УДК 531

ПЕРЕСЧЕТ ДОВЕРИТЕЛЬНЫХ ДИАПАЗОНОВ ХАРАКТЕРИСТИК ПРИ ПЕРЕХОДЕ ОТ НЕОГРАНИЧЕННЫХ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ К РАВНОМЕРНЫМ В ЗАДАЧАХ СИТУАЦИОННОГО МОДЕЛИРОВАНИЯ

Синицын Сергей Александрович
доктор технических наук, профессор, заведующий кафедрой
«Теоретическая и прикладная механика»
Российского университета транспорта (РУТ(МИИТ))

Аннотация

Схема поэтапного решения ситуационной задачи определяет последовательность случайных зависимых событий, определенных как реализации характеристических параметров рабочих моделей. В процессе решения ситуационной задачи каждая характеристика $Pr_i^{(j)}$ подлежит уточнению согласно диапазонному алгоритму. Помимо выбора значений параметров на каждом последующем этапе должны уменьшаться соответствующие доверительные диапазоны ΔPr_i. По мере уточнения диапазонов число содержащихся в них интервалов квантования уменьшается и в состоянии законченного решения интервал совпадает с диапазоном, то есть количество интервалов становится равным единице. Модель условных вероятностей позволяет использовать в данном исследовании характеристику условной энтропии, которая позволяет выполнять дополнительное уменьшение доверительных диапазонов на каждом последующем этапе решения ситуационной задачи. Каждый параметр модели определяется законом распределения вероятностей и величиной своего доверительного диапазона, который является зависимым от последовательности реализаций значений этого параметра на всех этапах решения задачи. Условная плотность строится для ситуаций, отражающих связь предыдущих реализаций при условии наступления последующих этапов моделирования. Задачи такого рода принято называть обратными задачами моделирования, которые позволяют прогнозировать на основе статистических связей значения характеристик ранних этапов с учетом требований конечных решений. Статистические характеристики могут быть получены на основе обработки экспериментальных данных с помощью известных соотношений теории вероятностей. Переход от условной энтропии к величине доверительного диапазона предлагается выполнять путем преобразования энтропии реального распределения в энтропию равномерного распределения. Методика такого преобразования приведена в этой статье.

Ключевые слова: характеристические параметры, ситуационные модели, доверительные диапазоны, законы распределения, условная энтропия, энтропийный диапазон, многоэтапные модели, условная плотность, обратные задачи моделирования, среднеквадратическая погрешность, энтропийная погрешность.

CALCULATION OF CONFIDENTIAL RANGE OF CHARACTERISTICS AT THE TRANSITION FROM UNLIMITED DISTRIBUTION LAWS TO UNIFORM IN SITUATIONAL SIMULATION PROBLEMS

Sergey A. Sinitsyn
Doctor of Technical Sciences, Professor, Head of Department
“Theoretical and Applied Mechanics”
Russian University of Transport (RUT (MIIT))

ABSTRACT

The stage-by-stage solution scheme for the situational problem determines the sequence of random dependent events defined as the implementation of the characteristic parameters of the working models. In the process of solving the situational problem, each characteristic Pri (j) is subject to refinement according to the range algorithm. In addition to the choice of parameter values, at each subsequent stage, the corresponding confidence ranges ΔPri should decrease. As the ranges are refined, the number of quantization intervals contained in them decreases and, in the state of the final solution, the interval coincides with the range, i.e., the number of intervals becomes equal to unity. The conditional probability model allows us to use the conditional entropy characteristic in this study, which allows us to further reduce the confidence ranges at each subsequent stage of solving the situational problem. Each parameter of the model is determined by the law of probability distribution and the value of its confidence range, which is dependent on the sequence of implementations of the values of this parameter at all stages of solving the problem. Conditional density is constructed for situations that reflect the relationship of previous implementations, provided that the subsequent stages of modeling occur. Tasks of this kind are usually called inverse modeling problems, which allow predicting, based on statistical relationships, the values of the characteristics of the early stages, taking into account the requirements of the final solutions. Statistical characteristics can be obtained based on the processing of experimental data using known relations of probability theory. The transition from conditional entropy to the value of the confidence range is proposed to be performed by converting the entropy of the real distribution into the entropy of a uniform distribution. The methodology for such a conversion is given in this article.

Keywords: characteristic parameters, situational models, confidence ranges, distribution laws, conditional entropy, entropy range, multi-stage models, conditional density, inverse modeling problems, mean square error, entropy error.

Условная энтропия $H(Pr_i^{(j)}/Pr_i^{(j-1)})$ определяет некоторую функциональную зависимость вероятности на интервале с бесконечными границами (-∞;+∞)(рис.1.)

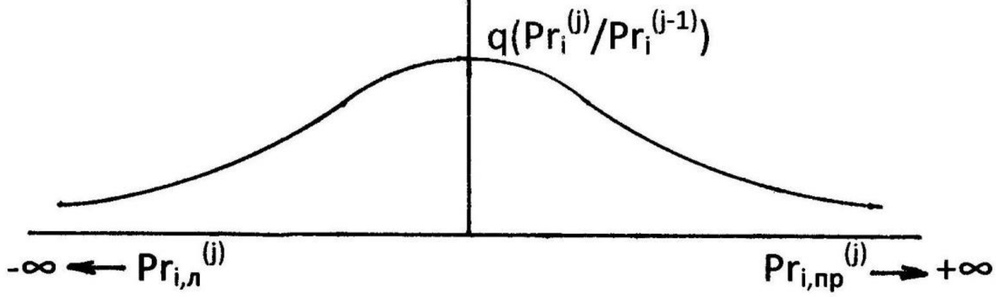

Рисунок 1. *Теоретическая область определения условной вероятности*

В наших ситуационных задачах на каждом этапе моделирования необходимо вычислять конечные значения доверительных диапазонов по каждой характеристике $\Delta Pr_i^{(j)}$, поэтому теоретический предел области определения функции энтропии $H(Pr_i^{(j)}/Pr_i^{(j-1)})$ должен быть ограничен конечными значениями левой и правой границ доверительного диапазона: $Pr_{i,Л}^{(j)}$; $Pr_{i,пр}^{(j)}$. Задача вычисления этих границ может быть решена несколькими способами [1, с. 33].

Одно из возможных решений основано на эквивалентном преобразовании энтропии бесконечного распределения в энтропию равновероятного или равномерного распределения с плотностью, обладающей резко очерченными границами. При этом в качестве основного критерия принимается признак равенства энтропий реального и равновероятного законов распределения [2, с. 258].

Функцией преобразования в ситуационных задачах является зависимость характеристик Pr_1, Pr_2, …, Pr_m от определяющих параметров исходного состояния субъекта: $Z_1, Z_2, \ldots, Z_m, \zeta_1, \zeta_2, \ldots, \zeta_m$:

$$Pr = F(Z,\zeta). \qquad (1)$$

Такая функция не может быть представлена строго детерминированной зависимостью в силу ряда особенностей ситуационных задач [3, с. 8]. Поэтому совокупности вполне определенных характеристик, составляющих массив исходных данных $\{Z\}$ может соответствовать множество конечных параметров $\{Pr^*\}$, которые принимают значения в некоторых интервалах шириной $\{\Delta Pr_{доп}^{(j)}\}$ относительно своих номинальных характеристик. Размеры промежуточных интервалов, названных доверительными диапазонами, зависят от степени конкретизации задачи на промежуточных этапах моделирования. То есть определяются свойствами конкретных моделей этапов решения ситуационной задачи.

Таким образом, на каждом этапе номинальная составляющая характеристики определена в своем доверительном диапазоне, отражающим степень ее неопределенности или точности решения. В этой схеме более важная роль отведена самому доверительному диапазону $\Delta Pr_i^{(j)}$, нежели номинальному значению характеристики $Pr_i^{(j)}$, которое не отражает точное решение задачи на любом промежуточном этапе.

Вероятные значения каждой i-й характеристики в любом j-м состоянии решения ситуационной задачи ограничены некоторым доверительным диапазоном $\Delta Pr_i^{(j)}$, в рамки которого эти значения обязаны уложиться. В таком случае говорят, что решение достоверно с заданной вероятностью или для доверительного диапазона – с заданной точностью.

Если не рассматривать равновероятные законы распределения, то каждая характеристика $Pr_i^{(j)}$ может обладать каким-то произвольным законом распределения. То есть в худшем случае законы распределения могут быть все различными. Приближенным описанием закона распределения может быть указание максимальной погрешности в ограниченном ряде реализаций характеристик $Pr_i^{(j)}$.

Законы распределения погрешностей для различных характеристик могут быть весьма разнообразными [4, с. 76]. Такое разнообразие создает значительные трудности при вычислении эффективного диапазона достоверности каждой характеристики модели этапа, который однозначно мог бы характеризовать величину неопределенности, остающейся после выполнения моделирования на данном этапе решения ситуационной задачи.

Например, в случае определения эквивалентных интервалов неопределенности характеристик $Pr_i^{(j)}$ для распределений, представленных на рисунках 2 и 3, их приходится произвольно приписывать ширине полосы доверительных вероятностей $\Delta Pr_i^{(j)}$.

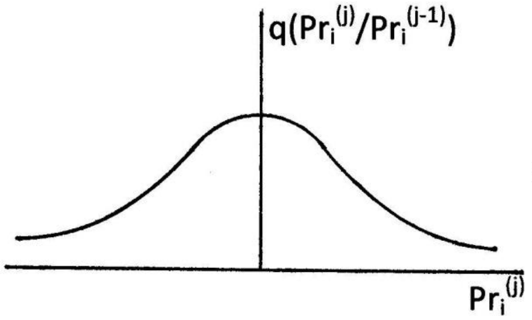

Рисунок 2. *Распределение типа пологой кривой*

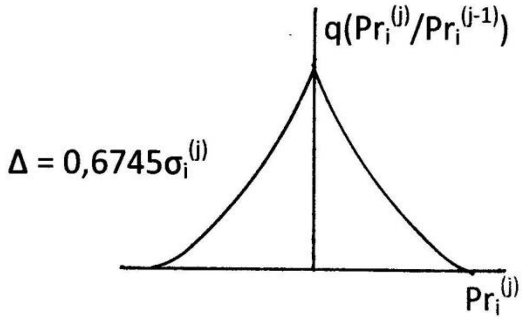

Рисунок 3. *Остронаправленное распределение*

Эта полоса может быть принята равной средней вероятностной погрешности. При этом, если закон распределения характеристик нормальный, то половина всех возможных значений будет меньше этой величины, а вторую половину будут составлять большие значения. В этом случае следует напомнить, что определяющая погрешность выбрана с доверительной вероятностью $Р_д$=0,5.

В то же время, применение энтропийной характеристики доверительных диапазонов позволяет получить простое решение задачи с требуемой доверительной вероятностью $Р_д$=0,95.

Количество информации, получаемое при моделировании на данном этапе решения ситуационной задачи, равно уменьшению неопределенности состояния, которая в нашем исследовании измеряется энтропией:

$$Inf(Pr^{(j)}) = H(Pr^{(j-1)}) - H(Pr^{(j)}/Pr^{(j-1)}), \qquad (2)$$

то есть равно разности энтропий в начале и конце этапа моделирования. При этом исходная неопределенность, как безусловная энтропия $H(Pr^{(j-1)})$ принимается зависящей только от распределения вероятностей значений характеристик $Pr^{(j-1)}$, то есть до начала выполнения этапа моделирования, и не зависит от распределения вероятностей тех же характеристик $Pr^{(j)}$ после выполнения этапа моделирования [5,с.199]. Остаточная неопределенность j-го этапа определена условной энтропией $H(Pr^{(j)}/Pr^{(j-1)})$ и зависит от распределения вероятностей на остаточном диапазоне Δ $Pr^{(j)}$. Поэтому при информационном решении ситуационных задач каждая из характеристик соседних этапов моделирования может исследоваться отдельно и независимо от остальных.

На рисунке 4 представлена плотность распределения $q(Pr_i^{(j)})$ для равномерного закона распределения. На рисунке левая граница распределения или, в нашем понимании доверительного диапазона, определена минимальным значением характеристики $Pr^{(j)}$, определенного рядом условий существования моделей данного уровня [6, с. 200], правая граница – максимальной величиной этой же характеристики того же j-го состояния ситуационной задачи.

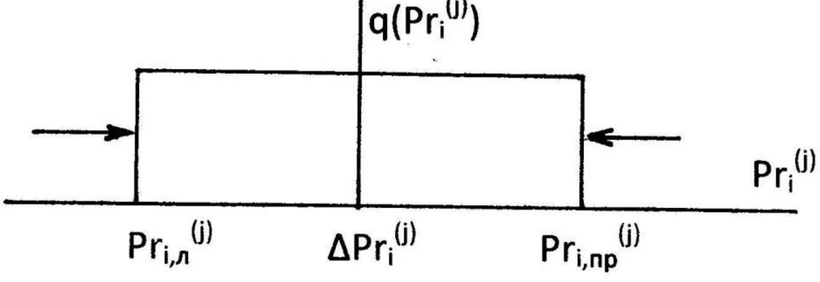

Рисунок 4. *Равномерное распределение вероятностей на характеристическом диапазоне параметра*

В этой задаче плотность распределения $f(Pr_i^{(j)})$ может быть записана по трем участкам:

$$f(Pr_i^{(j)}) = 0 \text{ при } Pr_i^{(j)} < Pr_{i,л}^{(j)} \text{ и } Pr_i^{(j)} > Pr_{i,пр}^{(j)}\text{, то есть при } |\, Pr_i^{(j)} \,| > \Delta Pr_i^{(j)} \text{ и}$$

$$f(Pr_i^{(j)}) = 1/\Delta Pr_i^{(j)} \quad \text{при} \quad |\, Pr_i^{(j)} \,| \le \Delta Pr_i^{(j)}. \qquad (3)$$

Энтропия j-го диапазона характеристики $Pr_i^{(j)}$ при равномерном законе распределения может быть определена по формуле:

$$H(Pr_i^{(j)}/Pr_i^{(j-1)}) = -\int_{Pri,л}^{Pri,п} 1/\Delta Pr_i^{(j)} \ln\Delta Pr_i^{(j)} dPr_i^{(j)} = \ln \Delta Pr_i^{(j)}. \qquad (4)$$

Величина среднеквадратической погрешности характеристики $Pr_i^{(j)}$ определяется соотношением:

$$\sigma_i^{(j)2} = \int_{Pri,л}^{Pri,п} Pr_i^{(j)2} f(Pr_i^{(j)}) d\, Pr_i^{(j)} = \Delta Pr_i^{(j)2}/12. \qquad (5)$$

Таким образом, для равномерного распределения вероятностей:

$$\sigma_i^{(j)} = \Delta Pr_i^{(j)}/\, 2\sqrt{3}\,; \qquad \Delta Pr_i^{(j)} = 2\sqrt{3}\,\sigma_i^{(j)}. \qquad (6)$$

Подставляя (6) в (4) получаем формулу для вычисления энтропии равномерного распределения:

$$H(Pr_i^{(j)}/Pr_i^{(j-1)}) = \ln 2\sqrt{3}\,\sigma_i^{(j)}. \qquad (7)$$

Если характеристики модели в доверительном диапазоне имеют нормальное распределение, то есть:

$$f(Pr_i^{(j)}) = 1/\sqrt{2\pi}\,\sigma_i^{(j)} \times e^{-\frac{Pri(j)2}{2\sigma i(j)2}}, \qquad (8)$$

то условная энтропия в конце этапа моделирования определится соотношением [7, с. 209]:

$$H(Pr_i^{(j)}/Pr_i^{(j-1)}) = \int_{-\infty}^{+\infty} f(Pr_i^{(j)})(\ln\sqrt{2\pi}\,\sigma_i^{(j)} + Pr_i^{(j)2}/\,2\sigma_i^{(j)2}) dPr_i^{(j)} =$$

$$= \ln\sqrt{2\pi}\,\sigma_i^{(j)} \int_{-\infty}^{+\infty} f\,(Pr_i^{(j)})\, dPr_i^{(j)} + 1/2\sigma_i^{(j)2}) \int_{-\infty}^{+\infty} f\,(Pr_i^{(j)2}) f(Pr_i^{(j)}) dPr_i^{(j)}. \qquad (9)$$

Поскольку:

$$\int_{-\infty}^{+\infty} f\,(Pr_i^{(j)})\, dPr_i^{(j)} = 1,$$

а второе слагаемое численно равно дисперсии:

$$\int_{-\infty}^{+\infty} f\,(Pr_i^{(j)2}) f(Pr_i^{(j)}) dPr_i^{(j)} = 2\sigma_i^{(j)2},$$

то условная энтропия может быть выражена формулой:

$$H(Pr_i^{(j)}/Pr_i^{(j-1)}) = \ln\sqrt{2\pi}\,\sigma_i^{(j)} + 1/2 = \ln\sqrt{2\pi}\,\sigma_i^{(j)} + \ln\sqrt{e} = \ln(\sqrt{2\pi e}\,\sigma_i^{(j)}). \qquad (10)$$

Формула для доверительного диапазона с нормальным распределением вероятностей отличается от ранее полученного выражения для равномерного распределения, резко ограниченного шириной полосы $\Delta Pr_i^{(j)}$, произведением под знаком логарифма [8, с. 62].

Таким образом, с информационной точки зрения неограниченное распределение в форме пологой кривой, например, нормальное, приводит к получению точно такого же значения энтропии, как и резко ограниченное равномерное распределение, если выполняется равенство:

$$\Delta Pr_i^{(j)} = \sqrt{2\pi e}\,\sigma_i^{(j)}. \qquad (11)$$

Полученный результат позволяет сделать важный вывод о том, что в качестве энтропии некоторого j-го этапа решения ситуационной задачи можно всегда рассматривать энтропию доверительного диапазона с равномерным законом распределения, которая вносит такую же неопределенность в решение задачи, что и энтропия фактического неограниченного закона распределения. При этом величина доверительного диапазона будет равна:

$$\Delta Pr_i^{(j)} = \exp H(Pr_i^{(j)}/Pr_i^{(j-1)}). \quad (12)$$

Возможность перехода от любого неограниченного закона распределения вероятностей к равномерному позволяет значительно упростить процедуру вычисления величин доверительных диапазонов при переходе от этапа к этапу моделирования в процессе решения ситуационных задач.

Список литературы

1. Левчук Т.В., Маслов А.А. Использование имитационного моделирования для анализа эксплуатационных испытаний программного обеспечения// История и перспективы развития транспорта на севере России. 2015. №1. С.32-34.
2. Синицын С.А., Гусарова О.Ф. Информационный подход к разработке и применению иерархических ситуационных моделей интерактивного интеллекта// Социология. 2019. №1. С.255-262.
3. Синицын С.А., Гусарова О.Ф. Информационные характеристики доверительных диапазонов параметров ситуационных моделей// Оригинальные исследования. 2019. Т.9. №4. С.4-12.
4. Левчук Т.В., Ким А.Р. Трехмерное моделирование при визуализации математических и технических задач// История и перспективы развития транспорта на севере России. 2016. №1. С.75-78.
5. Левчук Т.В. Современные пакеты прикладных программ в инженерной и научной деятельности //История и перспективы развитие транспорта на севере России. 2013. №1. С.196-200.
6. Левчук Т.В., Втулкин М.Ю., Череватый Д.Н. Применение интегрированных пакетов в частных задачах вычислительной математики //История и перспективы развития транспорта на севере России. 2014. №1. С.199-200.
7. Синицын С.А., Дубровин В.С. Конечные схемы распределения точечных множеств геометрических объектов// Современные проблемы совершенствования работы железнодорожного транспорта. №13. 2017. С.207-213.
8. Левчук Т.В. Эффективность внедрения информационных систем на железнодорожном транспорте//История и перспективы развития транспорта на севере России. 2012. №1. С.60-64.

References

1. Levchuk T.V., Maslov A.A. The use of simulation for analysis of operational testing of software // History and prospects of transport development in the north of Russia. 2015. No1. P. 32-34.
2. Sinitsyn S.A., Gusarova O.F. An informational approach to the development and application of hierarchical situational models of interactive intelligence // Sociology. 2019. №1. P. 255-262.
3. Sinitsyn S.A., Gusarova O.F. Information Characteristics of Confidence Ranges of Parameters of Situational Models // Original Research. 2019. V. 9. Number 4. P. 4-12.
4. Levchuk T.V., Kim A.R. Three-dimensional modeling in the visualization of mathematical and technical problems // History and prospects of transport development in the north of Russia. 2016. No1. P. 75-78.
5. Levchuk T.V. Modern packages of applied programs in engineering and scientific activities // History and prospects of transport development in the north of Russia. 2013. No1. P. 196-200.

6. Levchuk T.V., Vtulkin M.Yu., Cherevaty D.N. The use of integrated packages in particular problems of computational mathematics // History and prospects of transport development in the north of Russia. 2014. No1. P. 199-200.
7. Sinitsyn S.A., Dubrovin V.S. Finite distribution schemes for point sets of geometric objects // Modern problems of improving the work of railway transport. No. 13. 2017. P. 207-213.
8. Levchuk T.V. The effectiveness of the implementation of information systems in railway transport // History and prospects of transport development in the north of Russia. 2012. No1. P. 60-64.