Теорія Ймовір. та Матем. Статист.
Вип. 99, 2018

Theor. Probability and Math. Statist.
No. 99, 2019, Pages 199–210
https://doi.org/10.1090/tpms/1090
Article electronically published on February 27, 2020

EVALUATION OF EXTREME VALUES OF ENTROPY FUNCTIONALS

UDC 519.21

YU. S. MISHURA AND H. S. ZHELEZNIAK

Abstract. We consider the sum of two independent Wiener processes with a drift and construct a family of probability measures such that the drift with respect to each of them is zero. Among these measures, we search for those that minimize or maximize certain functionals, for example, entropy-type functionals.

1. Introduction and setting of the problem

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathsf{P})$ be a probability space equipped with a filtration. All objects considered below are assumed to be defined on this space.

Fix $T > 0$ and let P_1 be another probability measure on $(\Omega, \mathcal{F})$ such that the measure P_1 is absolutely continuous with respect to the measure P. By definition [1, pp. 121–130], if P_1 is absolutely continuous with respect to P, the entropy of the probability measure P_1 with respect to P is defined by

$$\mathsf{H}(\mathsf{P}_1 \mid \mathsf{P}) := \mathsf{E}\left[\frac{d\,\mathsf{P}_1}{d\,\mathsf{P}} \log \frac{d\,\mathsf{P}_1}{d\,\mathsf{P}}\right].$$

The main feature of the functional $\mathsf{H}(\mathsf{P}_1 \,|\, \mathsf{P})$ is that $\mathsf{H}(\mathsf{P}_1 \,|\, \mathsf{P}) \geq 0$ and $\mathsf{H}(\mathsf{P}_1 \,|\, \mathsf{P}) = 0$ if and only if $\mathsf{P}_1 = \mathsf{P}$.

Indeed, the function $h(x) = x \log x$ is strictly convex on $(0, \infty)$. Applying the Jensen inequality to the function $h(x)$, we obtain

$$\mathsf{H}(\mathsf{P}_1 \mid \mathsf{P}) = \mathsf{E}\left[h\left(\frac{d\,\mathsf{P}_1}{d\,\mathsf{P}}\right)\right] \geq h\left(\mathsf{E}\left[\frac{d\,\mathsf{P}_1}{d\,\mathsf{P}}\right]\right) \geq h(1) = 0.$$

This inequality becomes an equality if and only if $\mathsf{P}_1 = \mathsf{P}$.

In addition to the standard entropy, one can consider other functionals defined for densities of probability measures. For example, the functionals we are interested in appear when solving the following problem.

Assume that two independent Wiener processes

$$\mathrm{W}_1 = \{\mathrm{W}_1(t), t \in [0,T]\} \quad \text{and} \quad \mathrm{W}_2 = \{\mathrm{W}_2(t), t \in [0,T]\}$$

are defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathsf{P})$ and let $f \in L_2([0,T], \lambda_1)$ be a nonrandom nonnegative function where λ_1 denotes the Lebesgue measure in the real line. Consider the sum of these Wiener processes with a drift, that is, we consider the stochastic process of the following form:

$$S(t) = \mathrm{W}_1(t) + \mathrm{W}_2(t) + \int_0^t f(s)\,ds, \qquad t \in [0,T]. \tag{1}$$

2010 *Mathematics Subject Classification.* Primary 60G22, 60J65; Secondary 94A17.

Key words and phrases. Wiener process, Radon–Nikodym derivative, entropy functional, minimization, maximization.

Despite the sum of two independent Wiener processes is again a Wiener process multiplied by a constant, we prefer to view $S(t)$ as a sum of $W_1(t)$ and $W_2(t)$. Consider another probability measure $\widetilde{\mathsf{P}}$ such that the sum $S(t)$ in (1) is of the form

$$\widetilde{S}(t) = \widetilde{W}_1(t) + \widetilde{W}_2(t), \qquad t \in [0,T], \tag{2}$$

where $\widetilde{W}_1$ and $\widetilde{W}_2$ are two independent Wiener processes with respect to the measure $\widetilde{\mathsf{P}}$. Since the Wiener processes $\widetilde{W}_1$ and $\widetilde{W}_2$ are independent, the likelihood ratio splits into the product of two factors,

$$\frac{d\widetilde{\mathsf{P}}}{d\,\mathsf{P}} = \frac{d\,\mathsf{P}_1}{d\,\mathsf{P}} \times \frac{d\,\mathsf{P}_2}{d\,\mathsf{P}},$$

where the measures P_1 and P_2 are generated by the Wiener processes $\widetilde{W}_1$ and $\widetilde{W}_2$, respectively. In order to remove the drift $\int_0^t f(s)\,ds$, we represent the sum of the processes W_1 and W_2 with a drift as follows:

$$\widetilde{W}_1(t) = \mathrm{W}_1(t) + \int_0^t f_1(s)\,ds,$$
$$\widetilde{W}_2(t) = \mathrm{W}_2(t) + \int_0^t f_2(s)\,ds,$$

where

$$f_1(t) + f_2(t) = f(t), \qquad t \in [0,T],$$

and the functions f_i are such that

$$f_i \in L_2([0,T],\lambda_1), \qquad 0 \le f_i(t) \le f(t), \quad t \in [0,T].$$

Then

$$\frac{d\,\mathsf{P}_i}{d\,\mathsf{P}} = \exp\left\{\int_0^T f_i(s)\,dW_i(s) - \frac{1}{2}\int_0^T f_i^2(s)\,ds\right\}, \qquad i = 1,2,$$

by Girsanov's theorem. Our aim is to find the functions $f_1(t)$ and $f_2(t)$ such that they give the minimum or maximum to the functionals

$$\mathsf{E}\left[F\left(\frac{d\mathsf{P}_1}{d\,\mathsf{P}}, \frac{d\mathsf{P}_2}{d\,\mathsf{P}}\right)\right],$$

where $F(\cdot,\cdot)\colon \mathbb{R}^2 \to \mathbb{R}_+$ is a certain nonnegative measurable function. Another approach to the minimization of entropy functionals is presented in the paper [2]. When solving the problem, we consider two cases, where the function F

1) is differentiable and the standard Itô formula is applicable, and
2) defines an entropy-type functional containing the logarithms.

The paper is organized as follows. The theorem on the points of maximum and minimum is stated in Section 2 for a functional that depends on three probability measures and that is generated by a twice differentiable function. Extreme points are found in Section 3 for the entropy functionals that depend on three probability measures and that contain the logarithms. Concluding remarks are collected in Section 4.

2. Extreme values of the functionals for differentiable functions

We consider the case where the function F belongs to the space $C^2(\mathbb{R}^2)$ and the order of growth of it and those of its derivatives are at most power. Consider the following nonnegative martingales:

$$M_i(t) = \exp\left\{\int_0^t f_i(s)\,dW_i(s) - 1/2\int_0^t f_i^2(s)\,ds\right\}, \qquad i = 1,2.$$

It is obvious that $\frac{d\mathsf{P}_i}{d\,\mathsf{P}} = M_i(T)$, $i = 1,2$.

Lemma 2.1. *Let $F \in C^2(\mathbb{R}^2)$ and let there exist constants $k \in \mathbb{N}$ and $C > 0$ such that*

$$|F(x_1,x_2)|, \left|\frac{\partial F(x_1,x_2)}{\partial x_i}\right|, \left|\frac{\partial^2 F(x_1,x_2)}{\partial x_i^2}\right| \le C\|x\|^k, \qquad i = 1,2,$$

where $\|x\|^2 = x_1^2 + x_2^2$. Then

$$\mathsf{E}\left[F\left(\frac{d\mathsf{P}_1}{d\,\mathsf{P}}, \frac{d\mathsf{P}_2}{d\,\mathsf{P}}\right)\right] = F(1,1) + \frac{1}{2}\Bigg(\int_0^T \Bigg(\mathsf{E}\left(\frac{\partial^2 F(M_1(s),M_2(s))}{\partial x_1^2}\right) f_1^2(s) + \mathsf{E}\left(\frac{\partial^2 F(M_1(s),M_2(s))}{\partial x_2^2}\right) f_2^2(s)\Bigg) ds\Bigg). \tag{3}$$

Proof. We apply the standard Itô formula to the function F and martingales M_i, $i = 1,2$. Note that all stochastic integrals below exist and have finite moments of all orders in view of the assumption on the growth order. Hence

$$\begin{aligned} F\left(\frac{d\mathsf{P}_1}{d\,\mathsf{P}}, \frac{d\mathsf{P}_2}{d\,\mathsf{P}}\right) &= F(M_1(T), M_2(T)) \\ &= F(1,1) + \sum_{i=1,2}\int_0^T \frac{\partial F(M_1(s),M_2(s))}{\partial x_i} f_i(s)\,dW_i(s) \\ &\quad + \frac{1}{2}\sum_{i=1,2}\int_0^T \frac{\partial^2 F(M_1(s),M_2(s))}{\partial x_i^2} f_i^2(s)\,ds, \end{aligned}$$

whence the proof follows. ☐

It is reasonable to reduce the problem of evaluation of extreme values of the functional

$$\mathsf{E}\left[F\left(\frac{d\mathsf{P}_1}{d\,\mathsf{P}}\frac{d\mathsf{P}_2}{d\,\mathsf{P}}\right)\right]$$

to the case where the expectation on the right hand side of (3) can be evaluated explicitly. This, for example, can be done for the functional

$$F(x_1,x_2) = x_1^{k+2} + x_2^{l+2},$$

where k and l are integer nonnegative numbers. In this case,

$$\begin{aligned} \mathsf{E}\left(\frac{\partial^2 F(M_1(s),M_2(s))}{\partial x_1^2}\right) &= (k+2)(k+1)\mathsf{E}(M_1(s))^k \\ &= (k+2)(k+1)\mathsf{E}\exp\left\{k\int_0^s f_1(u)\,dW_1(u) - \frac{1}{2}k\int_0^s f_1^2(u)\,du\right\} \\ &= (k+2)(k+1)\exp\left\{\frac{1}{2}(k^2-k)\int_0^s f_1^2(u)\,du\right\} \end{aligned} \tag{4}$$

and similarly

$$\begin{aligned} \mathsf{E}\left(\frac{\partial^2 F(M_1(s),M_2(s))}{\partial x_2^2}\right) &= (l+2)(l+1)\mathsf{E}(M_2(s))^l \\ &= (l+2)(l+1)\mathsf{E}\exp\left\{l\int_0^s f_2(u)\,dW_2(u) - \frac{1}{2}l\int_0^s f_2^2(u)\,du\right\} \\ &= (l+2)(l+1)\exp\left\{\frac{1}{2}(l^2-l)\int_0^s f_2^2(u)\,du\right\}. \end{aligned} \tag{5}$$

Taking into account equalities (3)–(5), the problem reduces to the optimization (either maximization or minimization) problem for the functional

$$
\begin{aligned}
G &:= \int_0^T \left[\mathsf{E}\left(\frac{\partial^2 F(M_1(s), M_2(s))}{\partial x_1^2}\right) f_1^2(s) + \mathsf{E}\left(\frac{\partial^2 F(M_1(s), M_2(s))}{\partial x_2^2}\right) f_2^2(s)\right] ds \\
&= (k+2)(k+1) \int_0^T \exp\left\{\frac{1}{2}\left(k^2 - k\right) \int_0^s f_1^2(u)\,du\right\} f_1^2(s)\,ds \\
&\quad + (l+2)(l+1) \int_0^T \exp\left\{\frac{1}{2}\left(l^2 - l\right) \int_0^s f_2^2(u)\,du\right\} f_2^2(s)\,ds \\
&= k_1 \left(\exp\left\{\frac{1}{2}\left(k^2 - k\right) \int_0^T f_1^2(u)\,du\right\} - 1\right) \\
&\quad + l_1 \left(\exp\left\{\frac{1}{2}\left(l^2 - l\right) \int_0^T f_2^2(u)\,du\right\} - 1\right),
\end{aligned}
\tag{6}
$$

where

$$
k_1 = \frac{2(k+2)(k+1)}{k^2 - k} \quad \text{and} \quad l_1 = \frac{2(l+2)(l+1)}{l^2 - l}.
$$

In turn, it is reasonable to rewrite equality (6) as follows. Let $f_1(t) = \alpha(t) f(t)$, where $\alpha(t) \in [0,1]$ is a measurable function. Then $f_2(t) = (1 - \alpha(t)) f(t)$ and problem (6) reduces to the following one: find extreme values of the functional

$$
G(\alpha) := k_1 \exp\left\{k_2 \int_0^T \alpha^2(u) f^2(u)\,du\right\} + l_1 \exp\left\{l_2 \int_0^T (1 - \alpha(u))^2 f^2(u)\,du\right\},
\tag{7}
$$

where $k_2 = \frac{1}{2}(k^2 - k)$, $l_2 = \frac{1}{2}(l^2 - l)$. Without loss of generality we suppose that $k > l$.

Theorem 2.1. 1) *There exists a unique function $\alpha(s) = \alpha_0$ that minimizes the functional $G(\alpha)$; the function α_0 is such that*

$$
\alpha_0 = \frac{l_1 l_2 \exp\left\{l_2 c (1 - \alpha_0)^2\right\}}{k_1 k_2 \exp\left\{k_2 c \alpha_0^2\right\} + l_1 l_2 \exp\left\{l_2 c (1 - \alpha_0)^2\right\}},
$$

where $c = \|f\|^2_{L_2([0,T],\lambda_1)}$.

2) $\alpha_0 = \frac{1}{2}$ *for $k = l$.*

3) *The maximum of $G(\alpha)$ equals* $\max\left\{k_1 + l_1 \exp\left\{l_2 c\right\}, k_1 \exp\left\{k_2 c\right\} + l_1\right\}$.

Proof. We prove the first statement. First we substitute $\alpha(u) + \varepsilon\beta(u)$ instead of $\alpha(u)$ in (7), where $\varepsilon \in \mathbb{R}$ and β is an arbitrary bounded measurable function. In other words, we consider the functional $G(\alpha, \varepsilon) := G(\alpha + \varepsilon\beta)$. Put

$$
\begin{aligned}
G_1(\alpha, \varepsilon) &= \exp\left\{k_2 \int_0^T \left(\alpha(u) + \varepsilon\beta(u)\right)^2 f^2(u)\,du\right\}, \\
G_2(\alpha, \varepsilon) &= \exp\left\{l_2 \int_0^T \left(1 - \alpha(u) - \varepsilon\beta(u)\right)^2 f^2(u)\,du\right\}.
\end{aligned}
$$

Then $G(\alpha,\varepsilon) = k_1G_1(\alpha,\varepsilon) + l_1G_2(\alpha,\varepsilon)$. Now we evaluate the derivative with respect to ε:

$$\begin{aligned} G'_\varepsilon(\alpha,\varepsilon) &= k_1G'_1(\alpha,\varepsilon) + l_1G'_2(\alpha,\varepsilon) \\ &= 2k_1k_2G_1(\alpha,\varepsilon)\int_0^T (\alpha(u)+\varepsilon\beta(u))\,\beta(u)f^2(u)\,du \\ &\quad - 2l_1l_2G_2(\alpha,\varepsilon)\int_0^T (1-\alpha(u)-\varepsilon\beta(u))\,\beta(u)f^2(u)\,du. \end{aligned} \tag{8}$$

Since we are searching for the points of minimum, the inequality $G(\alpha,0)\le G(\alpha,\varepsilon)$ must hold for all $\varepsilon\in\mathbb{R}$. Thus we equate the derivative at zero in (8) to zero,

$$k_1k_2G_1(\alpha,0)\int_0^T \alpha(u)\beta(u)f^2(u)\,du - l_1l_2G_2(\alpha,0)\int_0^T (1-\alpha(u))\,\beta(u)f^2(u)\,du = 0,$$

or

$$\int_0^T [k_1k_2G_1(\alpha,0)\alpha(u) - l_1l_2G_2(\alpha,0) + l_1l_2G_2(\alpha,0)\alpha(u)]\,\beta(u)f^2(u)\,du = 0$$

for any bounded function β.

Putting $\beta(u) = (k_1k_2G_1(\alpha_1,0) + l_1l_2G_2(\alpha,0))\,\alpha(u) - l_1l_2G_2(\alpha,0)$ we get the following equation for $\alpha(u)$:

$$\alpha(u) = \frac{l_1l_2G_2(\alpha,0)}{k_1k_2G_1(\alpha,0)+l_1l_2G_2(\alpha,0)}. \tag{9}$$

Equation (9) means that $\alpha(u)=\alpha\in(0,1)$, where α is a constant such that

$$\alpha = \frac{l_1l_2\exp\{l_2(1-\alpha)^2c\}}{k_1k_2\exp\{k_2\alpha^2c\}+l_1l_2\exp\{l_2(1-\alpha)^2c\}}. \tag{10}$$

Here $c := \|f\|^2_{L_2([0,T],\lambda_1)}$, $k_1k_2=(k+2)(k+1)$, and $l_1l_2=(l+2)(l+1)$. The left hand side of (10) increases from 0 to 1. We simplify the right hand side of (10) to the form

$$\frac{1}{r\exp\{(k_2\alpha^2-l_2(1-\alpha)^2)\,c\}+1},\quad \text{where } r=\frac{k_1k_2}{l_1l_2}.$$

Next we investigate the behavior of the function $\varphi(\alpha) = \exp\{c\,(k_2-l_2)\,\alpha^2+2l_2c\alpha\}$. Since $k>l$, we have $k_2>l_2$ and thus the function $\varphi(\alpha)$ increases with respect to α. Hence the above fraction decreases from

$$\frac{l_1l_2\exp\{l_2c\}}{k_1k_2+l_1l_2\exp\{l_2c\}}$$

to

$$\frac{l_1l_2}{k_1k_2\exp\{k_2c\}+l_1l_2}.$$

Note that both these values belong to the interval $(0,1)$. Since the functions on the left and right hand sides of (10) are continuous, equation (10) has a unique solution. Denote this solution by α_0. Now we evaluate the second derivative $G''_{\varepsilon\varepsilon}(\alpha,\varepsilon)$ and find its sign at $\varepsilon=0$.

It is obvious that $G''_{\varepsilon\varepsilon}(\alpha,\varepsilon) = k_1 G''_1(\alpha,\varepsilon) + l_1 G''_2(\alpha,\varepsilon)$. Hence it is sufficient to investigate the sign of each term. For example, we obtain for $G''_1(\alpha,\varepsilon)$ that

$$\begin{aligned} G''_1(\alpha,\varepsilon)\Big|_{\varepsilon=0} &= \Bigg[2k_2 G'_1(\alpha,\varepsilon)\int_0^T (\alpha(u)+\varepsilon\beta(u))\,\beta(u)f^2(u)\,du \\ &\qquad + 2k_2 G_1(\alpha,\varepsilon)\int_0^T \beta^2(u)f^2(u)\,du\Bigg]\Bigg|_{\varepsilon=0} \\ &= 4k_2^2 G_1(\alpha,0)\left(\int_0^T \alpha(u)\beta(u)f^2(u)\,du\right)^2 \\ &\quad + 2k_2 G_1(\alpha,0)\int_0^T \beta^2(u)f^2(u)\,du > 0 \end{aligned} \tag{11}$$

by (8). Analogously,

$$\begin{aligned} G''_2(\alpha,\varepsilon)|_{\varepsilon=0} &= 4l_2^2 G_2(\alpha,0)\left(\int_0^T (1-\alpha(u))\beta(u)f^2(u)\,du\right)^2 \\ &\quad + 2l_2 G_2(\alpha,0)\int_0^T \beta^2(u)f^2(u)\,du > 0. \end{aligned} \tag{12}$$

Therefore the minimum of the functional $G(\alpha)$ is indeed attained at the point α_0, whence statement 1) follows. Statements 2) and 3) are obvious. □

3. Extreme values of entropy-type functionals

As mentioned above, the entropy of a probability measure with respect to another one is a nonnegative functional of the Radon–Nikodym derivative. Note that the Radon–Nikodym derivative equals zero if and only if the measures coincide. Such kind of functionals are said to be *entropy-type functionals*.

Turning back to the problem of removing the shift for the sum of two independent Wiener processes, we are going to find the measures P_1 and P_2 for which the sum of two entropy-type functionals

$$\mathsf{H}(\mathsf{P},\mathsf{P}_1,\mathsf{P}_2) = \mathsf{E}\left[\left(\frac{d\mathsf{P}_1}{d\mathsf{P}}\right)^2\left(-\log\frac{d\mathsf{P}_2}{d\mathsf{P}}\right)\right] + \mathsf{E}\left[\left(\frac{d\mathsf{P}_2}{d\mathsf{P}}\right)^2\left(-\log\frac{d\mathsf{P}_1}{d\mathsf{P}}\right)\right] \tag{13}$$

attains its extreme value, either minimum or maximum. Our interest to functionals (13) is two fold. First, we are interested in studying a functional whose points of minimum are not trivial. Second, it is desirable that the points of minimum of a functional can be evaluated explicitly.

Some properties of the functional $\mathsf{H}(\mathsf{P},\mathsf{P}_1,\mathsf{P}_2)$ that explain the name "entropy-type functional" are collected in Lemma 3.1.

Lemma 3.1. *The functional* $\mathsf{H}(\mathsf{P},\mathsf{P}_1,\mathsf{P}_2)$ *defined by equality* (13) *is nonnegative, that is,* $\mathsf{H}(\mathsf{P},\mathsf{P}_1,\mathsf{P}_2) \ge 0$. *Moreover,* $\mathsf{H}(\mathsf{P},\mathsf{P}_1,\mathsf{P}_2) = 0$ *if and only if* $\mathsf{P}_1 = \mathsf{P}$ *and* $\mathsf{P}_2 = \mathsf{P}$.

Proof. The functional $\mathsf{H}(\mathsf{P},\mathsf{P}_1,\mathsf{P}_2)$ is the sum of two terms

$$\mathsf{E}\left[\left(\frac{d\mathsf{P}_1}{d\mathsf{P}}\right)^2\left(-\log\frac{d\mathsf{P}_2}{d\mathsf{P}}\right)\right] \quad\text{and}\quad \mathsf{E}\left[\left(\frac{d\mathsf{P}_2}{d\mathsf{P}}\right)^2\left(-\log\frac{d\mathsf{P}_1}{d\mathsf{P}}\right)\right].$$

Further,

$$\mathsf{E}\left[\left(\frac{d\mathsf{P}_1}{d\,\mathsf{P}}\right)^2\left(-\log\frac{d\mathsf{P}_2}{d\,\mathsf{P}}\right)\right]=\mathsf{E}\left[\left(\frac{d\mathsf{P}_1}{d\,\mathsf{P}}\right)\right]^2\left(-\mathsf{E}\log\frac{d\mathsf{P}_2}{d\,\mathsf{P}}\right).$$

Since $\log x$ is a concave function,

$$\mathsf{E}\log\frac{d\mathsf{P}_2}{d\,\mathsf{P}}\le\log\mathsf{E}\frac{d\mathsf{P}_2}{d\,\mathsf{P}}=\log 1=0,$$

whence

$$\mathsf{E}\left[\left(\frac{d\mathsf{P}_1}{d\,\mathsf{P}}\right)^2\left(-\log\frac{d\mathsf{P}_2}{d\,\mathsf{P}}\right)\right]\ge 0.$$

The second term has analogous properties. □

Since the Wiener processes are independent, the functional $\mathsf{H}(\mathsf{P},\mathsf{P}_1,\mathsf{P}_2)$ can be rewritten with the help of the Itô formula as follows:

$$\mathsf{H}(\mathsf{P},\mathsf{P}_1,\mathsf{P}_2)=\frac{1}{2}\left(\exp\left\{\int_0^T f_1^2(t)\,dt\right\}\int_0^T f_2^2(t)\,dt+\exp\left\{\int_0^T f_2^2(t)\,dt\right\}\int_0^T f_1^2(t)\,dt\right).$$

The minimal and maximal values of the functional (13) are found in the paper [3] for the case where $f_1(t)=\alpha f(t)$ and $f_2(t)=(1-\alpha)f(t)$. Therefore the problem in [3] reduces to searching a point α at which the minimal value is attained. Now our aim is to search for extreme values of the same functional but in a wider class of functions, namely

$$f_1(t)=\alpha(t)f(t)\quad\text{and}\quad f_2(t)=(1-\alpha(t))f(t), \tag{14}$$

where $\alpha(t)\in[0,1]$ is a bounded measurable function. Recall the notation

$$c=\|f\|_{L_2([0,T],\lambda_1)}.$$

Theorem 3.1. 1) *If $c\le 2$, then the functional $\mathsf{H}(\mathsf{P},\mathsf{P}_1,\mathsf{P}_2)$ attains its minimal value $(c/4)e^{c/4}$ if*

$$f_1(t)=f_2(t)=\frac{f(t)}{2};$$

correspondingly, the functional attains its maximal value $c/2$ if (for example) $f_1(t)=1$.

2) *If $2<c<4$, then the functional attains its minimal value $\frac{1}{2}(1-x_1)\exp\{cx_1^2\}$ if $f_1(t)=x_1f(t)$ and $f_2(t)=(1-x_1)f(t)$, where x_1 is a unique root of the equation*

$$x\exp\{c(1-2x)\}+x-1=0$$

that belongs to the interval $(0,1/c)$. In view of the symmetry, the functional attains its minimal value if $f_1(t)=(1-x_1)f(t)$ and $f_2(t)=x_1f(t)$, as well. Correspondingly, the functional attains its maximal value $c/2$ if $2<c<\ln 16$ or $(c/4)e^{c/4}$ if $\ln 16\le c<4$.

3) *If $c\ge 4$, then, like the preceding case, the functional attains its minimal value if $f_1(t)=x_1f(t)$ and $f_2(t)=(1-x_1)f(t)$, where x_1 is a unique root of the equation*

$$x\exp\{c(1-2x)\}+x-1=0$$

that belongs to the interval $(0,1/c)$. The maximal value of the functional equals $(c/4)e^{c/4}$ if $c=4$ or

$$\frac{1}{4}\exp\left\{\frac{(c-2)-\sqrt{c(c-4)}}{2}\right\}$$
$$\times\left((c-2)-\sqrt{c(c-4)}-\left(\sqrt{c(c-4)}-(c-2)\right)\exp\left\{\sqrt{c(c-4)}\right\}\right)$$

if $c > 4$ *and* $f_1(t) = x_{4,5} f(t)$ *and* $f_2(t) = (1 - x_{4,5}) f(t)$, *where* $x_{4,5}$ *are the roots of the quadratic equation* $1 - xc + x^2 c = 0$ (*the roots are equal to*

$$\frac{1}{2} \pm \frac{1}{2}\sqrt{1 - \frac{4}{c}},$$

respectively).

Proof. Using equality (14), we rewrite the functional $\mathsf{H}(\mathsf{P}, \mathsf{P}_1, \mathsf{P}_2)$ in the following form:

$$\begin{aligned} \mathsf{H}(\mathsf{P}, \mathsf{P}_1, \mathsf{P}_2) = \frac{1}{2} \exp\left\{ \int_0^T f^2(t) \alpha^2(t)\, dt \right\} \int_0^T f^2(t)(1 - \alpha(t))^2\, dt \\ + \frac{1}{2} \exp\left\{ \int_0^T f^2(t)(1 - \alpha(t))^2\, dt \right\} \int_0^T f^2(t) \alpha^2(t)\, dt. \end{aligned}$$

As in the proof of Theorem 2.1, consider the functional H at $\alpha(t) + \varepsilon\beta(t)$, where $\beta \in L_2[0, T]$ is an arbitrary function and $\varepsilon \in \mathbb{R}$:

$$\begin{aligned} \mathsf{H}(\alpha + \varepsilon\beta) = \frac{1}{2} \exp\left\{ \int_0^T f^2(t)(\alpha(t) + \varepsilon\beta(t))^2\, dt \right\} \\ \times \int_0^T f^2(t)\big(1 - (\alpha(t) + \varepsilon\beta(t))\big)^2\, dt \\ + \frac{1}{2} \exp\left\{ \int_0^T f^2(t)\big(1 - (\alpha(t) + \varepsilon\beta(t))\big)^2\, dt \right\} \\ \times \int_0^T f^2(t)(\alpha(t) + \varepsilon\beta(t))^2\, dt. \end{aligned} \tag{15}$$

Put

$$I_1 = \exp\left\{ \int_0^T f^2(t)\alpha^2(t)\, dt \right\} \quad \text{and} \quad I_2 = \exp\left\{ \int_0^T f^2(t)(1 - \alpha(t)^2)\, dt \right\}.$$

If we are searching for a function α that minimizes the right hand side of (15), then obviously

$$\mathsf{H}(\alpha) \le \mathsf{H}(\alpha + \varepsilon\beta)$$

for all $\varepsilon \in \mathbb{R}$ and all $\beta \in L_2[0, T]$, whence we conclude that $\mathsf{H}(\alpha + \varepsilon\beta)$ attains its minimal value at $\varepsilon = 0$. Then we evaluate the first derivative of the functional $\mathsf{H}(\alpha + \varepsilon\beta)$ with respect to ε:

$$\begin{aligned} \mathsf{H}'(\alpha + \varepsilon\beta)\big|_{\varepsilon=0} = I_1 \int_0^T f^2(t)\alpha(t)\beta(t)\, dt \int_0^T f^2(t)(1 - \alpha(t))^2\, dt \\ - I_1 \int_0^T f^2(t)(1 - \alpha(t))\beta(t)\, dt \\ - I_2 \int_0^T f^2(t)(1 - \alpha(t))\beta(t)\, dt \int_0^T f^2(t)\alpha^2(t)\, dt \\ + I_2 \int_0^T f^2(t)\alpha(t)\beta(t)\, dt. \end{aligned} \tag{16}$$

Simplifying the right hand side of equality (16) and then equating it to zero we obtain the following equation:

$$
\begin{aligned}
(17)\qquad & I_1\left[\int_0^T f^2(t)\alpha(t)\beta(t)\,dt\int_0^T f^2(t)(1-\alpha(t))^2\,dt-\int_0^T f^2(t)(1-\alpha(t))\beta(t)\,dt\right]\\
& -I_2\left[\int_0^T f^2(t)(1-\alpha(t))\beta(t)\,dt\int_0^T f^2(t)\alpha^2(t)\,dt-\int_0^T f^2(t)\alpha(t)\beta(t)\,dt\right]=0.
\end{aligned}
$$

If $K=I_1/I_2$, then the left hand side of (17) can be rewritten in the following form:

$$
K:=\frac{\int_0^T f^2(t)(1-\alpha(t))\beta(t)\,dt\int_0^T f^2(t)\alpha^2(t)\,dt-\int_0^T f^2(t)\alpha(t)\beta(t)\,dt}{\int_0^T f^2(t)\alpha(t)\beta(t)\,dt\int_0^T f^2(t)(1-\alpha(t))^2\,dt-\int_0^T f^2(t)(1-\alpha(t))\beta(t)\,dt}
$$

or, which is the same,

$$
\begin{aligned}
& \int_0^T f^2(t)(1-\alpha(t))\beta(t)\,dt\int_0^T f^2(t)\alpha^2(t)\,dt-\int_0^T f^2(t)\alpha(t)\beta(t)\,dt\\
(18)\qquad & \quad =K\Bigg(\int_0^T f^2(t)\alpha(t)\beta(t)\,dt\int_0^T f^2(t)(1-\alpha(t))^2\,dt\\
& \qquad\qquad -\int_0^T f^2(t)(1-\alpha(t))\beta(t)\,dt\Bigg).
\end{aligned}
$$

Rearranging the terms in (18) to separate the factors containing $\beta(t)$, we obtain

$$
\begin{aligned}
& \left[\int_0^T f^2(t)\alpha^2(t)\,dt+K\right]\int_0^T f^2(t)\beta(t)(1-\alpha(t))\,dt\\
(19)\qquad & \quad =\left[1+K\int_0^T f^2(t)(1-\alpha(t))^2\,dt\right]\int_0^T f^2(t)\alpha(t)\beta(t)\,dt.
\end{aligned}
$$

If

$$
L:=\frac{\int_0^T f^2(t)\alpha^2(t)\,dt+K}{1+K\int_0^T f^2(t)(1-\alpha(t))^2\,dt},
$$

then equality (19) reduces to

$$
(20)\qquad \int_0^T f^2(t)\beta(t)[\alpha(t)-L(1-\alpha(t))\,dt=0.
$$

Since equality (20) holds for an arbitrary bounded function $\beta=\beta(t)$, one can put

$$
\beta(t)=(1+L)\,\alpha(t)-L.
$$

Then equation (20) reduces to

$$
(21)\qquad \int_0^T f^2(t)(\alpha(t)-L(1-\alpha(t)))^2\,dt=0,
$$

whence

$$
\alpha(t)=\alpha:=\frac{L}{1+L},
$$

that is,

$$
L=\frac{\alpha}{1-\alpha}.
$$

If α is a constant, then

$$K = \exp\{c(2\alpha - 1)\}, \qquad L = \frac{\alpha}{1-\alpha} = \frac{\alpha^2 c + \exp\{c(2\alpha-1)\}}{1 + \exp\{c(2\alpha-1)\}c(1-\alpha)^2},$$

and the latter equation is equivalent to

$$\left(\alpha^2 c - \alpha c + 1\right)\left((1-\alpha)\exp\{c(2\alpha-1)\} - \alpha\right) = 0. \tag{22}$$

It is proved in the paper [3] that the left hand side of equation (22) equals the derivative of the entropy functional up to the factor $e^{c\alpha^2}$ if the functional is considered only for constants $\alpha \in [0,1]$. Since we proved already that the extreme values of the entropy functional are attained if α is constant, we need

(a) to find the roots of the equation (22),
(b) to decide which of the roots correspond to the maximal values and which to minimal values,
(c) and then to compare the corresponding values with the values of the entropy functional for $\alpha = 0$ and $\alpha = 1$.

If $x = 1 - \alpha$, then the first factor on the left hand side of (22) does not change, while the second one is equal to $x\exp\{c(1-2x)\} + x - 1$. Then the functional is as follows:

$$\mathsf{H}(\mathsf{P}, \mathsf{P}_1, \mathsf{P}_2) = \frac{c}{2}\left((1-x)^2 \exp\left\{cx^2\right\} + x^2 \exp\left\{c(1-x)^2\right\}\right). \tag{23}$$

The values of the functional at the points 0 and 1/2 equal $c/2$ and $(c/4)e^{c/4}$, respectively.

The equation

$$x\exp\left\{c(1-2x)\right\} + x - 1 = 0 \tag{24}$$

is considered separately. The necessary properties of equation (22) are obtained in the paper [3]. The first factor has roots only if $c \geq 4$. If $c \leq 2$, then the second factor has a unique root $x = \frac{1}{2}$ which is a point of minimum. The maximal value is attained at the end points of the interval and is equal to $c/2$. This completes the proof of the first statement.

The case of $c > 2$ is considered in more detail below. Let $c \in (2,4)$. Put

$$z(x) = x\exp\left\{c(1-2x)\right\} + x - 1.$$

To find the extreme points of the function z we evaluate the first two derivatives,

$$z'(x) = \mathrm{e}^{c-2xc}(1-2xc) + 1,$$
$$z''(x) = 4c\,(xc-1)\,\mathrm{e}^{c-2xc}.$$

The second derivative changes its sign at the point $1/c$, namely $z''(x) < 0$ for $x \in [0, 1/c)$, and $z''(x) \geq 0$ for $x \in (1/c, 1]$. The sign of the first derivative $z'(x)$ is given by

$$z'(x) > 0 \ \text{ for } x \in [0, x_1), \quad \text{where } x_1 \in \left(0, \frac{1}{c}\right),$$
$$z'(x) < 0 \ \text{ for } x \in [x_1, x_2), \quad \text{where } x_2 \in \left(1 - \frac{1}{c}, 1\right),$$
$$z'(x) > 0 \ \text{ for } x \in [x_2, 1].$$

As a result, the function $z(x)$ is nondecreasing in the interval $[0, x_1)$; then it is nonincreasing in the interval $[x_1, x_2)$, and then again it is nondecreasing in $[x_2, 1]$. Since $z(0) = -1 < 0$, $z(1/2) = 0$, and $z(1) = \ = e^{-c} > 0$, we conclude that $z(x)$ has three roots: $x_1 \in (0, 1/c)$, $x_2 = \frac{1}{2}$, and

$$x_3 \in (1 - 1/c, 1).$$

Thus the functional attains its minimal value $\frac{1}{2}(1-x_1)\exp\{cx_1^2\}$ at the points x_1 and x_3 if $2<c<4$. Comparing the values of the functional at x_2 and at the end points of the interval we see that the maximal value is equal to $c/2$ if $c\in(2,\ln 16)$, or $(c/4)e^{c/4}$ if $c\in[\ln 16,4)$.

Now let $c\geq 4$. There are five roots of equation (22) in this case:

$$x_1\in(0,1/c)\,,\qquad x_2=\frac{1}{2},\qquad x_3\in(1-1/c,1)\,,\quad \text{and}\quad x_{4,5}=\frac{\sqrt{c}\pm\sqrt{c-4}}{2\sqrt{c}}.$$

Denote the left hand side of equation (22) by y. Since the function y is antisymmetric about $x=\frac{1}{2}$, we conclude that the points of local minimum are $x_1\in(0,1/c)$, $x_2=\frac{1}{2}$, and $x_3\in(1-1/c,1)$. It is sufficient to compare the values of the functional at the points x_1 and $x_2=\frac{1}{2}$. For this, we consider the function

$$g(x)=e^{cx^2}(1-x),\qquad x\in[0,1].$$

Its first derivative equals $g'(x)=e^{cx^2}(2cx-2cx^2-1)$ and has two roots

$$\tilde{x}_{1,2}=\frac{1}{2}\pm\frac{1}{2}\sqrt{1-\frac{2}{c}}.$$

Indeed, $\tilde{x}_1<\frac{1}{2}<\tilde{x}_2$ and $g'(x)>0$ for $x\in(\tilde{x}_1,\tilde{x}_2)$, while $g'(x)<0$ for $x\in(0,\tilde{x}_1)\cup(\tilde{x}_2,1)$.

Moreover, at the point $\tilde{x}_1$,

$$z(\tilde{x}_1)=\left(\frac{1}{2}-\frac{1}{2}\sqrt{1-\frac{2}{c}}\right)\left(\exp\left\{c\sqrt{1-\frac{2}{c}}\right\}+1\right)-1>0.$$

This inequality is equivalent to

$$\exp\left\{\sqrt{c^2-2c}\right\}>\frac{1+\sqrt{1-\frac{2}{c}}}{1-\sqrt{1-\frac{2}{c}}}.$$

The right hand side of the latter inequality is rewritten as follows:

$$\frac{1+\sqrt{1-\frac{2}{c}}}{1-\sqrt{1-\frac{2}{c}}}=\left(c-1-\sqrt{c^2-2c}\right),$$

whence, for $u=\sqrt{c^2-2c}$,

$$e^u>1+u+\frac{u^2}{2}=1+\sqrt{c^2-2c}+\frac{c^2}{2}-c>c-1-\sqrt{c^2-2c}\quad\text{for } c\geq 4.$$

This means that $\tilde{x}_1>x_1$. Note that $g(1)=1$ and $g'(x)<0$ for $x<\tilde{x}_1$.

Comparing the values of $g'(x)$ at the points 0 and $\frac{1}{2}$, we conclude that $g'(0)<g'\left(\frac{1}{2}\right)$. Moreover, the value of the functional at the point x_1 is less than the value at $\frac{1}{2}$.

If $c=4$, then $x_2=x_4=x_5=\frac{1}{2}$ and the maximal value is equal to $(c/4)e^{c/4}$. The maximal value is attained at the points that are the roots of the quadratic equation if $c>4$, namely

$$x_{4,5}=\frac{c\pm\sqrt{c^2-4c}}{2c}=\frac{\sqrt{c}\pm\sqrt{c-4}}{2\sqrt{c}}.$$

Now we determine the intervals that contain the roots x_4 and x_5. Comparing x_4 with $\frac{1}{c}$, we derive

$$\frac{1}{2}-\frac{1}{2}\sqrt{1-\frac{4}{c}}>\frac{1}{c}.$$

Thus $x_4 \in (1/c, 1/2)$ and $x_5 \in (1/2, 1-1/c)$. The maximal value is equal to

$$\frac{1}{4}\exp\left\{\frac{(c-2)-\sqrt{c(c-4)}}{2}\right\} \times \left((c-2)-\sqrt{c(c-4)}-\left(\sqrt{c(c-4)}-(c-2)\right)\exp\left\{\sqrt{c(c-4)}\right\}\right).$$

This is the maximal value of the functional, indeed. This follows from the following reasoning. The point of local minimum $x = 1/2$ lies between points of maximum x_4 and x_5. Hence the value of the function at the point $x_2 = 1/2$ is less than the value at each of the points x_4 and x_5. As mentioned above, if $c > 4$, then the value at $x = 1/2$ is equal to $(c/4)e^{c/4}$ and exceeds the value $c/2$ corresponding to the end points. The proof of the theorem is complete. □

4. Concluding remarks

A theorem on extreme values of a functional constructed from three probability measures is proved in the paper. The functional considered in the paper contains a twice continuously differentiable function. Also, extreme values are found for an entropy-type functional constructed from three probability measures.

Bibliography

[1] H. Föllmer and A. Schied, *Stochastic Finance: An Introduction in Discrete Time*, Walter de Gruyter, Berlin, 2002, pp. 121–130. MR1925197

[2] C. Leonard, *Minimization of entropy functionals*, J. Math. Anal. Appl. **346** (2008), no. 1, 183–204. MR2428283

[3] Y. Mishura and H. Zhelezniak, *Extreme measures for entropy functionals*, Bull. Taras Shevchenko National University of Kyiv. Series: Physics & Mathematics (2017), no. 4, 15–20.

Department of Probability Theory, Statistics, and Actuarial Mathematics, Faculty for Mechanics and Mathematics, Kyiv Taras Shevchenko National University, Volodymyrs'ka Street, 64/13, Kyiv 01601, Ukraine

Email address: myus@univ.kiev.ua

Department of Probability Theory, Statistics, and Actuarial Mathematics, Faculty for Mechanics and Mathematics, Kyiv Taras Shevchenko National University, Volodymyrs'ka Street, 64/13, Kyiv 01601, Ukraine

Email address: hanna.zhelezniak@gmail.com

Received 15/OCT/2018

Translated by N. N. SEMENOV