Available online at www.sciencedirect.com
ScienceDirect
C. R. Acad. Sci. Paris, Ser. I 346 (2008) 209–212
COMPTES RENDUS MATHEMATIQUE
http://france.elsevier.com/direct/CRASS1/

Probability Theory

The Neyman–Pearson lemma under g-probability ☆,☆☆

Shaolin Ji [a], Xun Yu Zhou [b,c]

[a] School of Mathematics and System Sciences, Shandong University, Jinan, Shandong, 250100, PR China
[b] Mathematical Institute, University of Oxford, 24–29 St Giles, Oxford OX1 3LB, UK
[c] Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong

Received 6 September 2007; accepted after revision 11 December 2007

Available online 9 January 2008

Presented by Paul Deheuvels

Abstract

The Neyman–Pearson fundamental lemma is generalized under g-probability. With convexity assumptions, a sufficient and necessary condition which characterizes the optimal randomized tests is obtained via a maximum principle for stochastic control. ***To cite this article: S. Ji, X.Y. Zhou, C. R. Acad. Sci. Paris, Ser. I 346 (2008).***

Résumé

Lemme de Neyman–Pearson généralisé pour les g-espérances. Le lemme fondamental de Neyman–Pearson est généralisé au cas de g-probabilités. Sous des hypothèses de convexité, une condition suffisante et nécessaire caractérisant le test randomisé optimal est obtenue au moyen du principe du maximum dans le cadre du contrôle stochastique. ***Pour citer cet article : S. Ji, X.Y. Zhou, C. R. Acad. Sci. Paris, Ser. I 346 (2008).***

1. Introduction

It is well known that the Neyman–Pearson fundamental lemma gives the most powerful statistical tests for simple hypothesis testing problems. However, to the best of our knowledge little is known about the nonlinear probability counterpart except Huber and Strassen's work [10] for 2-alternating capacities. Recently some special capacities are found to be non-2-alternating in the context of modeling super- and sub-pricing of contingent claims in an incomplete financial market (see [2,3]), and such capacities can be described by the so-called g-probability which is introduced through a nonlinear mathematical expectation, i.e., g-expectation by Peng [13]. Thus, it is of interest to study the simple hypothesis testing under g-probability.

☆ The authors thank the partial support from the National Basic Research Program of China (973 Program, No. 2007CB814900), the RGC Earmark Grant No. 418606, and a start-up fund at Oxford.
☆☆ This Note is the succinct version of a text on file for five years in the Academy Archives.
E-mail addresses: jsl@sdu.edu.cn (S. Ji), zhouxy@maths.ox.ac.uk (X.Y. Zhou).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2007.12.007

In this Note, with convexity assumptions, a sufficient and necessary condition which characterizes the optimal randomized test is obtained by using a stochastic maximum principle approach. Two examples are given to show the statistical interpretation of our results as well as an application to a shortfall risk minimization problem.

2. Problem formulation

Let $W(\cdot)$ be a standard d-dimensional Brownian Motion defined on a complete probability space $(\Omega, \mathcal{F}, P)$. The information structure is given by a filtration $F = \{\mathcal{F}_t\}_{0 \leqslant t \leqslant T}$, which is the augmented σ-algebra generated by the Brownian Motion $W(\cdot)$. Let $\mathcal{F}_T = \mathcal{F}$ and $g = g(y, z, t)$: $R^1 \times R^{1\times d} \times [0, T] \to R^1$ be a function satisfying

(H1) g is uniformly Lipschitz with respect to y, z;
(H2) $\forall (y, z) \in R^1 \times R^{1\times d}$, $g(y, z, t)$ is continuous in t and $\int_0^T g^2(0, 0, t)\,\mathrm{d}t < \infty$;
(H3) $\forall (y, t) \in R^1 \times [0, T]$, $g(y, 0, t) \equiv 0$.

Denote by $M^2(0, T; R^n)$, the space of all $\mathcal{F}_t$-adapted process $x(\cdot)$ with values in R^n, such that $E\int_0^T |x(t)|^2\,\mathrm{d}t < \infty$. Let $L^2(\Omega, \mathcal{F}_T, P)$ be the set of all $\mathcal{F}_T$-measurable random variable ξ with value in R^1, such that $E|\xi|^2 < \infty$.

Under assumptions (H1)–(H3), for each $\xi \in L^2(\Omega, \mathcal{F}, P)$, the following BSDE

$$y_t = \xi + \int_t^T g(y_s, z_s, s)\,\mathrm{d}s - \int_t^T z_s\,\mathrm{d}W_s \tag{1}$$

has a unique solution $(y, z) \in M^2(0, T; R^1) \times M^2(0, T; R^{1\times d})$ (see [5]).

Peng [13] introduced the notion of g-expectation and g-probability as follows:

Definition 2.1. Suppose g satisfies (H1)–(H3). Given $\xi \in L^2(\Omega, \mathcal{F}, P)$, let (y, z) be the solution of (1). The g-expectation of ξ is defined by $\mathcal{E}_g[\xi] \equiv y_0$.

Definition 2.2. Suppose g satisfies (H1)–(H3). Given $A \in \mathcal{F}$, the g-probability of A is defined by $P_g[A] = \mathcal{E}_g[1_A]$.

We call g the generating function of the g-expectation and the g-probability. Since g-expectation or g-probability depends on the choice of the generating function g, g can be regarded as the parameter of such a nonlinear probability measure. We denote a class of g-probability measures over $(\Omega, \mathcal{F})$ by P_G where G is a given class of generating functions.

Note that g-expectation and g-probability can be defined only via BSDE. Thus, in the following, we introduce hypothesis testing in the BSDE framework. Let $G_0 \subset G$ and assume that both G_0 and $G \setminus G_0$ are nonempty. The class of g-probability measures P_G or G is called the set of admissible hypotheses, and P_{G_0} or G_0 (resp. $P_{G\setminus G_0}$ or $G \setminus G_0$) is called the null hypothesis H_0 (resp. alternative hypothesis H_1). If G_0 (resp. $G \setminus G_0$) consists of only one element, then it is called simple; otherwise composite. In this note, we shall test a simple null hypothesis H_0: $g = g^1$ versus a simple alternative H_1: $g = g^2$, where g^1 (resp. g^2) is the only element in G_0 (resp. G_1). It is worth pointing out that, in the classical case the simple hypothesis testing problem is investigated by introducing the randomized test (see [4] or [9]). In our context, we give the corresponding definition as follows:

Definition 2.3. A randomized test is an $\mathcal{F}$-measurable random variable $\xi : \Omega \to [0, 1]$. ξ is called a randomized test for the simple hypothesis H_0: $g = g^1$ versus a simple alternative H_1: $g = g^2$ with a level of significance α $(0 \leqslant \alpha \leqslant 1)$, if $\mathcal{E}_{g^1}[\xi] \leqslant \alpha$.

A randomized test ξ with significance level α can be interpreted as follows: for an observation, i.e. a sample ω from the sample space Ω, the hypothesis H_0 is rejected (respectively accepted) with probability $\xi(\omega)$ (respectively $1 - \xi(\omega)$); hence $\mathcal{E}_{g^1}[\xi] \leqslant \alpha$ guarantees that the correct hypothesis $g = g^1$ is (wrongly) rejected with an "average" of at most α. Thus, in the context of the Neyman–Pearson fundamental lemma (see [4] or [9]), $\mathcal{E}_{g^1}[\xi]$ represents Type I error for ξ, i.e. rejecting g^1 when it is true while $\mathcal{E}_{g^2}[1-\xi]$ represents Type II error for ξ, i.e., rejecting g^2 when it is true.

Set $\mathcal{L}=\{\xi \mid \xi \in L^2(\Omega, \mathcal{F}, P)$ and $0 \leqslant \xi \leqslant 1\}$. Then for a given acceptable significance level $\alpha \in (0,1)$, our problem is to find an optimal randomized test ξ^* which solves the following optimization problem:

$$\inf_{\xi \in \mathcal{L}} \mathcal{E}_{g^2}[1-\xi],$$
$$\text{subject to } \mathcal{E}_{g^1}[\xi] \leqslant \alpha. \tag{2}$$

3. Main results

We assume

(H4) g^1 and g^2 are continuously differentiable in (y, z) and their derivatives are bounded;
(H5) g^1 and g^2 are convex with respect to (y, z).

Theorem 3.1. *Suppose that assumptions* (H2)–(H5) *are satisfied. Then there exists a randomized test which attains the minimum of problem* (2).

Let ξ^* be optimal to (2) with $(y_1^*(\cdot), z_1^*(\cdot))$ and $(y_2^*(\cdot), z_2^*(\cdot))$ being the corresponding solutions of (1). Denote the derivatives $g_y^i(y_i^*(t), z_i^*(t), t)$ (resp. $g_z^i(y_i^*(t), z_i^*(t), t)$) by $g_y^i(t)$ (resp. $g_z^i(t)$) for $i=1,2$.

Theorem 3.2. *Suppose that assumptions* (H2)–(H5) *are satisfied. A randomized test ξ^* is optimal if and only if there exist a constant $v>0$ and a random variable $b \in \mathcal{L}$ such that*

$$\xi^* = 1_{\{vm(T)<n(T)\}} + b 1_{\{vm(T)=n(T)\}} \quad \textit{and} \quad \mathcal{E}_{g^1}[\xi^*]=\alpha \tag{3}$$

where $m(\cdot)$ and $n(\cdot)$ are the solutions of the following adjoint equations

$$\begin{cases} \mathrm{d}m(t) = g_y^1(t)m(t)\,\mathrm{d}t + g_z^1(t)m(t)\,\mathrm{d}W(t), \\ \mathrm{d}n(t) = g_y^2(t)n(t)\,\mathrm{d}t + g_z^2(t)n(t)\,\mathrm{d}W(t), \\ m(0)=1, \quad n(0)=1. \end{cases} \tag{4}$$

Remark. We prove the above theorem via a maximum principle for a certain stochastic control problem. The maximum principle is derived by a terminal perturbation technique which is introduced in [6] and developed in [11] and [12]. Furthermore, ξ^* can be obtained by solving the following forward backward system.

Theorem 3.3. *Suppose that* (H2)–(H5) *hold. Then there exist a positive number v and a random variable $b \in \mathcal{L}$ such that the following forward–backward system*

$$\begin{cases} \mathrm{d}m(t) = g_y^1(y_1(t), z_1(t), t)m(t)\,\mathrm{d}t + g_z^1(y_1(t), z_1(t), t)m(t)\,\mathrm{d}W(t), \\ -\mathrm{d}y_1(t) = g^1(y_1(t), z_1(t), t)\,\mathrm{d}t - z_1(t)\,\mathrm{d}W(t), \\ m(0)=1, \quad y_1(T)=\xi^*, \\ \mathrm{d}n(t) = g_y^2(y_2(t), z_2(t), t)n(t)\,\mathrm{d}t + g_z^2(y_2(t), z_2(t), t)n(t)\,\mathrm{d}W(t), \\ -\mathrm{d}y_2(t) = g^2(y_2(t), z_2(t), t)\,\mathrm{d}t - z_2(t)\,\mathrm{d}W(t), \\ n(0)=1, \quad y_2(T)=1-\xi^*, \end{cases} \tag{5}$$

with constraints $\xi^* = 1_{\{vm(T)<n(T)\}} + b 1_{\{vm(T)=n(T)\}}$ and $\mathcal{E}_{g^1}[\xi^*]=\alpha$, has a solution (y_1, z_1, y_2, z_2). Furthermore, the obtained ξ^* is an optimal randomized test.

4. Applications

Example 4.1. We study the classical Neyman–Pearson lemma (see [4] or [9]) in our context. There exists a probability measure μ and a standard d-dimensional Brownian Motion, $W^\mu(\cdot)$, defined on the complete probability space $(\Omega, \mathcal{F}, \mu)$. For two real vectors θ and ϕ, we define $m(T) = \exp\{\theta' W^\mu(T) - \frac{1}{2}T\|\theta\|^2\}$, $n(T) =$

$\exp\{\phi' W^{\mu}(T) - \frac{1}{2}T\|\phi\|^2\}$. Suppose that a probability measure Q (resp. P) is absolutely continuous with respect to μ on $\mathcal{F}$ which admits the Radon–Nikodym derivative $m(T)$ (resp. $n(T)$). The purpose is to test the 'hypothesis' Q against an 'alternative' P. In other words, we try to find a randomized test which minimizes $E_P[1-\xi]$ subject to $E_Q[\xi] \leqslant \alpha$. By Girsanov's theorem, we have that $E_Q[\xi] = E_{\mu}[m(T)\xi]$, $E_P[\xi] = E_{\mu}[n(T)\xi]$. Set the generating functions $g^1(y,z,t) = z\theta$, $g^2(y,z,t) = z\phi$. It is easy to check that $E_Q[\xi] = \mathcal{E}_{g^1}[\xi]$, $E_P[\xi] = \mathcal{E}_{g^2}[\xi]$.

By Theorem 3.2, we see that $m(\cdot)$ and $n(\cdot)$ are just the solutions of the adjoint equations and the optimal randomized test has the form $\xi^* = 1_{\{vm(T)<n(T)\}} + b1_{\{vm(T)=n(T)\}}$.

Remark. Since $n(T)/m(T)$ is the likelihood ratio in this example, we have an interpretation of the adjoint processes in our general case: $n(T)/m(T)$ can be seen as a generalization of the "likelihood ratio". In Huber and Strassen [10], they derived a generalized Radon-Nikodym derivative of a capacity with respect to another capacity. Our results show that such generalized Radon–Nikodym derivative in our context is nothing else than the ratio of the adjoint processes. It should be noted that g in general may be random (see [13]). In this note, we assume that g is a deterministic function for simplicity, and all the theorems still hold when g is a proper stochastic process.

Example 4.2. In a complete financial market, we denote a contingent claim by H which is a nonnegative random variable in $L^2(\Omega,\mathcal{F},P)$ and an investor's wealth process by $X(t)$, $0 \leqslant t \leqslant T$. Without loss of generality, we suppose that the interest rate and the risk premium process are equal to 0. Thus, the unique equivalent martingale measure $P^* = P$ such that the price of H at time 0 is $H_0 = E_{P^*}[H]$. But if the seller's initial capital $\tilde{X}_0$ is smaller than H_0, then he can not perfectly hedge H and the shortfall is $-(H - X_T)^+$. More precisely, the readers may refer to [7]. For more references about optimal portfolio with constraints, the readers may see [1] and the references therein.

In this example, we introduce, as in [8], the convex risk measure $\rho_g : L^2(\Omega,\mathcal{F},P) \to R$ by $\rho_g(X) = \mathcal{E}_g[-X]$ where $X \in L^2(\Omega,\mathcal{F},P)$ is a financial position and the generator g is convex in (y,z). We suppose that the seller selects his portfolio to minimize the shortfall risk $\rho_g(-(H - X_T)^+)$.

In summary, we try to minimize $\rho_g(-(H - X_T)^+)$ subject to $X_0 \leqslant \tilde{X}_0$. Note that $X_0 = E_{P^*}[X_T]$ and the optimal X_T must satisfy $0 \leqslant X_T \leqslant H$ by the comparison theorem of BSDEs. Then the above problem is equivalent to $\inf_{0 \leqslant X_T \leqslant H} \mathcal{E}_g(H - X_T)$, subject to $E_{P^*}[X_T] \leqslant \tilde{X}_0$.

Note that E_{P^*} can be regarded as a trivial g-expectation. Similar analysis as in Theorem 3.2 shows that there exist a positive number v and a random variable b $(0 \leqslant b \leqslant H)$ such that the optimal terminal wealth X_T^* satisfies$X_T^* = H1_{\{v<n(T)\}} + b1_{\{v=n(T)\}}$ and $E_{P^*}[X_T^*] = \tilde{X}_0$, where $n(\cdot)$ is the solution of the corresponding adjoint equation.

Acknowledgements

The authors would like to thank the referee for his/her comments, and S. Peng, Z. Chen and M. Xu for their helpful discussions.

References

[1] P. Boyle, W. Tian, Portfolio management with constraints, Math. Finance 17 (2007) 319–343.
[2] Z. Chen, L. Epstein, Ambiguity, risk and asset returns in continuous time, Econometrica 70 (2002) 1403–1443.
[3] Z. Chen, R. Kulperger, Minimax pricing and Choquet pricing, Insurance: Mathematics and Economics 38 (2006) 518–528.
[4] J. Cvitanic, I. Karatzas, Generalized Neyman–Pearson lemma via convex duality, Bernoulli 7 (2001) 79–97.
[5] N. El Karoui, S. Peng, M.C. Quenez, Backward stochastic differential equations in finance, Math. Finance 7 (1997) 1–71.
[6] N. El Karoui, S. Peng, M.-C. Quenez, A dynamic maximum principle for the optimization of recursive utilities under constraints, Ann. Appl. Probab. 11 (2001) 664–693.
[7] H. Föllmer, P. Leukert, Efficient hedging: cost versus shortfall risk, Finance Stochast. 4 (2000) 117–146.
[8] E. Gianin, Risk measures via g-expectations, Insurance: Mathematics and Economics (2006) 19–34.
[9] S. Helmut, Mathematical Theory of Statistics, Walter de Gruyter & Co., Berlin, 1985.
[10] P. Huber, V. Strassen, Minimax tests and the Neyman–Pearson lemma for capacities, Ann. Statist. 1 (1973) 251–263.
[11] S. Ji, S. Peng, Terminal perturbation method for the backward approach to continuous-time mean-variance portfolio selection, Stochastic Process. Appl. (2008), 10.1016/j.spa.2007.07.005, in press.
[12] S. Ji, X. Zhou, A maximum principle for stochastic optimal control with terminal state constraints, and its applications, Commun. Inf. Systems 6 (2006) 321–337 (a special issue dedicated to Tyrone Duncan on the occasion of his 65th birthday).
[13] S. Peng, Backward stochastic differential equations and related g-expectation, in: N. El Karoui, L. Mazliak (Eds.), Backward Stochastic Differential Equations, in: Pitman Res. Notes Math. Ser., vol. 364, 1997, pp. 141–159.