A "WORST CASE" UNCERTAINTY SELECTION WITHIN A PROBABILISTIC CRITERION CONTROL PROBLEM STATEMENT

Kirill Chernyshov

V.A. Trapeznikov Institute of Control Sciences
65 Profsoyuznaya, Moscow 117997, Russia
E-mail: myau@ipu.rssi.ru

Abstract: A refined statement of a probabilistic criterion control problem, appeared in the literature at the ridge of the centuries, is proposed and the corresponding approach to solve it is derived. The approach is oriented to taking into account conditions of existence of the resulting domain of the admissible controls (non emptiness of the intersections of the "partial" domains), as well as to provide the conditions of unambiguous selection of the "worst" probabilistic distribution(s) of the plant output model variable. As a basic analytical tool, probability theory inequalities are applied. Various numerical examples are presented to confirm the theoretical inferences.

Keywords: Conditional probability, Criterion functions, Models, Stochastic control, Output variables, Plants, Probability density function, Regression relationships

1. PRELIMINARIES

In a paper of Bernatskii et al. (2000), a method to find a non-linear industrial plant robust control is proposed. Its essence is to use an assumption that the plant output model variable distribution belongs to a class of distributions, whose make-up is not known a priori. Representatives of the class are the conditional probability distributions whose parameters are determined at each step via experimental data. For the each representative, a domain of admissible controls (DAC) in accordance to a probabilistic criterion is determined. The intersection of the DACs received for separate representatives, to opinion of Bernatskii et al (2000), is not empty (what generically is wrong) and, by virtue of limitations imposed on the control vector components, forms a robust control domain. In turn, to select the representatives, an entropy based approach is, in particular proposed. The entropy approach is based on involving into the "working" class of distributions those representatives which have the maximal value of the entropy. But the entropy approach to robust control in (Bernatskii et al., 2000) leads nowhere since it is well known that for a given variance there exist unique density having the largest entropy among all the distribution densities, the Gaussian one, what, say, may be directly received from the fundamental paper of Janes (1957). Therefore there is nothing to select within the entropy based approach of Bernatskii et al. (2000). From another hand side, numerical examples in the same paper of Bernatskii et al. (2000) demolish the assumption on association of the magnitude of the entropy of the plant output model variable distribution and the set of the admissible controls. Thus, the approach of Bernatskii et al. (2000) should be considered as a delusion (Chernyshov, 2005).

In the present paper, a refined statement of the control problem of (Bernatskii et al., (2000) is proposed and the corresponding approach to solve it is derived. The approach is oriented to take into account conditions of existence of the domain of the admissible controls (non-emptiness of the intersections mentioned above), as well as to provide unambiguous selecting the "worst" probabilistic distribution of the plant output model variable. Various numerical examples are presented to confirm the theoretical inferences. Specifically, a condition to select not more than two representatives, (in the case of asymmetric distributions) or unique (in the case of symmetric distributions) representatives of the "worst" probability densities of the plant output model variable, and a domain of admissible controls corresponding to that density. Again, for the case of symmetric distribution densities, an extension of the domain of admissible controls has been obtained by use of modifications of the Gauss inequality.

2. REVISING A ROBUST CONTROL PROBLEM STATEMENT

Bernatskii et al. (2000) has proposed a method of finding a robust control for a non-linear industrial plant. A mathematical model of the output variable *y* of the plant, a technological process (TPr), is represented in the form

$$f = M\{y / \mathbf{U}, \mathbf{X} = \mathbf{X}_0\} = f(u_1, \ldots, u_m) \qquad (1)$$

where $\mathbf{U} = (u_1, \ldots, u_m)^T$ is the vector of control variables, $\mathbf{X} = (x_1, \ldots, x_n)^T$ is the vector of input controlled variables, and $M\{\cdot/\cdot\}$ stands for the conditional mathematical expectation. The entity of the method is using an assumption that the probability distribution density (PDF) of the plant output model variable belongs to a class of distributions *K* whose make-up is not known a priori. The class is represented by conditional probability distribution densities $p_j(y, f, \varphi),\ j = 1, \ldots, k$ whose parameters are determined at the each step (at the time interval of forming the control) via experimental data. Here φ is the conditional mean square deviation. "The objective condition on the plant performance is represented by the following probabilistic inequality

$$P\{y \in [A, B]\} \geq P_0 \qquad (2)$$

where P_0 is a given level of probability, which is considered technologically plausible, and $[A, B]$ is a technological tolerance on the output variable *y*, whose distribution belongs to the distribution class *K*." (Bernatskii et al., 2000 (p. 1004). For the each representative $p_j(y, f, \varphi),\ j = 1, \ldots, k$ of the class *K* in accordance to the probability criterion a domain of the admissible controls is determined, the so-called domain of the admissible controls. It is formed in the following manner. "For every representative in the distribution class under consideration, the following equation is solved with respect to *f* (under the assumption that $\varphi = \text{const}$):

$$\int_A^B p_j(y, f, \varphi)dy = P_0,\ j = 1, \ldots, k\ . \qquad (3)$$

For unimodal, symmetric, or asymmetric PDFs, which are commonly used in the description of the probabilistic properties of the output variables of a TPr, each of the equations (3) has two roots $f_{1j} > f_{2j}$. These two quantities define the interval $\left[f_{2j}, f_{1j}\right]$ of feasible values of the conditional mathematical expectation *f*, over which the objective condition (2) is fulfilled for the *j*-th representative." (Bernatskii et al., 2000 (p. 1004)).

"The intersection

$$\hat{F} = \bigcap_{j=1}^{k} \left[f_{2j}, f_{1j}\right] \qquad (4)$$

of these intervals for a single-output plant is non empty, since any particular value of *f* is a conditional mathematical expectation for all representatives in the class *K*. Using regression (1), intersection (4) is mapped on the space of controls; with account for technological limitations on the controls, this mapping defines a domain S_r of robust control in the space of controls" (Bernatskii et al., 2000 (p. 1005).

In turn, to select these representatives an entropy approach is, in particular, used by Bernatskii et al. (2000). The entropy approach is based on involving into the "working" class of distributions those representatives which have the maximal value of the entropy. "Since entropy is a measure of uncertainty in the system, the control in the case is synthesized against the worst (the most uncertain) conditions that can be proposed by the Nature. It is assumed that if the control problem may be solved against the worst conditions, than it can be also solved under more favorable (less uncertain) conditions" (Bernatskii et al., 2000 (pp. 1006-1007)).

Example 1. Let us demonstrate numerically that intersection (4) may be empty. The method is to derive plots of the functions

$$\Psi(f) = \int_A^B p_j(y, f, \varphi)dy,\ j = 1, \ldots, k \qquad (5)$$

for several densities $p_j(y, f, \varphi)$ under a fixed φ. In fact, figure 1 presents plots of the functions $\Psi(f)$ in (5) calculated for the Gaussian (solid line) and log-normal (dotted line) distributions at the following data: $A = 10^{-3}$, $B = 0.5$, $\varphi = 0.1$. It is clearly seen that, say, for the level $P_0 = 0.980$ intersection (4) is non-empty, for the level $P_0 = 0.985$ intersection (4) is empty for non-empty intervals $\left[f_{2j}, f_{1j}\right]$ and the robust control problem solution will not already exist, for the level $P_0 = 0.990$ already one of the two intervals $\left[f_{2j}, f_{1j}\right]$ is empty, for the level $P_0 = 0.995$ both the intervals $\left[f_{2j}, f_{1j}\right]$ are empty.

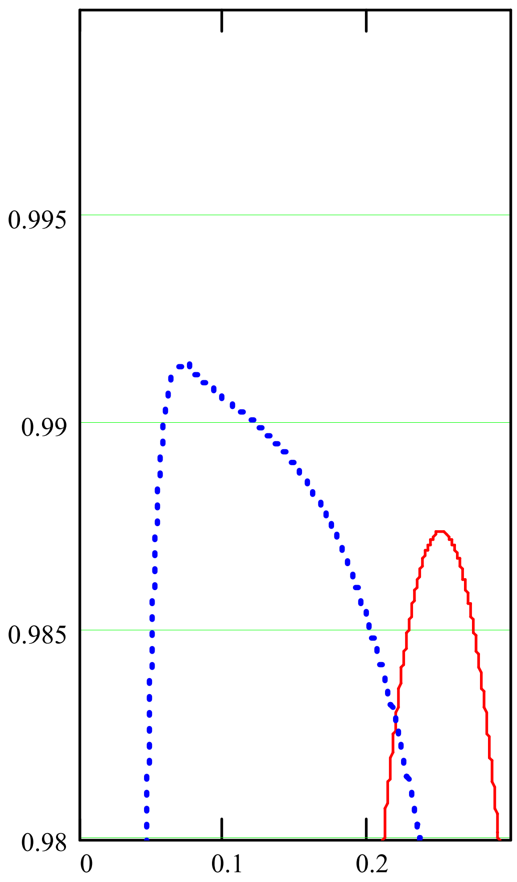


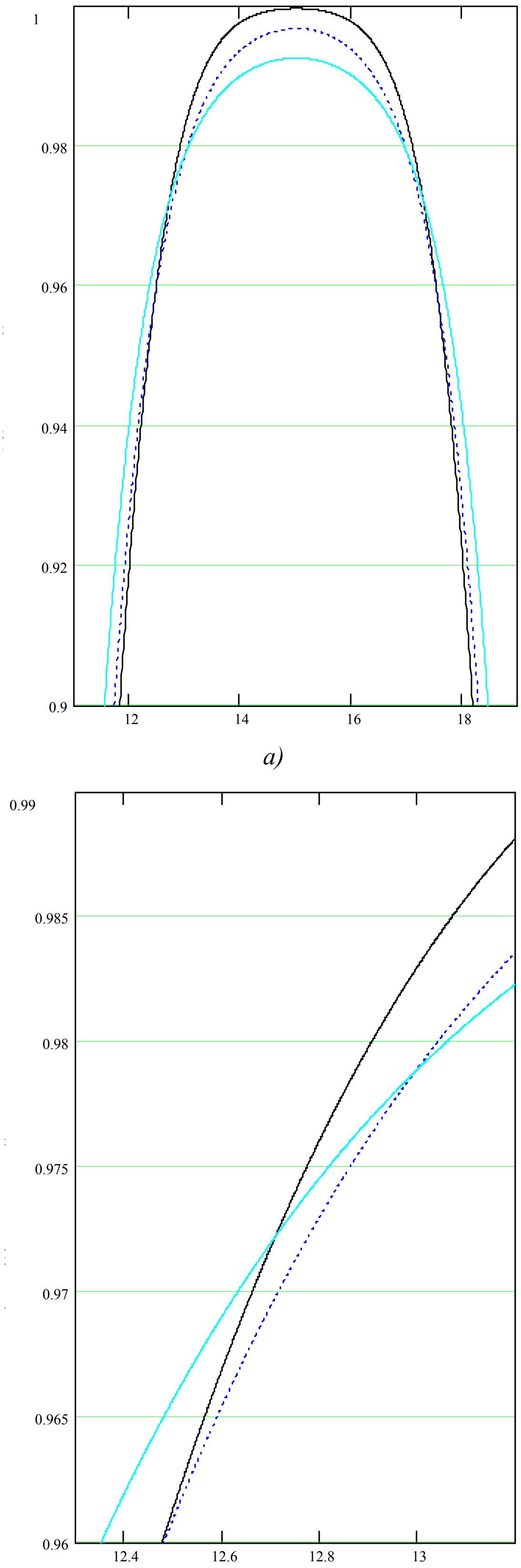
Fig. 1. Towards emptiness of intersection (4) for the Gaussian (solid line) and log-normal (dashed line) distributions at $A=10^{-3}$, $B=0.5$, $\varphi=0.1$.

Going back to the opinion of Bernatskii et al. (2000) that the intersection $\hat{F}$ in formula (4) "is non empty, since any particular value of f is a conditional mathematical expectation for all representatives in the class *K*", one should make an evident remark that the required level of the probability P_0 should not absolutely correspond to *any* particular value of the conditional mathematical expectation *f*.

Again, the entropy approach to robust control in (Bernatskii et al., 2000) leads nowhere since it is well known that for a given variance there exist unique density having the largest entropy among all the distribution densities, the Gaussian one, and there is nothing to select.

Example 2. Let in formula (4) the intersection $\hat{F}$ exist. Demonstrate numerically that determining $\hat{F}$ does not depend of the magnitude of the entropy of the distribution densities.

As well as above, figures 2 and 3 present plots of the functions $\Psi(f)$ calculated for the Gaussian (dark line), logistic (dotted line), and Student (light line) distributions at the following data: $A=10$, $B=20$. In the figures, $\varphi=\sqrt{2.0}$ for figure 2, and $\varphi=\sqrt{5.0}$ for figure 3. At that, figure 2b presents a refined scale representation of changing the density choice in the "neighbourhood" of the level of $P_0=0.980$. No special visual tools are required to insure that for a constant variance selecting the "worst case" distribution density depends on the required level of P_0 in (2) (or in (3)), but by no means on the entropy of the distribution.



b)

Fig. 2. Towards independence of determining $\hat{F}$ to the magnitude of the entropy of the distribution

densities for the Gaussian (dark line), logistic (dotted line) and Student (light line) distributions at $A = 10$, $B = 20$, $\varphi = \sqrt{2.0}$.

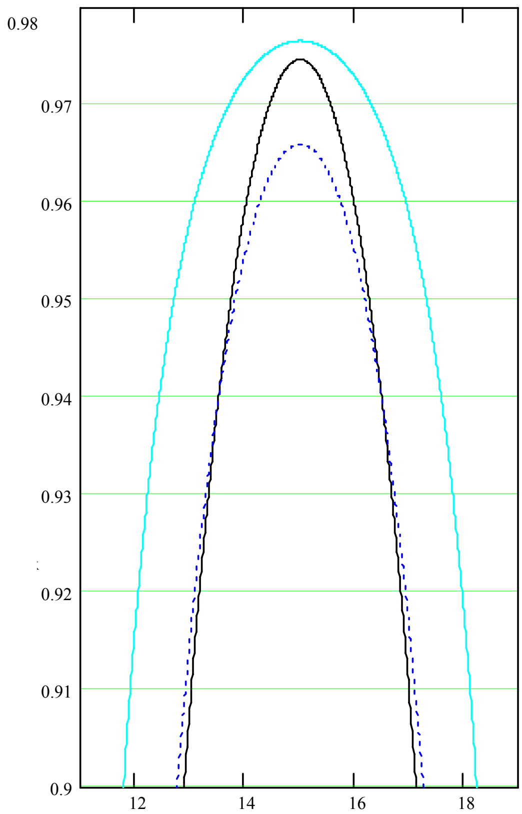


Fig. 3. Towards independence of determining $\hat{F}$ to the magnitude of the entropy of the distribution densities for the Gaussian (dark line), logistic (dotted line) and Student (light line) distributions at $A = 10$, $B = 20$, $\varphi = \sqrt{5.0}$.

3. AN IMPROVED APPROACH

In the present section, ways of improving the concept of Bernatskii et al. (2000) are proposed. The main step in that direction is determining conditions assuring non-emptiness of the intersection $\hat{F}$ in formula (4). At that, it is absolutely obvious that the width of the intervals $\left[f_{2j}, f_{1j}\right]$ in formula (4) and, finally, non-emptiness of their intersections are determined by two "parameters" of the initial problem statement, namely: P_0 and φ. As it has been clearly demonstrated by the presented examples, enlarging the probability P_0 or/and enlarging the mean squared deviation φ lead to emptiness of the intersection $\hat{F}$ in formula (4). Then, it would be naturally to assume an existence of a "balance" between the values of P_0 and φ, while meeting the balance would assure non-emptiness of the intersection $\hat{F}$ in formula (4).

Let, at first, the class K is formed by the representatives being symmetric unimodal distribution densities. Then, by use of the Tchebyshev inequality, taking into account the above presented notations, and by virtue of formula (2), one may write

$$1 - P_0 \le \frac{\varphi^2}{\left(\frac{B-A}{2}\right)^2}. \qquad (6)$$

At that, within the present paper context, just the condition of achieving the equality in formula (6) is of special interest. Thus, the value

$$\varphi_{\max}^2 = \left(\frac{B-A}{2}\right)^2 (1 - P_0) \qquad (7)$$

is that maximal magnitude of the variance at which one may guarantee non-emptiness of all the intervals $\left[f_{2j}, f_{1j}\right]$ in formula (4) given a probability P_0 and given a width of the interval $[A, B]$. For each $\varphi < \varphi_{\max}$ non-emptiness of the intersection $\hat{F}$ in formula (4) is even more so assured, both for the symmetric and asymmetric distributions, and, e.g. figure 4 for the Gumbel distribution density (i.e. $p(y) = \frac{1}{\beta}\exp\left(\frac{y-\alpha}{\beta} - \exp\left(\frac{y-\alpha}{\beta}\right)\right)$, where α and β are the distribution's parameters).

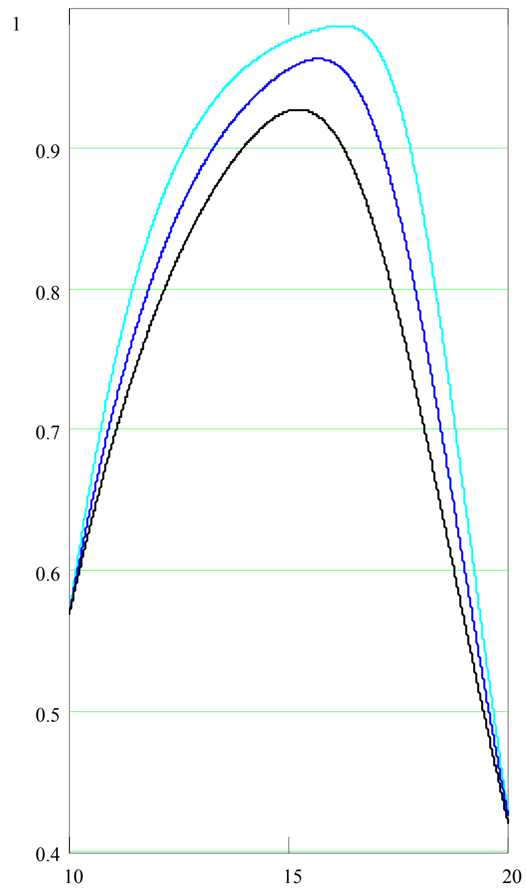


Fig. 4. Towards "embedding" of the intervals $\left[f_{2j}, f_{1j}\right]$ in (4) as φ increases for the Gumbel

distribution (at $A = 10$, $B = 20$, and $\varphi = 2$ (light line), $\varphi = 2.5$ (dotted line), $\varphi = 3$ (dark line)).

Let now the class *K* be represented by both symmetric and asymmetric unimodal distribution densities. Then, basing on the Tchebyshev inequality one may write (formally, more strict than that of (6)) condition

$$1 - P_0 \le \frac{M(y-\varsigma)^2}{(\min\{\varsigma - A, B - \varsigma\})^2}, \qquad (8)$$

where $M(\cdot)$ stands for the mathematical expectation symbol, ς is an arbitrary point from the interval $[A, B]$. At that, analogously to the preceding considerations, within the present problem context, just achieving the equality in formula (8) is of special interest (8). In turn, minimization of the expression standing in the right hand part of inequality (8) would only promote to achieving the equality. From the properties of the function $(\min\{\varsigma - A, B - \varsigma\})^2$ and by virtue of the equality $M(y-\varsigma)^2 = \varphi^2 + (M(y) - \varsigma)^2$, it follows directly that

$$\frac{M(y-\varsigma)^2}{(\min\{\varsigma - A, B - \varsigma\})^2} \ge \frac{\varphi^2}{\left(\frac{B-A}{2}\right)^2}, \qquad (9)$$

while the equality in formula (9) is achieved if and only if $\varsigma = \frac{A+B}{2}$ and $M(y) = \varsigma$.

Thus, from the above presented, it follows that the condition

$$\begin{cases} \varphi_{\max}^2 = \frac{1}{4}(B-A)^2(1-P_0) \\ \varphi^2 \le \varphi_{\max}^2 \end{cases}, \qquad (10)$$

is sufficient for non-emptiness of the intersection $\hat{F}$ in formula (4) for both symmetric and asymmetric (unimodal) densities from the class *K* (due to the fact that by virtue of condition (10) *all* the intervals $\left[f_{2j}, f_{1j}\right]$ contain the middle of the interval $[A, B]$).

More over, since the center of the interval $[A, B]$ is always "captured" by the intervals $\left[f_{2j}, f_{1j}\right]$ under condition (10), the condition of unimodality may thus be omitted. At that, of course, for a given distribution, only two roots of equation (3) being the closest to the point $\frac{A+B}{2}$ from the left and right should be chosen.

The above presented example 1 evidently demonstrates that under violation of condition (10) the middle of the interval $[A, B]$ may not belong to all the intervals $\left[f_{2j}, f_{1j}\right]$ in formula (4): within the conditions of example 1, condition (10) implies the following (upper) bound for the mean squared deviation: $\varphi \le \sqrt{\frac{(B-A)^2(1-P_0)}{4}} = 0.030561$ as $P_0 = 0.985$ and $\varphi \le \sqrt{\frac{(B-A)^2(1-P_0)}{4}} = 0.017635$ as $P_0 = 0.995$ what contradicts to the considered magnitude of the value $\varphi = 0.1$.

In the paper of Bernatskii et al. (2000), a hypothetical class of distribution K_h is given, from which one the selection of the "worst" densities is performed. To be involved into that class, the following distributions have been considered: uniform (note that it is not a unimodal distribution in contrast to the preliminary postulation of Bernatskii et al. (2000)), Gaussian, logistic, log-normal, triangular, Gumbel, extreme value ($p(y) = \frac{1}{b}\exp\left(\frac{a-y}{b} - \exp\left(\frac{a-y}{b}\right)\right)$, *a* and *b* are the distribution's parameters), Weibull, Tukey. At that, as a basic reason of non-involving a distribution into the class K_h, the paper of Bernatskii et al. (2000) refers to "complexity" of an explicit analytical expression for its entropy. Due to that reason, the paper of Bernatskii et al. (2000) has found to be non-reasonable to involve into the hypothetical class K_h the distributions of Weibull, Pearson, Johnson, gamma-distribution, and a number of other ones. Hardly ever such reasoning is acceptable as a justification (even, non-formal one) of a criterion to form the hypothetical class of distributions K_h. In contrast to the paper of Bernatskii et al. (2000), condition (10) derived in the present paper by no means restricts (even disregarding the condition of unimodality) forming the hypothetical class of distributions K_h, if one will take into account that a numerical (while another one is not required) solution of equations (3) possesses no difficulty for any explicitly given distributions $p_j(y, f, \varphi)$, $j = 1, \ldots, k, \ldots$.

4. IMPROVING CONDITION (10) FOR UNIMODAL DISTRIBUTIONS

In spite of its universal form which is suitable for all types of probability distribution densities, condition (10) looks rather rough and imposing exhaustively strict limitation to the admissible mean squared deviation with regard to the unimodal distributions. Under the condition of unimodality of the probability distribution functions, condition (10) can be refined by use of the above technique accompanied by some additional results in the field of the probability theory inequalities. Specifically, one should apply the following generalization (Vysochanskij and Petunin, 1985a) of the Gauss inequality:

$$P\{|\xi - x_0| \geq k\theta\} \leq \frac{4}{9k^2} \text{ for all } k > 0, \quad (11)$$

where $\theta^2 = M(\xi - x_0)^2$, x_0 is an arbitrary real number, and ξ is a random value having a unimodal distribution (the conventional Gauss inequality requires x_0 to be the mode coinciding with the mean of that unimodal distribution; and inequality (11) is also referred as the Vysochanskij-Petunin inequality).

Let, similarly to (8), ς be an arbitrary point from the interval $[A, B]$. Imposing in (11) $\xi \overset{\Delta}{=} y, x_0 \overset{\Delta}{=} \varsigma$, and $k = \frac{\min\{\varsigma - A, B - \varsigma\}}{\sqrt{M(y-\varsigma)^2}}$ one directly obtain from (11)

$$1 - P_0 \leq \frac{4}{9} \frac{M(y-\varsigma)^2}{(\min\{\varsigma - A, B - \varsigma\})^2}. \quad (12)$$

At that, analogously to the preceding Section considerations, within the present problem context, just achieving the equality in formula (12) is of special interest. In turn, minimization of the expression standing in the right hand part of inequality (12) would only promote to achieving the equality. Acting in the above Section manner, the minimum of the right-hand part of formula (12) is achieved if and only if $\varsigma = \frac{A+B}{2}$ and $M(y) = \varsigma$. Thus from (12) and the presented considerations, the following sufficient condition which should be imposed on the admissible magnitude of the variance φ^2 is

$$\begin{cases} \varphi_{\max}^2 = \frac{9}{16}(B-A)^2(1-P_0) \\ \varphi^2 \leq \varphi_{\max}^2 \end{cases}. \quad (13)$$

Going back to the above considered numerical illustration to example 1, one can conclude that under violation of condition (13) the middle of the interval $[A, B]$ may not belong to all the intervals $\left[f_{2j}, f_{1j}\right]$ in formula (4): condition (13) implies the following (upper) bound for the mean squared deviation: $\varphi \leq 0.045842$ as $P_0 = 0.985$ and $\varphi \leq 0.026453$ as $P_0 = 0.995$ what also contradicts to the considered magnitude of the value $\varphi = 0.1$.

5. CONCLUSIONS

Conditions (10), (13) derived, provide thus a possibility to select within the class *K* not more than two (in the case of asymmetric distributions) or one (in the case of symmetric distributions only) "worst" distribution densities of the plant output model variable and a domain of the admissible robust controls, corresponding to that density. In entity, the first (equality) parts of conditions (10) and (13) also determine the corresponding limit admissible magnitudes $P_0^{\max}$, and $\Delta_{\min}$ of the probability P_0, and the width Δ of the interval $[A, B]$ respectively: $P_0^{\max} = g_1(\varphi, \Delta)$, $\Delta_{\min} = g_2(\varphi, P_0)$, meeting to one of them under given the rest two magnitudes is a sufficient condition of non-emptiness of intersection (4).

If conditions of the solved practical problem statement enable one to restrict the consideration to the class of unilateral criterion only, say $P\{y \leq B\} \geq P_0$, than for the unimodal distribution densities a condition being similar to (13) may be obtained by use of corresponding generalizations of probabilistic inequalities (e.g. that of (Vysochanskij and Petunin, 1985b); what, in turn, will enable one to extend the domain of the admissible robust controls.

REFERENCES

Bernatskii F.I., Dobrodeev D.L., and Pashchenko F.F. (2000), "On the Reduction of Capacity of the Distributional Class in Robust Control of Technological Plants", Automation and Remote Control, vol. 61, no. 6, part 2, pp. 1003-1011.

Chernyshov K.R. (2005), "Information Theory within Identification and Control Problems: an Unexpected Review", Stability and Control Processes Conference SCP 2005 dedicated to 75th birthday anniversary of V.I. Zubov, June 29 - July 1, 2005, Saint-Petersburg State University, Saint Petersburg, Russia, pp. 1000-1009.

Janes E.T. (1957), "Information theory and statistical mechanics", Phys. Rev., vol. 106, pp. 620-630.

Vysochanskij D.F. and Petunin Yu.I. (1985a), "Generalization of Gauss inequality for unimodal distributions", Theory Probab. Math. Stat., vol. 31, pp. 29-34.

Vysochanskij D.F. and Petunin Yu.I. (1985b), "Improvement of the unilateral 3σ-rule for unimodal distributions" Generalization of Gauss inequality for unimodal distributions", Dokl. Akad. Nauk Ukr. SSR, Ser. A., no. 1, pp. 6-8. (in Russian); English summary available in European Mathematical Society, FIZ Karlsruhe & Springer Verlag.