$$

around $x = 0$ (at-the-money), and explicit expressions of Γ and its Fenchel-Legendre transform $I(x)$, we obtain:

$$\sigma_0^{imp}(x) \;=\; \sqrt{y_0}\Big[1 + \frac{\rho\sigma}{4y_0}x + \frac{\sigma^2}{24y_0^2}\Big(1 - \frac{5}{2}\rho^2\Big)x^2 + O(x^3)\Big]$$

(Durrleman (2004))

Some variations and extensions

- Corrections terms for the small-time asymptotics in the Heston model: Forde, Jacquier, Lee (2010),

- Small time asymptotics for fast-mean reverting SV models: Feng, Forde, Fouque (2009)

- Affine stochastic volatility models with jumps: see forthcoming talk by A. Jacquier

Lecture III. Large deviations in risk management

1. Large portfolio losses in credit risk

2. Long term investment

- Asymptotic arbitrage and large deviations

- Beating a benchmark: a large deviations approach

1. Large portfolio losses in credit risk

• Basic problem in measuring portfolio credit risk: Determine the distribution of losses from default over a fixed horizon.

- Credit portfolios are often large: e.g. exposure to thousands of obligors

- Default probabilities of high-quality credits are small!

▶ Rare but significant large loss events.

▶ Computation of the small probabilities of large losses: relevant for calculation of VaR, and related risk measures.

Notations

n = number of obligors to which portfolio is exposed,

Y_k = default indicator (= 1 if default, 0 otherwise) for k-th obligor,

$p_k = p$ = marginal probability that k-th obligor defaults, i.e. $p_k = \mathbb{P}[Y_k = 1]$,

$c_k = 1$ = loss resulting from default of the k-th obligor,

$L_n = c_1 Y_1 + \ldots + c_n Y_n$ = total loss from defaults.

$\rightarrow$ Estimation of tail probabilities:

$$\mathbb{P}[L_n > \ell_n]$$

in the limiting regime at increasingly high loss thresholds ℓ_n, and rarity of large losses resulting from a large number n of obligors and multiple defaults.

Dependence modelling among obligors: Normal copula model

$$
\begin{aligned}
Y_k &= 1_{\{X_k > x_k\}}, \\
X_k &= \rho Z + \sqrt{1-\rho^2}\varepsilon_k, \quad k = 1, \ldots, n.
\end{aligned}
$$

where $Z \rightsquigarrow \mathcal{N}(0,1)$, ε_k are independent $\mathcal{N}(0,1)$ distribution, and Z is independent of ε_k, $k = 1, \ldots, n$.

Z: systematic risk factor, common to all obligors

ε_k: idiosyncratic risk factor associated with the k-th obligor.

$\rho \in [0,1)$: factor loading on the single factor Z.

Case of independent obligors: $\rho = 0$

The default indicators Y_k are i.i.d. $\rightsquigarrow \mathcal{B}(p)$

$\to L_n \rightsquigarrow \mathcal{B}(n,p)$, and $\frac{L_n}{n} \to p$.

$\to$ The loss event $\{L_n \geq l_n\}$ becomes rare (without being trivially impossible) if e.g. we let $\ell_n = nq$ with $q \in (p, 1)$.

Case of independent obligors: $\rho = 0$

The default indicators Y_k are i.i.d. $\rightsquigarrow \mathcal{B}(p)$

$\rightarrow$ $L_n \rightsquigarrow \mathcal{B}(n,p)$, and $\frac{L_n}{n} \rightarrow p$.

$\rightarrow$ The loss event $\{L_n \geq l_n\}$ becomes rare (without being trivially impossible) if e.g. we let $\ell_n = nq$ with $q \in (p, 1)$.

▶ **Cramer's theorem**: large deviation of loss probability

$$\begin{aligned} \lim_{n \rightarrow} \frac{1}{n} \ln \mathbb{P}[L_n \geq nq] &= -R(\mathcal{B}(q)|\mathcal{B}(p)) \\ &= -q \ln\left(\frac{q}{p}\right) - (1-q)\ln\left(\frac{1-q}{1-p}\right) < 0. \end{aligned}$$

Remark: Estimation of tail probability $\mathbb{P}[L_n > nq]$

By denoting $\Gamma(\theta) = \ln(1 - p + pe^\theta)$ the Log-Laplace of $\mathcal{B}(p)$, we have an IS (unbiased) estimator of $\mathbb{P}[L_n \geq nq]$ by taking the average of independent replications of

$$\exp(-\theta L_n + n\Gamma(\theta))1_{L_n \geq nq}$$

where L_n is sampled with a default probability $p(\theta) = \mathbb{P}_\theta[Y_k = 1] = pe^\theta/(1 - p + pe^\theta)$.

Moreover, this estimator is asymptotically optimal, as n goes to infinity, for the choice of parameter $\theta_q \geq 0$ attaining the argmax in $\theta q - \Gamma(\theta)$.

Case of dependent obligors: $\rho > 0$

Conditionally on the factor Z, the default indicators Y_k are i.i.d. with Bernoulli distribution of parameter:

$$
\begin{aligned}
p(Z) &= \mathbb{P}[Y_k = 1|Z] = \mathbb{P}[\rho Z + \sqrt{1-\rho^2}\varepsilon_k > -\Phi^{-1}(p)|Z] \\
&= \Phi\Big(\frac{\rho Z + \Phi^{-1}(p)}{\sqrt{1-\rho^2}}\Big),
\end{aligned}
$$

with Φ c.d.f. of $\mathcal{N}(0,1)$.

$\rightarrow$ $\frac{L_n}{n}$ converges in law to $p(Z)$ valued in $(0,1)$.

$\rightarrow$ The event $\{L_n \geq l_n\}$ becomes rare (without being trivially impossible) if

$$
l_n = nq_n, \quad \text{with } q_n < 1,\ q_n \nearrow 1 \text{ as } n \rightarrow \infty.
$$

We shall consider:

$$
1 - q_n = O(n^{-a}), \quad \text{with } 0 < a \leq 1.
$$

Theorem

$$\lim_{n\to\infty} \frac{1}{\ln n} \ln \mathbb{P}[L_n \geq nq_n] \;=\; -a\frac{1-\rho^2}{\rho^2}.$$

Comments:

• The loss probability decays like $n^{-\gamma}$, with $\gamma = a(1-\rho^2)/\rho^2$.

• The decay rate is determined by the effect of the dependence structure in the Gaussian copula model:

- When ρ is small (weak dependence between sources of credit risk), large losses occur very rarely, which is formalized by a high decay rate.

- In the opposite case, this decay rate is small when ρ tends to one, which means that large losses are most likely to result from systematic risk factors.

Sketch of proof of the upper-bound.

We introduce the conditional Log-Laplace of Y_k :

$$\begin{aligned} \Gamma(\theta, z) &= \ln \mathbb{E}\left[e^{\theta Y_k} | Z = z\right] \\ &= \ln(1 - p(z) + p(z)e^{\theta}). \end{aligned}$$

Sketch of proof of the upper-bound.

We introduce the conditional Log-Laplace of Y_k :

$$\begin{aligned}\Gamma(\theta, z) &= \ln \mathbb{E}\Big[e^{\theta Y_k} | Z = z] \\ &= \ln(1 - p(z) + p(z)e^{\theta}).\end{aligned}$$

Then, for any $\theta \geq 0$, we get by Chebichev's inequality,

$$\mathbb{P}[L_n \geq nq_n | Z] \;\leq\; \mathbb{E}\Big[e^{\theta(L_n - nq_n)} | Z\Big] \;=\; e^{-n(\theta q_n - \Gamma(\theta, Z))},$$

so that by taking supremum over θ and taking expectation:

$$\mathbb{P}[L_n \geq nq_n] \;\leq\; \mathbb{E}[e^{-n\Gamma^*(q_n, Z)}] \;=:\; \mathbb{E}[e^{F_n(Z)}]$$

where $F_n(z) = -n\Gamma^*(q, z)$, and

$$\begin{aligned}\Gamma^*(q, z) &= \sup_{\theta \geq 0}[\theta q - \Gamma(\theta, z)] \\ &= \begin{cases} 0, & \text{if } q \leq p(z) \\ q \ln\left(\frac{q}{p(z)}\right) + (1 - q)\ln\left(\frac{1-q}{1-p(z)}\right), & \text{if } p(z) < q \leq 1. \end{cases}\end{aligned}$$

Shift the factor mean of Z to reduce the variance of $e^{F_n(Z)}$: introduce the change of measure $\mathbb{P}_\mu$ under which $Z \rightsquigarrow \mathcal{N}(\mu, 1)$:

$$\begin{aligned}\mathbb{P}[L_n \geq nq_n] \;&\leq\; \mathbb{E}[e^{F_n(Z)}] \;=\; \mathbb{E}_\mu\Big[e^{F_n(Z)-\mu Z+\frac{1}{2}\mu^2}\Big] \\ &\leq\; \mathbb{E}_\mu\Big[e^{F_n(\mu)+(F_n'(\mu)-\mu)Z-\mu F_n'(\mu)+\frac{1}{2}\mu^2}\Big],\end{aligned}$$

by concavity of F_n.

Shift the factor mean of Z to reduce the variance of $e^{F_n(Z)}$: introduce the change of measure $\mathbb{P}_\mu$ under which $Z \rightsquigarrow \mathcal{N}(\mu, 1)$:

$$
\begin{aligned}
\mathbb{P}[L_n \geq nq_n] &\leq \mathbb{E}[e^{F_n(Z)}] = \mathbb{E}_\mu\Big[e^{F_n(Z)-\mu Z+\frac{1}{2}\mu^2}\Big] \\
&\leq \mathbb{E}_\mu\Big[e^{F_n(\mu)+(F_n'(\mu)-\mu)Z-\mu F_n'(\mu)+\frac{1}{2}\mu^2}\Big],
\end{aligned}
$$

by concavity of F_n. Choose $\mu = \mu_n$ solution to:

$$
\mu_n = \arg\max_{\mu\in\mathbb{R}}[F_n(\mu) - \frac{1}{2}\mu^2], \quad \text{i.e. } F_n'(\mu_n) = \mu_n,
$$

so that

$$
\mathbb{P}[L_n \geq nq_n] \leq e^{F_n(\mu_n)-\frac{1}{2}\mu_n^2}.
$$

- Rate of convergence of (μ_n):

  - $\mu_n \sim z_n$, where z_n is the solution to $p(z_n) = q_n$ from which F_n is constant equal to zero.

  - Since q_n converges to 1, this implies $z_n \to \infty$.

  - By writing that $O(n^{-a}) = 1 - q_n = 1 - p(z_n) = 1 - \Phi\left(\frac{\rho z_n + \Phi^{-1}(p)}{\sqrt{1-\rho^2}}\right)$, and the standard estimate $1 - \Phi(d) \sim \varphi(d)/d$, we obtain:

$$\lim_{n\to\infty} \frac{\mu_n^2}{\ln n} = \lim_{n\to\infty} \frac{z_n^2}{\ln n} = 2a\frac{1-\rho^2}{\rho^2}.$$

▶ Recalling the bound:

$$\mathbb{P}[L_n \geq nq_n] \;\; \leq \;\; e^{F_n(\mu_n)-\frac{1}{2}\mu_n^2},$$

we deduce the large deviation upper bound:

$$\limsup_{n\to\infty} \frac{1}{\ln n} \ln \mathbb{P}[L_n \geq nq_n] \;\; \leq \;\; -\frac{1}{2} \lim_{n\to} \frac{\mu_n^2}{\ln n} \;\; = \;\; -a\frac{1-\rho^2}{\rho^2}.$$

Remark: Estimation of tail probability $\mathbb{P}[L_n > nq_n]$

• Conditionally on Z, we have an IS (unbiased) estimator of $\mathbb{P}[L_n \geq nq_n|Z]$ with

$$\exp(-\theta_{q_n}(Z)L_n + n\Gamma(\theta_{q_n}(Z), Z))1_{L_n \geq nq_n},$$

where L_n is sampled with a default probability $p(\theta_{q_n}(Z), Z) = p(Z)e^{\theta_{q_n}(Z)}/(1 - p(Z) + p(Z)e^{\theta_{q_n}(Z)})$, and $\theta_{q_n}(Z) \geq 0$ attaining the argmax in $\theta q_n - \Gamma(\theta, Z)$.

Remark: Estimation of tail probability $\mathbb{P}[L_n > nq_n]$

• Conditionally on Z, we have an IS (unbiased) estimator of $\mathbb{P}[L_n \geq nq_n|Z]$ with

$$\exp(-\theta_{q_n}(Z)L_n + n\Gamma(\theta_{q_n}(Z), Z))1_{L_n \geq nq_n},$$

where L_n is sampled with a default probability $p(\theta_{q_n}(Z), Z) = p(Z)e^{\theta_{q_n}(Z)}/(1 - p(Z) + p(Z)e^{\theta_{q_n}(Z)})$, and $\theta_{q_n}(Z) \geq 0$ attaining the argmax in $\theta q_n - \Gamma(\theta, Z)$.

• We further apply IS to the factor $Z \sim \mathcal{N}(0,1)$ under $\mathbb{P}$, by shifting the factor mean to μ, and then considering the estimator

$$\exp(-\mu Z + \frac{1}{2}\mu^2)\exp(-\theta_{q_n}(Z)L_n + n\Gamma(\theta_{q_n}(Z), Z))1_{L_n \geq nq_n},$$

where Z is sampled from $\mathcal{N}(\mu, 1)$.

Moreover, this this estimator is asymptotically optimal, as n goes to infinity, for the choice of parameter $\mu = \mu_n$.

References

• Dembo A., Deuschel J.D. and D. Duffie (2004): "Large portfolio losses", *Finance and Stochastics*, 8, 3-16.

• Glasserman P. (2004): "Tail approximations for portfolio credit risk", *Journal of Derivatives*, 24-42.

• Glasserman P., Kang W. and P. Shahabuddin (2007): "Large Deviations of Multifactor Portfolio Credit Risk", *Mathematical Finance*, vol. 17, 345-379.

Other related works

• Portfolio loss process: Leijdekker, Mandjes, Spreij (2009)

• Rare event losses in CDO tranches: Sowers (2009)

2. Long term investment

• Optimal investment in a financial market when the time horizon T tends to infinity

▶ Exponential growth of the terminal wealth X_T for $T \to \infty$.

▶ Asymptotic arbitrage and large deviations

▶ Beating a benchmark over long run

Diffusion model for stock price

$$dS_t \;=\; \Sigma(S_t)\Big(dW_t + \lambda(S_t)dt\Big),$$

on $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, with W a standard Brownian motion.

λ is the **market price of risk**: the stock's rate of return per unit volatility.

We assume that the Doléans-Dade exponential process

$$Z_t \;=\; \exp\left(-\int_0^t \lambda(S_u)dW_u - \frac{1}{2}\int_0^t |\lambda(S_u)|^2 du\right), \quad t \geq 0,$$

is a martingale, so that from Girsanov's theorem, for each $T > 0$, the measure $\mathbb{Q}_T$ on $\mathcal{F}_T$ defined by

$$\frac{d\mathbb{Q}_T}{d\mathbb{P}} \;=\; Z_T,$$

is a probability measure equivalent to $\mathbb{P}$ on $(\Omega, \mathcal{F}_T)$ s.t. $(S_t)_{0\leq t\leq T}$ is a local martingale under $\mathbb{Q}_T$: **equivalent martingale measure**.

Question

• Which features of the above model imply exponential growth of a well chosen portfolio wealth, as time horizon goes to infinity?

▶ Some conditions have to be imposed on the market price of risk. Indeed, if the market price of risk vanishes, then the stock price is a local martingale, and one cannot systematically win by betting on a local martingale in an admissible way.

Question

• Which features of the above model imply exponential growth of a well chosen portfolio wealth, as time horizon goes to infinity?

▶ Some conditions have to be imposed on the market price of risk. Indeed, if the market price of risk vanishes, then the stock price is a local martingale, and one cannot systematically win by betting on a local martingale in an admissible way.

• We say that S has a **non-trivial market price of risk** if there is $c > 0$ s.t.

$$\lim_{T\to\infty} \mathbb{P}\Big[\frac{1}{T}\int_0^T |\lambda(S_t)|^2 dt < c\Big] \;=\; 0.$$

▶ This condition is trivially satisfied for the BS model with nonzero constant market price of risk $\lambda \neq 0$. Also satisfied for ergodic diffusion processes with invariant measure μ and if λ is not μ a.s. equal to zero, since

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T |\lambda(S_t)|^2 dt \;=\; \int |\lambda(x)|^2 \mu(dx).$$

▶ If S has a non-trivial market price of risk, then there is an **asymptotic arbitrage** in the following sense: there exists $\gamma > 0$, s.t. for each $\varepsilon > 0$, for T large enough ($T \geq T_\varepsilon$), one can find some (admissible) terminal wealth X_T, starting from zero initial capital s.t.

$$(i) \quad X_T \; \geq \; -e^{-\gamma T}, \qquad (ii) \quad \mathbb{P}[X_T \geq e^{\gamma T}] \; \geq \; 1 - \varepsilon. \tag{4}$$

Interpretation: one may achieve exponential growth of a portfolio X_T with probability close to 1 (equal to $1 - \varepsilon$) as T goes to infinity (for $T \geq T_\varepsilon$), with an exponentially decreasing maximal potential loss.

However, the relation between ε and T_ε is not clarified, and the terminal wealth X_T is not explicitly given.

• We say that the **market price of risk satisfies a large deviations estimate** if there are constants c_1, $c_2 > 0$ s.t.

$$\limsup_{T\to\infty} \frac{1}{T} \ln \left(\mathbb{P}\left[\frac{1}{T} \int_0^T |\lambda(S_t)|^2 dt \le c_1 \right] \right) \;<\; -c_2.$$

• We say that the **market price of risk satisfies a large deviations estimate** if there are constants c_1, $c_2 > 0$ s.t.

$$\limsup_{T\to\infty} \frac{1}{T} \ln \left(\mathbb{P}\Big[\frac{1}{T}\int_0^T |\lambda(S_t)|^2 dt \le c_1\Big]\right) \;<\; -c_2.$$

▶ This condition is trivially satisfied for the BS model with nonzero constant market price of risk $\lambda \neq 0$. More generally for ergodic processes by Donsker-Varadhan large deviations results.

▶ One may expect to strenghten asymptotic arbitrage result in (**??**) with an exponential decay in time for the probability of falling short of the exponential growth portfolio:

$$\mathbb{P}[X_T < e^{\gamma_1 T}] \;\sim\; e^{-\gamma_3 T}, \quad i.e. \;\; \mathbb{P}[X_T \geq e^{\gamma_1 T}] \;\sim\; 1 - e^{-\gamma_3 T}$$

Illustration with the BS model: constant market price of risk $\lambda \neq 0$.

Take $\gamma \in (0, \lambda^2/2)$, $0 < \gamma_1 < \gamma$, and set for all $T > 0$, $A_T = \{Z_T \geq e^{-\gamma T}\}$, $\alpha_T = e^{\gamma_1 T}\mathbb{Q}[A_T^c]/\mathbb{Q}[A_T]$. Then, the claim

$$X_T \;=\; e^{\gamma_1 T} 1_{A_T^c} - \alpha_T 1_{A_T}$$

is an admissible terminal wealth attainable from zero initial capital, and satisfies for any $0 < \gamma_2 < \gamma - \gamma_1$:

$$X_T \;\geq\; -e^{-\gamma_2 T}, \quad \text{for large } T,$$

$$\lim_{T \to \infty} \frac{1}{T} \ln \mathbb{P}[X_T < e^{\gamma_1 T}] \;=\; -\frac{1}{2}\Big(\frac{\lambda}{2} - \frac{\gamma}{\lambda}\Big)^2.$$

Constructive proof.

- Set $A_T := \{Z_T \geq e^{-\gamma T}\}$. Then $\mathbb{Q}[A_T^c] = \int_{\{Z_T \leq e^{-\gamma T}\}} Z_T d\mathbb{P} \leq e^{-\gamma T}$, and so:

$$\begin{aligned} \alpha_T \;:=\; e^{\gamma_1 T}\frac{\mathbb{Q}[A_T^c]}{\mathbb{Q}[A_T]} &\leq e^{-(\gamma-\gamma_1)T}\frac{1}{1-e^{-\gamma T}} \\ &\leq e^{-\gamma_2 T} \qquad \text{for large } T \quad \text{if } \gamma_2 < \gamma - \gamma_1. \end{aligned}$$

Constructive proof.

• Set $A_T := \{Z_T \geq e^{-\gamma T}\}$. Then $\mathbb{Q}[A_T^c] = \int_{\{Z_T \leq e^{-\gamma T}\}} Z_T d\mathbb{P} \leq e^{-\gamma T}$, and so:

$$\begin{aligned} \alpha_T \;:=\; e^{\gamma_1 T} \frac{\mathbb{Q}[A_T^c]}{\mathbb{Q}[A_T]} &\leq e^{-(\gamma-\gamma_1)T} \frac{1}{1 - e^{-\gamma T}} \\ &\leq e^{-\gamma_2 T} \qquad \text{for large } T \quad \text{if } \gamma_2 < \gamma - \gamma_1. \end{aligned}$$

• Consider the claim: $X_T := e^{\gamma_1 T} 1_{A_T^c} - \alpha_T 1_{A_T}$. Then,

$$\begin{aligned} X_T &\geq -\alpha_T \geq -e^{-\gamma_2 T} \\ \mathbb{E}^{\mathbb{Q}}[X_T] &= e^{\gamma_1 T} \mathbb{Q}[A_T^c] - \alpha_T \mathbb{Q}[A_T] = 0. \end{aligned}$$

$\rightarrow$ By martingale representation theorem, X_T **is a terminal wealth attainable from zero initial capital**.

• Large deviations estimate for the portfolio:

Since $X_T := e^{\gamma_1 T}\mathbf{1}_{A_T^c} - \alpha_T \mathbf{1}_{A_T}$, we see that

$$\{X_T < e^{\gamma_1 T}\} \;=\; A_T \;:=\; \{Z_T \geq e^{-\gamma T}\}.$$

• Large deviations estimate for the portfolio:

Since $X_T := e^{\gamma_1 T} 1_{A_T^c} - \alpha_T 1_{A_T}$, we see that

$$\{X_T < e^{\gamma_1 T}\} \;=\; A_T \;:=\; \{Z_T \geq e^{-\gamma T}\}.$$

Thus,

$$\begin{aligned}
\mathbb{P}[X_T < e^{\gamma_1 T}] &= \mathbb{P}[Z_T \geq e^{-\gamma T}] \\
&= \mathbb{P}\Big[-\lambda W_T - \frac{\lambda^2}{2} T \geq -\gamma T\Big] \\
&= \Phi\Big(-\Big(\frac{\lambda}{2} - \frac{\gamma}{\lambda}\Big)\sqrt{T}\Big).
\end{aligned}$$

By using again the estimate $\Phi(-d) \sim \varphi(d)/d$ as d goes to infinity, we obtain:

$$\lim_{T\to\infty} \frac{1}{T} \ln \mathbb{P}[X_T < e^{\gamma_1 T}] \;=\; -\frac{1}{2}\Big(\frac{\lambda}{2} - \frac{\gamma}{\lambda}\Big)^2.$$

Remark and extensions

• Explicit form of the terminal wealth

• The key point in the large deviation estimate for the terminal wealth is the exponential decay of the probabilities

$$\mathbb{P}[Z_T \geq e^{-\gamma T}] \;=\; \mathbb{P}\Big[\frac{1}{T}\ln Z_T = -\frac{1}{T}\int_0^T \lambda(S_t)dW_t - \frac{1}{2T}\int_0^T |\lambda(S_t)|^2 dt \;\geq\; -\gamma\Big].$$

Remark and extensions

• Explicit form of the terminal wealth

• The key point in the large deviation estimate for the terminal wealth is the exponential decay of the probabilities

$$\mathbb{P}[Z_T \geq e^{-\gamma T}] \;=\; \mathbb{P}\Big[\frac{1}{T}\ln Z_T = -\frac{1}{T}\int_0^T \lambda(S_t)dW_t - \frac{1}{2T}\int_0^T |\lambda(S_t)|^2 dt \;\geq\; -\gamma\Big].$$

▶ In general, under some ergodic properties, this should follow in principle from Donsker-Varadhan large deviations results.

▶ For particular affine ergodic processes, this can be derived directly from explicit computation of the limiting moment generating functions of $\ln Z_T/T$ and Ellis-Gärtner theorem.

Example: geometric Ornstein-Uhlenbeck process for the stock price (Platen-Rebolledo model)

$$S_t \;=\; \mathrm{exp}(Y_t),$$

where Y is the stationary Ornstein-Uhlenbeck process defined by

$$dY_t \;=\; -\kappa Y_t dt + \sigma dW_t,$$

with parameters $\kappa > 0$, $\sigma > 0$.

$\rightarrow$ Market price of risk: $\lambda(S_t) = -\frac{1}{\sigma}\left(\kappa Y_t - \frac{\sigma^2}{2}\right)$.

Example: geometric Ornstein-Uhlenbeck process for the stock price (Platen-Rebolledo model)

$$S_t \;=\; \mathrm{exp}(Y_t),$$

where Y is the stationary Ornstein-Uhlenbeck process defined by

$$dY_t \;=\; -\kappa Y_t dt + \sigma dW_t,$$

with parameters $\kappa > 0$, $\sigma > 0$.

$\rightarrow$ Market price of risk: $\lambda(S_t) = -\frac{1}{\sigma}\Big(\kappa Y_t - \frac{\sigma^2}{2}\Big)$.

$\rightarrow$ One can then compute explicitly the limiting moment generating functions of $\ln Z_T/T$ (see Florens-Pham 99), and get the large deviations estimate:

$$\lim_{T\to\infty}\frac{1}{T}\ln\mathbb{P}[X_T < e^{\gamma_1 T}] \;=\; \lim_{T\to\infty}\frac{1}{T}\mathbb{P}[Z_T \geq e^{\gamma T}] \;=\; -\frac{\Big(\frac{\sigma^2}{8}+\frac{\kappa}{4}-\gamma\Big)}{\frac{\sigma^2}{8}+\frac{\kappa}{2}-\gamma}.$$

for $0 < \gamma_1 < \gamma < \frac{\sigma^2}{8} + \frac{\kappa}{4}$.

Beating a benchmark: a large deviations approach

• Classical portfolio selection problem rely on expected utility criterion: e.g. Samuelson, Merton.

$\rightarrow$ need to specify the degree of risk aversion of the investor, which is by nature subjective.

• Alternative popular approach: performance of the portfolio relative to the achievement of a given benchmark or index

$\rightarrow$ Beating a benchmark: maximize the probability for the portfolio value to exceed a given index

We look at such outperformance criterion when time horizon goes to infinity:

$\rightarrow$ Of practical interest for institutional managers with long term horizon, e.g. mutual funds

• Infinite horizon problems are usually more tractable than finite horizon problems

$\rightarrow$ provide good insight for management problems with long but finite horizons

Framework

X_t^π: portfolio value with a proportion π invested in stock

B_t: index or benchmark

- The portfolio's performance is measured by the ratio: $\frac{X_t^\pi}{B_t}$

- The ratio $\frac{X_t}{B_t}$ typically grows in time at an exponential rate

$\rightarrow$ The relevant quantity over a long term horizon T is the logarithm of the wealth/index ratio:

$$\bar{X}_T^\pi \;=\; \frac{1}{T} \ln \frac{X_T^\pi}{B_T}$$

A large deviations portfolio selection criterion

• Given a threshold x, the outperformance probability is:

$$\mathbb{P}[\bar{X}_T^\pi \geq x].$$

$\rightarrow$ This outperformance probability decays exponentially fast:

$$\mathbb{P}[\bar{X}_T^\pi \geq x] \;\simeq\; e^{-I(x,\pi)T}, \qquad \text{as } \; T \rightarrow \infty.$$

$\rightarrow$ The lower is the decay rate function $I(x, \pi)$, the more chance there is of realizing an index outperformance:

A large deviations portfolio selection criterion

• Given a threshold x, the outperformance probability is:

$$\mathbb{P}[\bar{X}_T^\pi \geq x].$$

$\rightarrow$ This outperformance probability decays exponentially fast:

$$\mathbb{P}[\bar{X}_T^\pi \geq x] \;\simeq\; e^{-I(x,\pi)T}, \qquad \text{as } \; T \rightarrow \infty.$$

$\rightarrow$ The lower is the decay rate function $I(x,\pi)$, the more chance there is of realizing an index outperformance:

▶ The asymptotic criterion for outperforming the index is:

$$v(x) \;:=\; \sup_\pi \lim_{T\rightarrow\infty} \frac{1}{T} \ln \mathbb{P}[\bar{X}_T^\pi \geq x]$$
$$(= \; -\inf_\pi I(x,\pi))$$

$\rightarrow$ Non standard large deviations control problem!

A dual problem

Formal derivation:

Given a portfolio policy π, the rate function $I(.,\pi)$ associated to the large deviations of the wealth/index log-ratio $\bar{X}_T^\pi$, i.e.

$$\mathbb{P}[\bar{X}_T^\pi \geq x] \;\; \simeq \;\; e^{-I(x,\pi)T}$$

should be given by the Donsker-Varadhan formula:

$$I(x,\pi) \;\; = \;\; \sup_\lambda [\lambda x - \Gamma(\lambda,\pi)]$$

where $\Gamma(.,\pi)$ is the Log-Laplace function:

$$\Gamma(\lambda,\pi) \;\; = \;\; \limsup_{T\to\infty} \frac{1}{T} \ln \mathbb{E}[e^{\lambda T \bar{X}_T^\pi}]$$

Formal derivation of the dual problem (Ctd)

The large deviations criterion can then be written as:

$$\begin{aligned} v(x) \;:=\; \sup_{\pi} \limsup_{T\to\infty} \frac{1}{T} \ln \mathbb{P}[\bar{X}_T^{\pi} \geq x] \;&=\; -\inf_{\pi} I(x,\pi) \\ &=\; -\inf_{\pi} \sup_{\lambda} [\theta x - \Gamma(\lambda,\pi)], \end{aligned}$$

Formal derivation of the dual problem (Ctd)

The large deviations criterion can then be written as:

$$
\begin{aligned}
v(x) \;:=\; \sup_{\pi} \limsup_{T\to\infty} \frac{1}{T} \ln \mathbb{P}[\bar{X}_T^{\pi} \geq x] &= -\inf_{\pi} I(x,\pi) \\
&= -\inf_{\pi} \sup_{\lambda} [\theta x - \Gamma(\lambda,\pi)],
\end{aligned}
$$

▶ By interverting infinum and supremum (!), we get the **duality relation**

$$
v(x) \;=\; -\sup_{\lambda} [\lambda x - \Gamma(\lambda)],
$$

with the **dual control problem** on the moment generating function:

$$
\Gamma(\lambda) \;=\; \sup_{\pi} \Gamma(\lambda,\pi) \;=\; \sup_{\pi} \limsup_{T\to\infty} \frac{1}{T} \ln \mathbb{E}[e^{\lambda T \bar{X}_T^{\pi}}].
$$

$\rightarrow$ This dual problem is rewritten after a change of probability measure as a **risk-sensitive** control problem, and can be solved by dynamic programming Bellman PDE methods.

Optimal portfolio

This duality relation also suggests the following strategy for obtaining the optimal portfolio:

• Solve the dual risk-sensitive control problem: $\Gamma(\lambda)$ and find the associated optimal control $\hat{\pi}(\lambda)$.

• The solution to the large deviations portfolio selection $v(x)$ is then given by

$$v(x) \;=\; -\sup_{\lambda}[\lambda x - \Gamma(\lambda)], \tag{5}$$

with an optimal control determined $\pi^*(x)$ by

$$\pi^*(x) \;=\; \hat{\pi}(\lambda(x)), \qquad \text{where } \lambda(x) \text{ attains the supremum in (4)}, i.e. \quad \Gamma'(\lambda(x)) \;=\; x.$$

▶ This formal derivation can be proved rigorously.

Connection with classical portfolio selection and risk aversion

The dual problem may be written also as:

$$\Gamma(\lambda) = \lim_{T\to\infty} \frac{1}{T} \ln \mathbb{E}\left[U_\lambda\left(X_T^{\widehat{\pi}(\lambda)}\right)\right],$$

where $U_\lambda(x) = x^\lambda$ is a power utility function with Constant Relative degree of Risk Aversion (CRRA): $1-\lambda$.

Connection with classical portfolio selection and risk aversion

The dual problem may be written also as:

$$\Gamma(\lambda) \;=\; \lim_{T\to\infty} \frac{1}{T} \ln \mathbb{E}\left[U_\lambda\left(X_T^{\widehat{\pi}(\lambda)}\right)\right],$$

where $U_\lambda(x) = x^\lambda$ is a power utility function with Constant Relative degree of Risk Aversion (CRRA): $1 - \lambda$.

▶ The duality relation with the Lagrange multiplier $\lambda(x)$ can then written formally as:

$$P[\bar{X}_T^{\pi^*(x)} \geq x] \;\approx\; \mathbb{E}\left[U_{\lambda(x)}\left(X_T^{\pi^*(x)}\right)\right] e^{-\lambda(x)xT},$$

$\rightarrow$ $1 - \lambda(x)$ can be interpreted as a CRRA for an investor who has an over-performance target level x, and it is decreasing with x.

$\rightarrow$ Relate the target level of growth rate to the degree of relative risk aversion in expected utility theory.

Example

A one-factor stock market model:

$$\frac{dS_t^0}{S_t^0} = r(Y_t)dt, \quad \frac{dS_t}{S_t} = \mu(Y_t)dt + \sigma(Y_t)dW_t,$$

with a factor Y as an Ornstein-Uhlenbeck ergodic process:

$$dY_t = \kappa(\theta - Y_t)dt + \vartheta dW_t^1, \quad d < W^1, W > = \rho dt$$

$\kappa > 0$, $\vartheta > 0$.

For simplicity, constant benchmark: $B = 1$.

Dual problem: control problem on the limiting logarithm moment generating function

• Wealth process $X = X^\pi$ controlled by π proportion invested in stock S:

$$dX_t = X_t\Big[\Big(r(Y_t) + (\mu - r)(Y_t)\pi_t\Big)dt + \pi_t\sigma(Y_t)dW_t\Big]$$

▶ Moment generating function of the wealth logarithm $\bar{X}_T^\pi = \frac{1}{T}\ln X_T^\pi$:

$$J_T(\lambda, \pi) := \mathbb{E}\Big[\exp(\lambda T \bar{X}_T^\pi)\Big] = \mathbb{E}^{\mathbb{Q}^\pi}\Big[\exp\Big(\int_0^T \ell(\lambda, Y_t, \pi_t)dt\Big)\Big]$$

with $\ell(\lambda, y, \pi) = \lambda r(y) + \lambda\pi(\mu - r)(y) - \frac{\lambda(1-\lambda)}{2}(\sigma(y)\pi)^2$, and

$$dY_t = (\kappa(\theta - Y_t) + \lambda\rho\vartheta\sigma(Y_t)\pi_t)dt + \vartheta dW_t^\pi,$$

Dual problem: control problem on the limiting logarithm moment generating function

• Wealth process $X = X^\pi$ controlled by π proportion invested in stock S:

$$dX_t = X_t\Big[\Big(r(Y_t) + (\mu - r)(Y_t)\pi_t\Big)dt + \pi_t\sigma(Y_t)dW_t\Big]$$

▶ Moment generating function of the wealth logarithm $\bar{X}_T^\pi = \frac{1}{T}\ln X_T^\pi$:

$$J_T(\lambda,\pi) := \mathbb{E}\Big[\exp(\lambda T\bar{X}_T^\pi)\Big] = \mathbb{E}^{\mathbb{Q}^\pi}\Big[\exp\Big(\int_0^T \ell(\lambda, Y_t, \pi_t)dt\Big)\Big]$$

with $\ell(\lambda,y,\pi) = \lambda r(y) + \lambda\pi(\mu - r)(y) - \frac{\lambda(1-\lambda)}{2}(\sigma(y)\pi)^2$, and

$$dY_t = (\kappa(\theta - Y_t) + \lambda\rho\vartheta\sigma(Y_t)\pi_t)dt + \vartheta dW_t^\pi,$$

▶ Dual control problem: $\Gamma(\lambda) := \sup_\pi \limsup_{T\to\infty}\frac{1}{T}\ln J_T(\lambda,\pi) \to$

$$\Gamma(\lambda) = \sup_\pi \limsup_{T\to\infty}\frac{1}{T}\ln\mathbb{E}^{\mathbb{Q}^\pi}\Big[\exp\Big(\int_0^T \ell(\lambda, Y_t, \pi_t)dt\Big)\Big].$$

$\rightarrow$ **Risk-sensitive control problem**: Fleming, Mc Eneaney, Nagai, Sheu, ...

Formal derivation of the Bellman equation $\leftrightarrow$ **risk-sensitive control** (on infinite horizon)

• Value function for the standard finite-time horizon problem:

$$v_\lambda(T, y) \;:=\; \sup_\pi \mathbb{E}_y^{\mathbb{Q}^\pi}\Big[\exp\Big(\int_0^T \ell(\lambda, Y_t, \pi_t)dt\Big)\Big]$$

$\rightarrow$ Hamilton-Jacobi-Bellman equation (HJB) for v_λ from dynamic programming

• Inverting $\sup_\pi$ and $\limsup_{T\to\infty}$, we expect that $\Gamma(\lambda) = \limsup_{T\to\infty} \frac{1}{T} \ln v_\lambda(T, y)$, and we make the heuristic logarithmic transformation:

$$\ln v_\lambda(T, y) \;=\; \Gamma(\lambda)T + \varphi_\lambda(y).$$

▶ Substitute into the HJB of v_λ

Bellman equation for the dual risk-sensitive control problem

$\rightarrow$ search for a pair $(\Gamma(\lambda), \varphi_\lambda)$ solution to:

$$\begin{aligned}\Gamma(\lambda) &= \frac{1}{2}\vartheta^2\varphi_\lambda''(y) + \frac{1}{2}|\vartheta\varphi_\lambda'(y)|^2 + k(\theta - y)\varphi_\lambda'(y) + \lambda r(y) \\ &\quad + \max_{\pi\in\mathbb{R}}\left[\lambda\sigma(y)\pi(\rho\vartheta\varphi_\lambda'(y) + \frac{(\mu - r)(y)}{\sigma(y)}) - \frac{\lambda(1-\lambda)}{2}(\sigma(y)\pi)^2\right].\end{aligned}$$

with suitable growth condition on φ_λ.

$\rightarrow$ Optimal control given by:

$$\hat{\pi}(\lambda, y) = \frac{(\mu - r)(y)}{\sigma(y)^2(1-\lambda)} + \frac{\rho\vartheta\varphi_\lambda'(y)}{\sigma(y)(1-\lambda)}.$$

Bellman equation for the dual risk-sensitive control problem

$\rightarrow$ search for a pair $(\Gamma(\lambda), \varphi_\lambda)$ solution to:

$$\begin{aligned}\Gamma(\lambda) &= \frac{1}{2}\vartheta^2\varphi_\lambda''(y) + \frac{1}{2}|\vartheta\varphi_\lambda'(y)|^2 + k(\theta - y)\varphi_\lambda'(y) + \lambda r(y) \\ &\quad + \max_{\pi\in\mathbb{R}}\left[\lambda\sigma(y)\pi(\rho\vartheta\varphi_\lambda'(y) + \frac{(\mu - r)(y)}{\sigma(y)}) - \frac{\lambda(1-\lambda)}{2}(\sigma(y)\pi)^2\right]. \end{aligned} \tag{6}$$

with suitable growth condition on φ_λ.

$\rightarrow$ Optimal control given by:

$$\hat{\pi}(\lambda, y) = \frac{(\mu - r)(y)}{\sigma(y)^2(1-\lambda)} + \frac{\rho\vartheta\varphi_\lambda'(y)}{\sigma(y)(1-\lambda)}.$$

Remark.

No unique pair solution to HJB equation (5), even up to a constant for φ_λ! A verification theorem is required to select the good one, and to ensure that it provides indeed the candidate solution to the control problem.

Bellman equation for the dual risk-sensitive control problem

$\rightarrow$ search for a pair $(\Gamma(\lambda), \varphi_\lambda)$ solution to:

$$\begin{aligned}\Gamma(\lambda) &= \frac{1}{2}\vartheta^2\varphi_\lambda''(y) + \frac{1}{2}|\vartheta\varphi_\lambda'(y)|^2 + k(\theta - y)\varphi_\lambda'(y) + \lambda r(y) \\ &\quad + \max_{\pi\in\mathbb{R}}\left[\lambda\sigma(y)\pi(\rho\vartheta\varphi_\lambda'(y) + \frac{(\mu - r)(y)}{\sigma(y)}) - \frac{\lambda(1-\lambda)}{2}(\sigma(y)\pi)^2\right].\end{aligned}$$

with suitable growth condition on φ_λ.

$\rightarrow$ Optimal control given by:

$$\hat{\pi}(\lambda, y) = \frac{(\mu - r)(y)}{\sigma(y)^2(1-\lambda)} + \frac{\rho\vartheta\varphi_\lambda'(y)}{\sigma(y)(1-\lambda)}.$$

▶ For $\mu(y)$ and $r(y)$ linear in y, σ constant $\rightarrow$ explicit solutions:

we look for a quadratic solution φ_λ:

$$\varphi_\lambda(y) = \frac{1}{2}A(\lambda)y^2 + B(\lambda)y.$$

Explicit calculations

Black-Scholes model: $dS_t = S_t(\mu dt + \sigma dW_t)$.

The solution to the dual problem is

$$\Gamma(\lambda) \;=\; \frac{1}{2}\frac{\lambda}{1-\lambda}\left(\frac{\mu - r}{\sigma}\right)^2, \quad \lambda \in [0,1).$$

The solution to the (primal) large deviations problem is

$$v(x) \;=\; \begin{cases} -(\sqrt{x} - \sqrt{\bar{x}})^2, & \text{if } x \geq \bar{x} := \frac{1}{2}\left(\frac{\mu-r}{\sigma}\right)^2 \\ 0, & \text{if } x < \bar{x}, \end{cases}$$

and the optimal portfolio proportion is constant given by:

$$\pi_t^*(x) \;=\; \begin{cases} \sigma\sqrt{2x}, & \text{if } x \geq \bar{x} \\ \frac{\mu-r}{\sigma^2}, & \text{if } x < \bar{x}. \end{cases}$$

$\rightarrow$ CPP (Constant Proportion Portfolio) strategy

Platen-Rebolledo model: $S = e^Y$: $dY_t = \kappa(\theta - Y_t)dt + \sigma dW_t$

The solution to the dual problem is

$$\Gamma(\lambda) \;=\; \frac{\kappa}{2}\Big[1 - \sqrt{1-\lambda}\Big] + \frac{\lambda}{2}\left(\frac{\kappa\theta - r + \frac{1}{2}\sigma^2}{\sigma}\right)^2, \quad \lambda \in [0,1).$$

The solution to the (primal) large deviations problem is

$$v(x) \;=\; \begin{cases} -\frac{(x-\bar{x})^2}{x-\bar{x}+\frac{\kappa}{4}}, & \text{if } \; x \geq \bar{x} := \frac{1}{2}\left(\frac{\kappa\theta - r + \frac{1}{2}\sigma^2}{\sigma}\right)^2 + \frac{\kappa}{4} \\ \quad 0, & \text{if } \; x < \bar{x}, \end{cases}$$

and the optimal portfolio proportion is:

$$\pi_t^*(x) \;=\; \begin{cases} -\frac{[4(x-\bar{x})+\kappa]}{\sigma} Y_t + \frac{\kappa\theta - r + \frac{1}{2}\sigma^2}{\sigma^2}, & \text{if } x \geq \bar{x} \\ \\ -\frac{\kappa}{\sigma} Y_t + \frac{\kappa\theta - r + \frac{1}{2}\sigma^2}{\sigma^2}, & \text{if } x < \bar{x}. \end{cases}$$

References

▶ Asymptotic arbitrage and large deviations

• Föllmer H. and W. Schachermayer (2008): "Asymptotic arbitrage and large deviations", *Math. Fin. Econ.*, **1**, 213-249.

• Florens D. and H. Pham (1999): "Large deviations in estimation of an Ornstein-Uhlenbeck model", *Journal of Applied Probability*, **36**, 60-77.

▶ Optimal long term investment and large deviations approach

• Pham H. (2003a): "A large deviations approach to optimal long term investment'", *Finance and Stochastics*, **7**, 169-195.

• Pham H. (2003b): "A risk-sensitive control dual approach to a large deviations control problem", *Systems and Control Letters*, **49**, 295-309.

▶ **Extensions and variations**:

• Hata H. and J. Sekine (2005): "Solving long term investment problems with Cox-Ingersoll-Ross interest rates", *Advances in Mathematical Economics*, **8**, 231-255.

• Hata H., Nagai H. and S. Sheu (2009): "Asymptotics of the probability minimizing a down-side risk", to appear in *Annals of Applied Probablity*.

• Guasoni P. and S. Robertson (2009): "Portfolio and risk premia for the long run", Preprint.

• Knispel T. (2010): "Asymptotics of robust utility maximization", Preprint, University of Leibniz.

Lecture notes forthcoming ....

THANK YOU FOR YOUR ATTENTION