Kolnogorov A. V. (Yaroslav-the-Wise-Novgorod State University, Veliky Novgorod, Russia). **On the minimax approach to the Gaussian multi-armed bandit**[1]**.**

We develop the limiting description of the minimax control [1] to the case of Gaussian multi-armed bandit; the case of three actions is considered for simplicity. Minimax strategy and risk are sought as Bayesian relative to the worst-case prior distribution.

Theorem. *Let us solve backwards a partial differential equation*

$$\min_{\ell=1,2,3}\left\{\frac{\partial r}{\partial t_\ell}+\frac{1}{2t_\ell^2}\times\left(\frac{\partial^2 r}{\partial S_\ell^2}-2\frac{\partial^2 r}{\partial S_\ell\partial S_{\overline{\ell}}}+\frac{\partial^2 r}{\partial S_{\overline{\ell}}^2}\right)+g_\ell(S,t,\varrho(w))\right\}=0,$$

where $\overline{\ell}-1\equiv\ell+1\,(\mathrm{mod}\ 3)$*,* $t_\ell\ge t_0$*, functions* $g_\ell\left(S,t,\varrho(w)\right)$ *depend on current statistics* S_ℓ*, times of applying actions* t_ℓ *and prior distribution* $\varrho(w)$ *(*$\ell=1,2,3$*). Initial condition is* $r(S_1,t_1,S_2,t_2,S_3,t_3)=0$ *if* $t_1+t_2+t_3=1$*. Bayesian strategy chooses action corresponding to the current smaller value on the left-hand side of the equation. For the Bayesian risk the estimate* $\lim_{K\to\infty}K^{-1/2}R_K^B=r(\varrho(w))$ *holds, where* $r(\varrho(w))=\lim_{t_0\downarrow 0}\iint_{-\infty}^{\infty}r(S_1,t_0,S_2,t_0,-S_1-S_2,t_0)dS_1dS_2$*. Minimax strategy and risk coincide with Bayesian ones, for which* $r(\varrho(w))$ *is maximal relative to* $\varrho(w)$*.*

REFERENCES

[1] A.V. Kolnogorov, “On a limiting description of robust parallel control in a random environment”, *Automation and Remote Control*, **76**:7, (2015), 1229–1241.

[1] The research was carried out using the infrastructure of the Shared Research Facilities “High Performance Computing and Big Data” (CKP “Informatics”) of FRC CSC RAS (Moscow).