Large Deviations Applications to Finance

Modern large deviations theory, pioneered by Donsker and Varadhan [12], concerns the study of rare events and it has become a common tool for the analysis of stochastic systems in a variety of scientific disciplines and in engineering. The theory developed by Donsker and Varadhan is a generalization of Laplace’s principle and Cramér’s theorem. Here we concentrate in applications to finance and risk management.

Large deviations theory is discussed in this encyclopedia. A wealth of monographs discuss the topic in detail, see for instance (Dembo and Zeitouni [9], Deuschel and Strook [10], Dupuis and Ellis [11]). Here, we shall discuss applications to mathematical finance, that includes option pricing, risk management and portfolio optimization (the reader may also consult [21], which includes details behind many of the topics touched here). Before discussing these applications, however, let us just provide a brief introduction to some basic concepts underlying the theory of large deviations.

Definition: A sequence of random objects \((Z_n : n \geq 0)\) taking values on some topological space \(S\), satisfies a large deviations principle with rate function \(J(z) : z \in S\), if for each Borel measurable set \(A \in S\)

\[
\sup_{z \in A^c} \inf_{n \to \infty} - \frac{1}{n} \log P (Z_n \in A) \leq \lim_{n \to \infty} - \frac{1}{n} \log P (Z_n \in A) \leq \inf_{z \in A} J(z),
\]

and \(J(\cdot)\) is non-negative and upper semicontinuous. Here \(A^c\) and \(\overline{A}\) stand respectively for the interior and closure of \(A\).

Throughout our discussion we may assume that \(S\) is a Polish space (i.e. a separable, completely metrizable topological space). One can show that if a large deviations principle holds, then there exists a deterministic element \(\tilde{z} \in S\) such that \(Z_n \to \tilde{z}\) (where \(\to\) denotes weak convergence). In many applications, \(S\) is a function space, so \(Z_n\) is a stochastic process and \(\tilde{z}\) is the asymptotically most likely path.

A large deviations principle is intuitively interpreted as having the formal approximation

\[
P (Z_n \approx z) \approx \exp (-n J(z)),
\]

for \(z\) outside a neighborhood of \(\tilde{z}\). The previous approximation does not carry any rigorous meaning, nevertheless, the formal use of (1) often allows to infer large deviations properties of stochastic systems that can be later justified rigorously using large deviations theory. The rigorous meaning behind (1) is given by Varadhan’s lemma (see Dembo and Zeitouni [9], p. 137), which states that, in the presence of a large deviations principle, for every continuous function \((F(z) : z \in S)\) bounded from below we have that

\[
\log \frac{1}{n} E \exp (-n F(Z_n)) \to - \inf_{z \in S} (J(z) + F(z))
\]

as \(n \to \infty\). Varadhan’s lemma is a generalized version of Laplace’s principle. Indeed, note that in view of representation (1), Varadhan’s lemma makes perfect intuitive sense after applying Laplace’s principle since

\[
E \exp (-n F(Z_n)) \approx \int \exp (-n (J(z) + F(z))) dz 
\]

\[
\approx \exp \left(- \inf_{z \in S} (J(z) + F(z))\right).
\]

The solution to the optimization problem \(\inf_{z \in S} (J(z) + F(z))\) is the so-called minimum energy path or optimal path. Formally, if we put \(F(z) = \infty \times I_A(z)\), then \(E \exp (-n F(Y)) = P(Y \in A)\) and if \(\tilde{z} \notin A\) and regularity assumptions on the set \(A\) hold Varadhan’s lemma yields

\[
P (Z_n \in A) \approx \exp \left(- \inf_{z \in A} J(z)\right).
\]

A related class of asymptotic approximations arises in the setting of dynamical systems with a small random perturbation, for instance,

\[
dZ_{\varepsilon} (t) = \mu (Z_{\varepsilon} (t)) dt + \varepsilon \sigma (Z_{\varepsilon} (t)) dB (t);
\]

\[
Z_{\varepsilon} (0) = z_{\varepsilon} (0);
\]

assuming the necessary regularity conditions for the existence of a solution to the previous SDE (stochastic differential equations) driven by a standard Brownian motion \(B(\cdot)\). Note that, formally, if we send \(\varepsilon \to 0\) and assume that \(z_{\varepsilon} (0) \to \tilde{z} (0)\), we obtain convergence of \(Z_{\varepsilon}\) to a deterministic dynamical system satisfying \(\tilde{z} (t) = \mu (\tilde{z} (t))\) given \(\tilde{z} (0)\). Therefore, under appropriate conditions one would expect the existence of a theory backing up approximations such as (1) given by an appropriate rate function \(I(\cdot)\). Such theory indeed can be developed and it is...
known as Freidlin-Wentzel theory (see, Dembo and Zeitouni [9], p. 212 or Shwartz and Weiss [22], p. 129). The rest of our discussion concentrates on application of large deviations ideas in finance.

**Option pricing**

Direct use of large deviations approximations, Large deviations principles are applied in finance in order to develop approximations for option prices. It is easier to explain the techniques with a simple example. Consider the problem of pricing a digital knock-in call option with maturity time $T$ under a Black-Scholes economy, in particular,

$$\alpha_T = P \left( \min_{0 \leq u \leq T} B (s) \leq -a, B (T) > b \right)$$

(2)

for some barrier values $a, b > 0$, and $B (\cdot)$ is a standard Brownian motion. This probability can be, of course, evaluated in closed form (see for instance [25], [26] or equation (4) below), but we shall illustrate the use of large deviations theory here. We will develop approximations that can be asymptotically validated when the time to maturity and the barriers are large. First, we embed the problem in a large deviations setting by introducing an appropriate scaling

$$\alpha_T = P \left( \min_{0 \leq u \leq 1} Z_T (u) \leq -c, Z_T (1) \geq d \right),$$

where $a = -Te$ and $b = Td$, and $Z_T (u) = B (uT) / T$, $u \in [0, 1]$. The rate function of $Z_T (\cdot)$ is defined on the space $C_1 := C_1 ([0, 1])$ of absolutely continuous functions and takes the form

$$J (z) = \frac{1}{2} \int_0^1 \dot{z} (u)^2 \, du.$$

Note that $Z_T \to \bar{z} = 0$ uniformly in probability. A formal application of (1) combined with Laplace’s principle then yields

$$\alpha_T = \exp \left( -T \inf_{z \in A} J (z) + o (T) \right),$$

(3)

where

$$A = \{ z \in C_1 : z (0) = 0, \min_{0 \leq u \leq 1} z (u) \leq -c, z (1) \geq d \}.$$

One then can show directly using standard techniques from calculus of variations (applying Euler-Lagrange’s principle) that the optimal path is a piecewise linear function. Note that (1) also suggests that

the optimal path is the most likely way in which the particular large deviations event, given by $A$, can occur. In our simple example one can directly evaluate such optimal path using the reflection principle. In particular, note that

$$\alpha_T = P (B (T) > (2a + b)T) = P (Z_T (1) > 2c + d)$$

(4)

which allows to conclude (again by reflection) that

$$z^* (t) = I (t \leq t^*) \theta_+ t + (\theta_+ t - c) I (1 \leq t \leq t^*),$$

(5)

where $\theta_- = -(2c + d)$, $\theta_+ = -\theta_-$ and $t^* = c/(2c + d)$. Moreover, we have that $J (z^*) = (2c + d)^2 / 2$. It is important to note using (4) and elementary properties of the Gaussian density that

$$\alpha_T \sim \frac{\exp (-TJ (z^*))}{(2\pi)^{1/2} T^{1/2}}$$

(6)

as $T \to \infty$. In particular, it is worth noting that the large deviations results such as (3) provide only rough approximations because no information is given about premultiplying terms like the factor $T^{-1/2}$ which appears in (6).

Enhancing Monte Carlo simulations. As we indicated before, large deviations approximations typically provide only logarithmic asymptotics (i.e. only the exponential rate of decay is identified without any additional information). Sharp asymptotics (i.e. asymptotics with information about premultiplying factors such as (6)) can only be developed under additional problem structure. Let us continue discussing pricing issues in the context of barrier options; the one described in previous paragraph is just an example of many types. In evaluating barrier options an important ingredient relates to calculations of exit probabilities. For instance, in Baldi et al [5] such calculations are treated combining both Monte Carlo methods and large deviation principles. The model in use is a geometric Brownian motion under the risk neutral measure

$$dS_t = rS_t \, dt + \sigma S_t \, dB_t, \quad S_0 = x.$$  

(7)

Using this model, we consider computing the probability that the process $S$ solving (7) does not hit any of two barriers, a lower or an upper barrier, which are suitable positive twice continuously differentiable functions ($l(t), t \geq 0$) and ($u(t), t \geq 0$), respectively.
The approach suggested by Baldi et al [5] consists in using sharp large deviations to reduce bias in Monte Carlo simulation. The simulation of $S$ is done by making an equi-distance partition of the time interval $[0, T]$, where $T$ indicates the expiration of the contract. Then, the sample path is the collection

$$S_{t_{i+1}} = S_{t_i} \exp \left( (r - \frac{\sigma^2}{2}) t_i + \sigma (B_{t_{i+1}} - B_{t_i}) \right),$$

$i = 0, 1, \ldots, m$; where $t_0 = 0 < t_1 < \ldots < t_m = T$ and $t_{i+1} - t_i = \epsilon > 0$.

Using sharp asymptotics, it is possible to find approximations of the exit probabilities over small intervals, i.e. when $\epsilon$ becomes small. Let $\varsigma \in (l(t_i), u(t_i))$ and $y \in (l(t_{i+1}), u(t_{i+1}))$, then the approximation appears as follows

$$p_i^\epsilon := P(\exists t \in [t_i, t_{i+1}] : S_t \notin ([l(t), u(t)]) \mid S_{t_i} = \varsigma, S_{t_{i+1}} = y) = f(t_i, \varsigma, y, \epsilon)(1 + O(\epsilon)),$$

where the function $f$ is indeed known explicitly:

$$f(t_i, \varsigma, y, \epsilon) = \exp \left( \frac{-Q(t_i, \varsigma, y)}{\epsilon} - R(t_i, \varsigma, y) \right),$$

with

$$Q(t_i, \varsigma, y) = \begin{cases} \frac{2(u(t_i) - \varsigma)(u(t_i) - y)}{\sigma^2} & \text{if } \varsigma + y > u(t_i) + l(t) \\ \frac{2(c(t_i) - (u(t_i) - l(t))}{\sigma^2} & \text{if } \varsigma + y < u(t_i) + l(t) \end{cases}$$

and

$$R(t_i, \varsigma, y) = \begin{cases} \frac{2(u(t_i) - \varsigma)u'(t_i)}{\sigma^2} & \text{if } \varsigma + y > u(t_i) + l(t) \\ \frac{2(c(t_i) - (u(t_i) - l(t))}{\sigma^2} & \text{if } \varsigma + y < u(t_i) + l(t). \end{cases}$$

The final estimator, with reduced bias, for the probability that $S$ does not hit the barrier in the interval $[0, T]$ is constructed by simulating $N$ i.i.d. replications $(S^{(j)} : j \leq N)$ of the process $S$ and obtaining

$$\frac{1}{N} \sum_{j=1}^{N} \prod_{i=0}^{m-1} I \left( S_{t_i}^{(j)} \in (l(t_i), u(t_i)) \right) p \left( t_i, S_{t_i}^{(j)}, S_{t_{i+1}}^{(j)} \right),$$

where

$$p \left( t_i, S_{t_i}^{(j)}, S_{t_{i+1}}^{(j)} \right) = 1 - f \left( t_i, S_{t_i}^{(j)}, S_{t_{i+1}}^{(j)}, \epsilon \right).$$

Large deviations analysis can also aid the application of variance reduction techniques for Monte Carlo simulation via importance sampling, as we will discuss next.

Importance sampling. Once again, we concentrate on a concrete problem in order to illustrate the use of large deviations in the design of importance sampling. Consider the problem of calculating (2). The optimal path (5) suggests a particular way in which one could bias the occurrence of the large deviations event; here we also consider time maturity and barrier values to be large, thus we would be dealing with small probabilities. In particular, consider a probability measure, say $\mathbb{Q}$, under which a process $(Y(u) : 0 \leq u \leq T)$ follows a Brownian motion with drift $\theta_-$ up to time $\tau_{-a} = \inf \{ s \geq 0 : Y(u) \leq -Ta \}$, and from time $\tau_{-a}$ up to time $T$ (assuming $\tau_{-a} < T$) the drift of $Y(\cdot)$ is switched to the value $\theta_+$. Using the process $Y$ we obtain the representation

$$\alpha_T = E_Q \left( \min_{0 \leq s \leq T} Y(s) \leq -a, Y(T) > bT \right) \times h.$$

The expression involving the exponentials in the previous display is nothing but the likelihood ratio between the Wiener measure $\mathbb{P}$ (corresponding to the process $(B(s) : 0 \leq s \leq T)$) and the measure $\mathbb{Q}$ on the set $I(\tau_{-a} < T)$. More precisely,

$$\frac{d\mathbb{P}}{d\mathbb{Q}} I(\tau_{-a} < T) = I(\tau_{-a} < T) \times e^{-\theta_- a + \theta_+^2 (\tau_{-a} - \tau_{-a})/2}(T) + \theta_+(T) - \tau_{-a}/2.$$

The use of importance sampling proceeds by simulating, say, $N$ i.i.d. replications $(Y_i : i \leq N)$ of the process $(Y(s) : 0 \leq s \leq T)$ and estimating

$$P(\min_{0 \leq s \leq T} B(s) \leq -a, B(T) > b) \text{ via}$$

$$\alpha_T^{(N)} = \frac{1}{N} \sum_{i=1}^{N} \left( I(\tau_{-a}^{(i)} < T, Y_i(T) > bT) \right) \times h^{(i)},$$

where

$$h^{(i)} = e^{-\theta_- a + \theta_+^2 (\tau_{-a} - \tau_{-a})/2)(T)}(Y_i(T) + a) + \theta_+(T) - \tau_{-a}/2.$$

Here $\tau_{-a}^{(i)}, i = 1, \ldots, N$ are the i.i.d. replications of $\tau_{-a}$ obtained from the corresponding i.i.d. processes.
The likelihood ratio $L$ and Brownian motion under $Q$ process (see, for instance, Asmussen and Glynn [3], Ch. 7). Additional extensions of this methodology are applicable to other types of processes. However, to design optimal importance sampling estimators it is required to compute the associated optimal path under the corresponding rate function of the underlying process. Research in this direction has been developed recently (see Guasoni and Robertson [17] and Asmussen and Glynn [3]).

Let us now give a broader description of the importance sampling technique in option pricing. When pricing an option, the question is to find the expectation $I_g = E(g(S_t : 0 \leq t \leq T))$, \begin{equation}
I_g = E(g(S_t : 0 \leq t \leq T)), \tag{8}
\end{equation}
where $g$ is a payoff function (possible path dependent), and $S$ is governed by a suitable stochastic differential equation. Let us concentrate here in the geometric Brownian motion given by (7). The naive Monte Carlo estimator based on $N$ replications is
\begin{equation}
\tilde{I}_g^N = \frac{1}{N} \sum_{i=1}^{N} g(S_t^{(i)} : 0 \leq t \leq T),
\end{equation}
where the $S_t^{(i)}$s are i.i.d. replications (see e.g. [3] as a general reference on stochastic simulation) of the process $S$ following (7).

The importance sampling procedure turns (8) into
\begin{equation}
I_g = E_Q(g(S_t^Q : 0 \leq t \leq T) \times L_T^{-1}),
\end{equation}
where
\begin{equation}
dS_t^Q = (r + \sigma \phi_t) S_t^Q dt + \sigma S_t^Q dB_t^Q \tag{9}
\end{equation}
and the process $\left(B_t^Q: 0 \leq t \leq T\right)$ is a standard Brownian motion under $Q$. The relation between $B_t^Q$ and $B_t$ is given by
\begin{equation}
dB_t = \phi_t dt + dB_t^Q.
\end{equation}
The likelihood ratio $L_T^{-1} = dP/dQ$ satisfies
\begin{equation}
L_T^{-1} = \exp \left( - \int_0^T \phi_s dB_s - \frac{1}{2} \int_0^T \phi_s^2 ds \right)
= \exp \left( - \int_0^T \phi_s dS_s^Q - \frac{1}{2} \int_0^T \phi_s^2 ds \right).
\end{equation}
Under this specifications, the importance sampling estimator takes the form
\begin{equation}
I_{g, IS}^N = \frac{1}{N} \sum_{i=1}^{N} g(S_t^{(i)} : 0 \leq t \leq T) L_T^{-1}(i),
\end{equation}
where $S_t^{(i)}$ and $L_T^{-1}(i)$ are obtained from the corresponding i.i.d. copies of the process $S^Q$ which follows the evolution equation (9). Therefore, the ultimate problem is to find a process $\phi$ (might or not be deterministic) that guaranties an efficient estimation of $I_g$ in terms of variance reduction. As we indicated in the simple barrier example, large deviation techniques provide tools to find such process $\phi$. In the next subsection we will discuss another example.

\textbf{Freidlin-Wentzel theory}

Let us suppose that we are dealing with a European option, thus (8) becomes
\begin{equation}
I_g(x) = E(g(S_T)|S_0 = x),
\end{equation}
and we write $I_g(x)$ to emphasize the dependence on the initial position.

Using \textbf{Itô’s formula}, under measure $Q$, the variance of $I_{g, IS}^N$ is given by
\begin{equation}
\text{Var}_Q(I_{g, IS}^N) = \frac{1}{N} E_Q \left( \int_0^T L_t (\sigma I_t^g(S_t^Q) + \phi_t I_t^g(S_t^Q))^2 dt \right),
\end{equation}
where $I_t^g$ is the derivative with respect to $x$. One can readily see that if $\phi_s = -\sigma I_t^g(S_t^Q)/I_t^g(S_t^Q)$ the variance of the estimator would vanish. Unfortunately, the function $I_t^g(\cdot)$ is precisely what we want to find, but the idea is to find an approximation to $I_t^g(\cdot)$, say $\tilde{I}_t(\cdot)$, and consider $\tilde{I}_t = -\sigma \tilde{I}_t \left( S_t^Q \right) / I_t \left( S_t^Q \right)$ which can be used to generate a change of measure that reduces the variance of the associated importance sampling estimator. For instance, Fourniè et al ([14]) suggest developing a parametric expansion for $I_t^g(x)$ as a function of $\sigma$ as $\sigma \searrow 0$. In turn, such an expansion is based, for options out of the money, on the Freidlin-Wentzel theory.

\textbf{Varadhan-Laplace principle}

Recall that under the geometric Brownian motion, model (7), the path simulation at discrete-time points satisfies

\begin{equation}

\end{equation}

\[ S_{t+1} = S_t \exp \left( \left( r - \frac{\sigma^2}{2} \right) \epsilon + \sigma \sqrt{\epsilon} Z_i \right), \]

where \( Z_i \sim N(0,1) \) for \( i = 1, \ldots, n \). Let \( Z = (Z_1, \ldots, Z_n)^\top \) (\( ^\top \) stands for the transpose), then we denote by \( G(Z) \) the payoff of the option under a sample \( Z \), for instance, for the Asian option we have \( G(Z) = \max(0, \frac{1}{n} \sum_{i=1}^n S_i - K) \). In this case the partition indicates the times when the process is monitored to take the average.

The procedure developed in Glasserman et al [15] is to change the mean of \( Z \) from 0 to a vector \( \mu \) to obtain an estimator of \( I_g \). Multiplying each sample by the corresponding importance sampling weight or likelihood ratio the estimator then is

\[ I_g^N = \frac{1}{N} \sum_{i=1}^N G(Z^{(i)}) \exp \left( -\mu^\top Z^{(i)} + \frac{1}{2} \mu^\top \mu \right), \]

where \( Z^{(i)} \) is the vector \( (Z_1^{(i)}, \ldots, Z_n^{(i)})^\top \) of independent random variables such that \( Z_j^{(i)} \sim N(\mu_j, 1), j = 1, \ldots, n \). The \( Z^{(i)} \)’s, \( i = 1, \ldots, N \) are i.i.d. replications. To choose \( \mu \) it suffices to minimize the variance of \( I_g^N \), which turns out to be equivalent to minimizing

\[ E \left( G(Z)^2 \exp \left( -\mu^\top Z + \frac{1}{2} \mu^\top \mu \right) \right), \tag{10} \]

where \( Z \) here is an \( n \)-dimensional vector of i.i.d. standard Gaussian random variables. It is not an easy task to solve this optimization problem, but it can be tackled using Varadhan-Laplace asymptotics. The idea is to scale \( Z \) by \( \sqrt{\epsilon} \) and asymptotically estimate the expectation in (10) as \( \epsilon \to 0 \). This approach makes it simpler to find the optimal \( \mu \), at least as an approximation, which is asymptotically correct in an environment of small volatility, i.e. \( \sigma \) close to 0 in (7).

Another interesting application of large deviations principles is in the context of index option pricing. Consider an index \( H \) composed of \( m \) stocks \( (S_1, \ldots, S_m) \). The index at time \( t \) is computed as \( H(t) = \sum_{i=1}^m w_i S_i \), where \( w_1, \ldots, w_m \) are constants.

Pricing, for instance, a European call on the index involves knowing the so-called local volatility function of the index. What is proposed in Avellaneda et al [4] is to approximate such function using Varadhan’s principle in order to handle it in a simpler way, and therefore, pricing the option.

**Risk management**

The use of large deviations theory for computational purposes also arises in the context of risk management. For instance, Dembo, Deuschel and Duffie [8] developed approximations based on large deviations theory for the tail of a loss distribution, a relevant assumption is that the individual losses are conditionally i.i.d., given the state of the economy and their identifying class (these parameters can be given by the specific industry or business line).

**Credit risk.** Typically, an important task in credit risk theory involves computing the distribution of losses in a portfolio composed of several debt contracts. More precisely, in a portfolio composed of \( n \) obligors, the question is to calculate the tail probability of \( L_n = c_1 Y_1 + \ldots + c_n Y_n \),

\[ P(L_n > l), \tag{11} \]

where the \( Y_i \)’s are Bernoulli random variables such that \( P(Y_i = 1) = p_i = 1 - P(Y_i = 0) \), and indicating that the \( i \)th obligor (for \( i = 1, \ldots, n \)) may or may not default. The \( c_i \)’s represent the loss resulting from the default, and \( l \) is a threshold. Generally, the number of obligors is large, and surpassing threshold \( l \) may be a rare event which represents significant losses, in which case one may use large deviations theory to approximate the probability in (11). In Glasserman et al [16] such probability is approximated under a multifactor Gaussian copula model and using large deviation theory; here we present some results that can be found in [21].

Suppose that the variables \( Y_i, i = 1, \ldots, n \) are triggered by other variables \( X_i, i = 1, \ldots, n \) that might or not be related, this is done in the following way: \( Y_i = 1_{\{X_i > x_i\}}, i = 1, \ldots, n \) and the vector \( (X_1, \ldots, X_n) \) is normally distributed. The parameters \( x_i \) are such that \( P(X_i > x_i) = p_i \) for \( i = 1, \ldots, n \). The correlations among variables \( \{X_1, \ldots, X_n\} \) are determined by the following single factor model:

\[ X_i = \rho Z + \sqrt{1-\rho^2}Z_i, i = 1, \ldots, n, \]

where \( \rho \in [0, 1) \) and \( Z, Z_1, \ldots, Z_n \) are independent standard normal r.v.s.

When \( \rho = 0 \) there is no correlation, and using Cramér’s theorem one can find an asymptotic for-
mula. Suppose that \( p_i = p, i = 1, \ldots, n \), then
\[
\lim_{n \to \infty} \frac{1}{n} \log P(L_n > nq) = -q \log(q/p) \log \left( \frac{1 - q}{1 - p} \right),
\]
where \( l = nq \) and \( q \in (p, 1) \).

When \( \rho > 0 \), i.e. there is dependence among obligors, it is also possible to derive an asymptotic result. Indeed, if \( q_n = 1 - n^{-a} \) with \( a \in (0, 1) \), it is shown in [21] that
\[
\lim_{n \to \infty} -\frac{1}{n} \log P(L_n > nq_n) = -a \frac{1 - \rho^2}{\rho^2}.
\]

A more elaborated model of this type is treated in [16].

We have discussed the use of large deviations theory in several computational settings, both in pricing and risk assessment. Another application of large deviations, also in the context of risk management, arises in the theoretical analysis of risk measures.

**Risk measures.** Financial institutions are constantly worried about the quantification of the risk in their portfolios of assets: stocks, bonds, credits and options. A financial asset is generally represented by a random variable, say \( X \), and characterized by the probability measure \( \mu \) that governs the outcome of \( X \), that is, we write \( \mathcal{L}(X) = \mu \) whenever \( P(X \in A) = \mu(A) \). The risk associated to \( X \) is measured by means of a so-called risk measure \( \rho \), which is a real valued function on the space of random variables that satisfies certain properties (such as monotonicity, subadditivity, positive homogeneity and translation invariance); see Artzner et al [2].

Assume that it is not possible to deal with the probability measure \( \mu \) directly, but instead we have independent samples \( X_1, \ldots, X_n \) of \( X \). Then, one may want to approximate \( \rho(X) \) by calculating \( \rho(X^{(n)}) \); in this case \( X^{(n)} \) is a random variable and its law \( \mathcal{L}(X^{(n)}) \) is the empirical measure
\[
\mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}.
\]

Here \( \delta_x \) stands for the Dirac measure at \( x \).

This approach generates an error, and the question suggested in Weber [27] is to quantify the error by considering the asymptotics, as \( n \to \infty \), of
\[
P \left( \left| \rho(X) - \rho(X^{(n)}) \right| > \epsilon \right), \text{ for } \epsilon > 0.
\]

Then, for positive \( \epsilon \) and when \( n \to \infty \), the previous quantity becomes small, and large deviation principles can be used to deal with this probability. An important result used for this is Sanov’s theorem (see e.g. [9]).

**Optimal Investment.**

We now move to investment. Maximizing revenue in investments under a certain risk profile (which could be induced by a specific set of assets types) has been a common question for financial investors. To tackle the problem usual approaches are the mean-variance analysis of Markowitz or the use of utility functions. Alternatively, a criterion can be set by considering asymptotic performance of the portfolio over long time horizons (see for instance [1]). One has to decide how to measure such performance, and one way to do this is by estimating, given a fixed investment policy, the probability of performing well, (see [13, 7]). The optimization problem can be set as selecting the investment policy that minimizes the probability of ending up below a threshold or that maximizes the probability of surpassing it. Since the problem involves calculating an asymptotic probability, it is appealing to use large deviations techniques (one can find this approach in [23, 24, 19, 20]).

Below, as an illustration, we describe the method described by Pham in [19] and [20]. Given an investment policy \( \alpha \in \mathcal{A} \) (where \( \mathcal{A} \) is the set of admissible investment policies), we consider the rate of return of the associated wealth process (i.e. the logarithm of the value of the portfolio obtained by applying the policy \( \alpha \)). We denote such rate of return process by \( X(\alpha) = (X_t(\alpha), t \geq 0) \). The aim is to maximize over \( \mathcal{A} \) the probability \( P(X_t(\alpha)/t \geq x) \) in the long term, which is the probability of a rare event as \( t \to \infty \). Here, level \( x \in \mathbb{R} \) represents a benchmark that an investor wants to achieve; in [7] it is considered an stochastic benchmark, such as an index.

It is natural to attempt using large deviations theory to find a (static) policy \( \alpha^* \) from
\[
v(x, \alpha^*) := \sup_{\alpha \in \mathcal{A}} v(x, \alpha),
\]

where
\[
v(x, \alpha) = -\lim_{t \to \infty} \frac{1}{t} \log P(X_t(\alpha)/t > x).
\]

Under regularity conditions on the drift and the volatility of the process \( X_t(\alpha) \) (for instance if the
drift and diffusion coefficients are bounded and continuously differentiable), one can evaluate \( v(x) \) via
\[
v(x,a) = \sup_{\theta} (\theta x - \Gamma(\theta,a)),
\]
where
\[
\Gamma(\theta,a) := \lim_{t \to \infty} \frac{1}{t} \ln E \left( e^{\theta X_t(a)} \right).
\]
This approach is considered in [19]. For instance, if \( X_t(a) = \alpha Y(t) \) where \( \alpha \in \mathbb{R} \) and \( Y(t) = rt + \sigma B(t) \) with \( r > 0 \) and \( B(\cdot) \) is a standard Brownian motion, we have that
\[
\Gamma(\theta,a) = \theta \alpha r + \frac{(\theta \alpha \sigma)^2}{2},
\]
and then the problem is solved when \( \alpha = x/r \) (see [21]). More sophisticated and also explicit calculations can be found in [19].

References


Jose H. Blanchet

and

Carlos G. Pacheco-González