

УДК 519.6

А.П.Гожий, И.И.Коваленко

## СИСТЕМНЫЙ ПОДХОД К ВЫБОРУ СТАТИСТИЧЕСКИХ МЕТОДОВ ПРИ ГЕНЕРАЦИИ И АНАЛИЗЕ СЦЕНАРИЕВ

**Введение.** Для анализа принципиально новых проектов и прогноза фундаментальных и прикладных исследований разрабатываются сценарии, в основе которых лежит построение так называемого «дерева целей» или «прогнозного графа», которые отображают как инновационную потребность, так и инновационную альтернативу, в том числе оценку необходимых ресурсов и производственных возможностей.

При разработке сценариев желательна многовариантность описания будущей ситуации, что дает возможность анализировать и определять наиболее рациональную стратегию воздействия на ситуацию. Для разработки сценариев привлекаются различные группы высококвалифицированных специалистов, аналитиков, консультантов, экспертов, и лиц принимающих решения (ЛПР), которые в своей работе выполняют определенные функции.

Очевидно, что при разработке сценариев ни один из имеющихся методов изолированно, вне связи с другими не может дать исчерпывающих результатов. Поэтому стремление к использованию отдельных методов в определенной последовательности с установлением определенных взаимосвязей между ними, является необходимостью. В современной трактовке такой процесс получил название технологическое предвидение [1], которое и формируется на основе методологии системного анализа. Необходимо отметить, что вопросы формализации процедур построения сценариев на основе системного применения ряда инструментальных методов являются наименее разработанными, по ней практически отсутствует библиография, сослаться можно, пожалуй, лишь на работу [2].

**Постановка задачи.** В данной статье рассматривается системный подход к построению сценариев, в рамках которого рассматриваются следующие вопросы:

систематизированы и выделены основные группы так называемых «инструментальных методов», используемых для

построения и анализа сценариев (методы генерации информации качественного характера; графовые модели, необходимые для построения деревьев целей или событий; вероятностно-статистические методы для получения оценок событий на графах и методы сворачивания сценариев; методы поддержки принятия решений по выбору лучшего из множества альтернативных сценариев).

предложено с позиций системного подхода рассматривать определенную совокупность используемых инструментальных методов в виде системы методов (СМ), между элементами которой (в качестве элементов выступают отдельные методы) существуют определенные связи или формы взаимодействия;

на основе анализа СМ рассмотрены процедуры формирования системных технологий выбора различных вариантов и последовательностей методов в соответствии с установленными критериями и принципами оптимальности их использования.

**Основная часть.** На начальном этапе изучения проблемы и её качественного анализа используются методы генерации информации качественного характера (мозговая атака, метод аналогий, метод контрольных вопросов, морфологический анализ, эвристические методы и др.), с помощью которой анализируют характерные особенности, определяют направления исследований, формулируют важнейшие критерии для выбранной проблемы (рис.1).

На этапе написания сценариев используются методы теории графов (построение дерева целей, дерева событий, прогнозный граф и др.), что дает возможность получить топологию цепочек событий, связанных причинно-следственными связями. Для количественного оценивания вероятностей событий, из которых формируется сценарий, и последующего его сворачивания применяются вероятностно-статистические методы (стратегия Байеса, многошаговый алгоритм Байеса, методы параметрической, робастной и непараметрической статистики).

После написания сценариев и оценивания их реалистичности последним важным этапом является их представление группе лиц, принимающих решения. В данной ситуации применяются методы поддержки принятия решений (однокритериальные методы при объективных моделях, многокритериальные методы при объективных моделях и многокритериальные методы при субъективных моделях).

Представим некоторую совокупность методов, используемую на различных этапах построения сценариев которую, можно представить в виде системы методов, между элементами которой (в качестве элементов выступают отдельные методы), существуют определенные связи или формы взаимодействия. Выявление таких связей и форм взаимодействия с целью формирования процедуры по применению определенной их совокупности будем называть системной технологией выбора методов генерации и анализа сценариев [3].

Математически задачу выбора (В) можно описать парой:  $B = \langle M, O \rangle$ , где априори задано  $M = \{M_i\}$ ,  $i = 1, n$  – множество исходных методов генерации и анализа сценариев (М), О – принцип оптимальности, задаваемый лицом, принимающим решение (ЛПР) из совокупности требований по допустимости ( $O_{\text{доп.}}$ ) и критериальных требований (далее критериев), причем  $O = O_{\text{доп.}} \cup O_{\text{кр.}}$ ,  $O_{\text{доп.}} \cap O_{\text{кр.}} = \emptyset$ .

Требования по допустимости  $O_{\text{доп.}}$  представляют собой требования к характеристикам объектов (методов) выбора (В) посредством отношений R-типа:  $R = \{=, \neq, <, >, \leq, \geq, \langle, \rangle\}$ .

Критериями  $O_{\text{кр.}}$ , формируемые ЛПР, отображают его целевые устремления с учетом информационного обеспечения на данном этапе процедуры выбора. Решением задачи выбора (В) будем называть подмножество  $M_{\text{opt}} \in M$ , полученное с помощью принципа оптимальности О в следующей последовательности:

$$\begin{array}{ccc} M & \rightarrow & M_{\text{доп.}} \rightarrow M_{\text{opt.}} \\ & \uparrow & \uparrow \\ & O_{\text{доп.}} & O_{\text{кр.}} \end{array}$$

то есть вначале выделяется множество допустимых методов  $M_{\text{доп.}}$ , удовлетворяющих требованиям по допустимости, а затем уже на этом усеченном множестве проводится выбор методов по критериям  $O_{\text{кр.}}$ .

С учётом изложенного, рассмотрим один из возможных подходов системного применения статистических методов, при чём в качестве **критерия по допустимости** будем рассматривать априорные информационные уровни задания функции распределения вероятностей.

1. Задание строгих параметрических моделей распределений вероятностей, описываемых известными законами (например, нормальным законом) может быть обеспечено объединением

следующих информационных составляющих, например, в символах алгебры логики:

$$F \sim I_1 \wedge I_2 \wedge I_3 \wedge I_4,$$

где  $F$  – функция распределения вероятностей;  $I_1$  – информация об общей схеме (механизме) формирования случайных величин данной природы;  $I_2$  – область значений случайной величины;  $I_3$  – примеры реальных признаков, подчиняющихся данному закону;  $I_4$  – аналитическое задание (модель) закона и его определяющие параметры. В данной ситуации для обработки данных могут быть использованы **методы параметрической статистики**

**2. Задание полной окрестности искажённых параметрических моделей** распределений вероятности базируется на том, что распределения реально наблюдаемых случайных величин в подавляющем большинстве отличны от строгих параметрических моделей (например, от нормальных). Такие модели можно охарактеризовать как «искажённые» параметрические модели (наличие «хвостов», многомодальность, выраженные асимметрия или эксцесс).

Подобные искажения обязаны своим появлением наличием, как правило, в исследуемых данных определённой доли (обычно относительно небольшой) так называемых «засоряющих» значений. Формально такие модели представляются симметричными и несимметричными смесями распределений вероятностями вида [ 7 ]:

$$F(X) = (1 - \varepsilon)N_0(X; \bar{X}_0; \sigma_0^2) + \varepsilon N_1(X; \bar{X}_1; \sigma_1^2)$$

$$\bar{X}_0 = \bar{X}_1$$

$$F(X) = (1 - \varepsilon)N_0(X; \bar{X}_0; \sigma_0^2) + \varepsilon N_1(X; \bar{X}_1; \sigma_1^2; \theta),$$

$$\bar{X}_0 \neq \bar{X}_1$$

где  $N_0(X; \bar{X}_0; \sigma_0^2)$  - плотность основного нормального распределения со средним значением  $\bar{X}_0$  и дисперсией  $\sigma_0^2$ ;  $\varepsilon$  - доля «засоряющих» значений (уровень неоднородности), определяющих распределение  $N_1$  с соответствующими параметрами  $\bar{X}_1$  и  $\sigma_1^2$ ;  $\theta$  - параметр сдвига распределения  $N_1$  относительно  $N_0$ .

В данном случае целесообразно использовать **методы робастной (устойчивой) статистики** с целью получения несмещённых (малосмещённых) и эффективных оценок параметров положения и масштаба основного распределения  $N_0$ .

Рисунок 1 - Системное использование методов при построении сценариев

**3. Задание только различий между распределениями вероятностей** рассматривается в рамках **непараметрической статистики**, которая не делает предположение о принадлежности функции распределения тем или иным параметрическим семействам распределений, и её задачи формулируются в терминах только **различий** между классами или внутри класса неизвестных распределений.

Применение методов параметрической статистики при выполнении первого критерия по допустимости, сводится, как правило, к оцениванию параметров положения, масштаба и других числовых характеристик законов распределения (рис.2). Использование же методов робастной и непараметрической статистики при реализации двух других критериев требует наличия **критериев оптимальности**. Для первой группы методов в качестве таких критериев могут выступать смещённость и эффективность оценок, искажённость по асимметрии ( $A_s$ ) и эксцессу ( $E_x$ ) распределений. Критерии оптимального применения методов непараметрической статистики (как правило, методы проверки непараметрических статистических гипотез), определяются постановками основных типов непараметрических задач.

Рассмотрим одну из них, которая получила название задачи согласия. Пусть задано известное непрерывное распределение  $F(X)$ . Из неизвестного распределения  $G(X)$  берётся выборка  $X_1, X_2, \dots, X_n$  и формируются гипотезы:

- нулевая  $H_0 : F(X)=G(X)$
- альтернативная  $H_0 : F(X)\neq G(X)$

Для решения такой задачи используются, например, строгий непараметрический метод Колмогорова-Смирнова [5,6].

Всё изложенное, позволяет сформировать системную технологию применения статистических методов, структурная схема которой представлена на рис.2. Это позволит оптимизировать процесс работы экспертов при анализе сценариев.

Предложенное системное использование методов позволяет генерировать и эффективно оценивать разнообразную информацию при построении и анализе сценариев при исследованиях сложных систем различной природы.

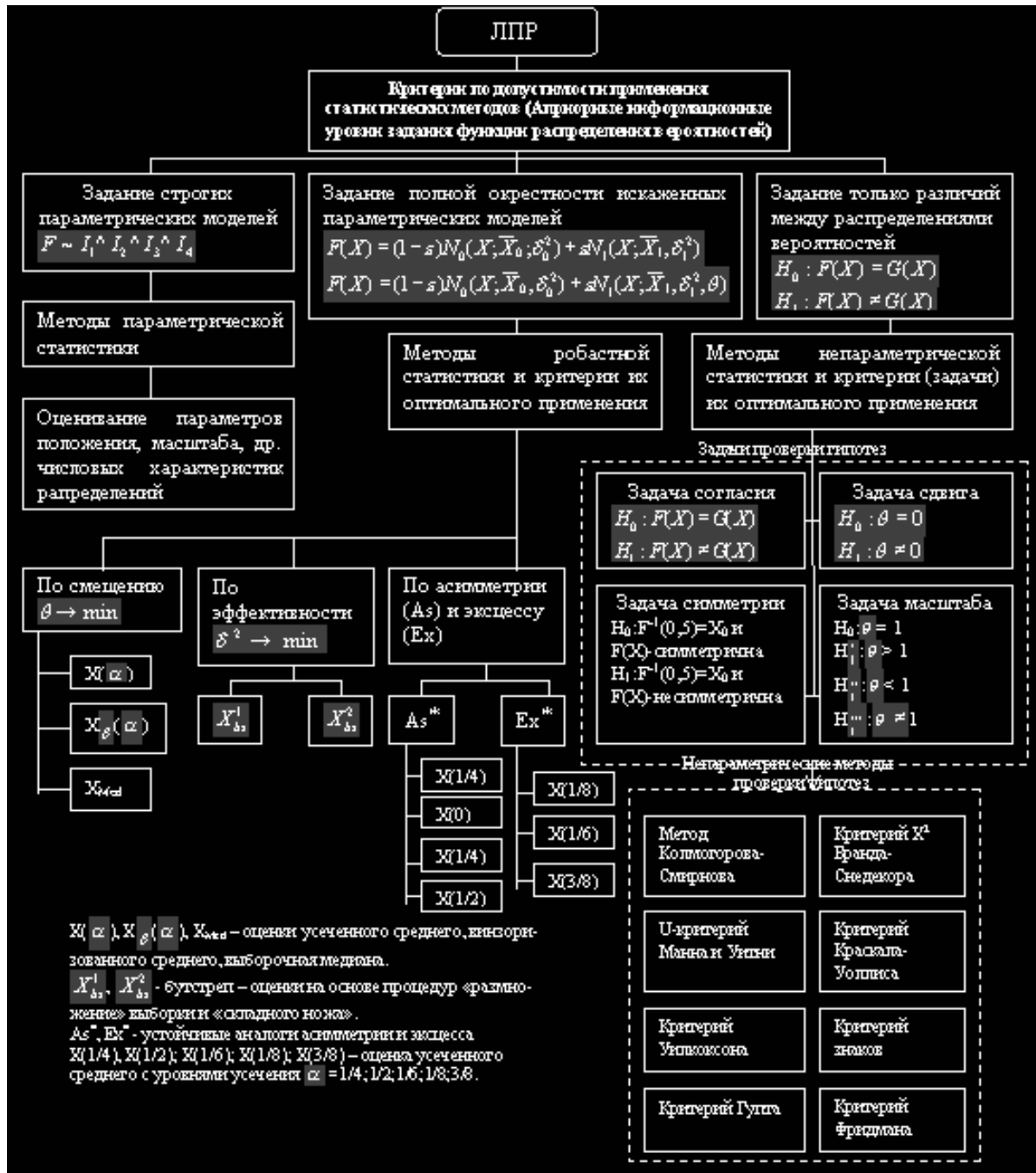


Рисунок 2 - Структурная схема системной технологии применения статистических методов в анализе сценариев

### ЛИТЕРАТУРА

1. М.З. Згуровський Системна методологія передбачення - Київ: “Політехніка”, 2001. – 49 с.
2. М.З.Згуровський, Н.Д. Панкратова Технологическое предвидение.- Київ:“Політехніка”, 2005. – 154 с.
3. Гожий А.П., Коваленко И.И. Системные технологии генерации и анализа сценариев // Автоматика, Автоматизация, Электротехнические комплексы и системы, № 2(16), 2005.- 89-96 с.
4. И.И.Коваленко, А.П. Гожий, Т.В. Пономаренко Системный подход к выбору методов анализа сложных физических систем// Материалы

международной научн. Конференции ISDMIT-2005, Евпатория, 2005, с.75-78.

5. Тарасенко Ф.П. Непараметрическая статистика. – Томск.: Из-во ТГУ,1976.
6. Холлендер М., Вулф Д. А. Непараметрические методы статистики. - М.: Финансы и статистика, 1983.
7. Хьюбер П. Робастность в статистике: Пер. с англ.-М.: Мир, 1984. - 304 с

Получено \_\_.\_\_. 2006 г.