Технические науки

УДК 51-7:612.16

Аббаскулиев Айдын Сахим оглы

кандидат технических наук, доцент,

доцент кафедры Фундаментальных наук

Высшая военная школа Азербайджана имени Гейдара Алиева

AbbasguliyevAydin

Phylosophy Doctor on Technical Sciences, Docent,

Docent of the Department of Fundamental Sciences

Azerbaijan Higher Military School named after Heydar Aliyev

Гулиева Афет Исмаил кызы

старший преподаватель кафедры Фундаментальных наук

Высшая военная школа Азербайджана имени Гейдара Алиева

Guliyeva Afet

Senior Teacher of the Department of Fundamental Sciences

Azerbaijan Higher Military School named after Heydar Aliyev

Исаева Дурдана Севдигюл кызы

преподаватель кафедры Фундаментальных наук

Высшая военная школа Азербайджана имени Гейдара Алиева

Isayeva Durdana

Teacher of the Department of Fundamental Sciences

Azerbaijan Higher Military School named after Heydar Aliyev

Назарова Айнур Наджми кызы

преподаватель кафедры Фундаментальных наук

Высшая военная школа Азербайджана имени Гейдара Алиева

Nazarova Aynur

Teacher of the Department of Fundamental Sciences

Azerbaijan Higher Military School named after Heydar Aliyev

ОБ ОДНОМ ПОДХОДЕ К ПОСТРОЕНИЮ СТАТИСТИЧЕСКИХ ГРАФИКОВ
ABOUT ONE APPROACH TO BUILDING STATISTICAL GRAPHS

Аннотация. *Статья посвящена исследованию статистических графиков с нечеткими значениями. Рассмотрена задача оценки степени обьективности выбранного алгоритма применительно к статистическому таблицу. С этой целью предлагается использовать математического аппарата «if A then B else C». Проводится анализ полученных результатов. Даются рекомендации по применению предлагаемого алгоритма.*

Ключевые слова: *статистические таблицы, статистические графики, нечеткий алгоритм, лица, принимающего решения, диалог лингвистических переменных.*

Summary. *This article deals with the investigation of the statistical qraphic with fizzy (illegible) meaning. It has been looked through the assessment task of the objectivity level of selected alqorithm can be applied to the statististical table.For this purpose it is affered to use mathematical apparatus "If A then B else C". It carried out the analysis of the received results. It is given a recommendation to accept offering alqorithm.*

Key words: *Statistical table, statistical graphic, fuzzy algorithm, figures, accepting decisions, variable linguistical dialogue.*

Введение. Статистические графики обычно строят в качестве дополнения к статистическим таблицам, но могут быть построены и самостоятельно. Они представляют собой условные изображения статистических данных посредством линий, геометрических фигур, рисунков или картосхем.

Интеллектуальная обработка статистических графиков, которые играют очень важную роль при исследовании статистических данных системы, требует учета человеческого фактора. Другими словами, оптимизация должна происходить при участии лица, принимающего решения (ЛПР).

В настоящее время разработаны пакеты прикладных программ компьютерной графики, которые значительно облегчают решение задач статистического исследования. Наиболее часто используемыми пакетами прикладных программ являются: Microsoft Excel; STATISTIKA 6.0; Harvard-graphics. Имеющиеся недостатки при обработке статистических данных системы с помощью этих пакет прикладных программ, можно устранить с помощью математического аппарата – нечеткого алгоритма. Это дало возможность вести интеллектуальную обработку статистических данных системы.

Известно, что используемые при интеллектуальной обработке статистических данных системы, лингвистические данные более информативны, чем обычные данные; параметры, характеризующие статистических данных системы имеют большой спектр термов (норма, ниже нормы, очень ниже нормы, выше нормы, очень выше нормы, и т.д.).

При принятии решения это хорошая база для применения компьютерной диалоговой системы человеком, определенные нарушения в системе статистических данных выдаются на языке лингвистических переменных. Кроме того, если полученные с помощью математического аппарата « if A then B else C » результаты нас не устраивают с точки зрения точности, тогда множество лингвистических термов расширяется, т.е. увеличивается количество критериев с помощью экспликация (словесное описание содержания графика – название графика, подписи к масштабным шкалам и пояснения к отдельным частям графика).

Кроме экспликации элементами графика являются графический образ (столбики, линии, точки и т.д.), поле графика, пространственные ориентиры (система координат), масштабные ориентиры**.**

Статистические графики классифицируются по задачам, решаемым с их помощью, а так же – по форме графического образа.

Прежде, чем приступать к построению графика, необходимо чётко уяснить поставленную перед собой задачу и затем только выбрать графический образ, который лучше всего будет способствовать её решению.

Существует следующие виды статистических графиков по поставленным задачам: диаграммы сравнения, диаграммы динамики, диаграммы контроля выполнения плана, диаграммы структуры, графики вариационных рядов, графики зависимости варьирующих признаков, статистические карты.

Очень важно выбор вид статистических графиков (A) по форме графического образа (B). Здесь:

A: Линейные графики + Плоскостные графики + Объёмные графики.

B: статистические кривые + столбиковые + полосовые + секторные + круговые + квадратные + прямоугольные + фигурные + точечные + фоновые + поверхностные распределения (пирамиды, параллелепипеды, сегментные графики).

Построим фрагмент алгоритма выбора вида статистического графика с помощью математического аппарата if A then B else C.

if «форма графического образа» = «статистистические кривые»

then «вид графика» должен быть «линейные» else

if «форма графического образа» = «фигурные»

then «вид графика» должен быть «плоскостные» else

if «форма графического образа» = «сегментные графики»

then «вид графика» должен быть «объемные».

Более всего распространённой системой координат при построении статистических графиков является система прямоугольных координат. При этом оптимальным соотношением сторон поля графика по осям абсцисс и ординат для зрительного восприятия считается соотношение, равное от 1,5:1,00 до 1,3:1,00, которое называется «правилом золотого сечения».

Введём исходную информацию для построения графиков. На базе исходной информации, а также дополнительной, сделаем постановку отдельных заданий. Затем представим методику построения всех видов графиков.

Для сравнения однотипных объектов по одноимённым признакам используются диаграммы сравнения. По форме графического образа здесь чаще всего используются диаграммы: столбиковые; полосовые (ленточные); фигурные и реже: круговые; квадратные; прямоугольные.

Столбики символизируют собой сравниваемые объекты, строятся на горизонтальной оси. Количество столбиков определяется числом сравниваемых объектов. Ширина столбика может быть произвольной, но обязательно одинаковой для всех столбиков. Высота столбика строится в соответствии с масштабной шкалой, построенной на вертикальной оси, и отражает величину изучаемого показателя. Числа, характеризующие величину показателя, помещаются внутри каждого столбика или над ним.

В отличие от предыдущей в полосовых диаграммах сравнения столбики строятся на вертикальной оси – оси ординат. Полосовая (ленточная) диаграмма представляет ряд простирающихся вдоль оси абсцисс полос одинаковой ширины.

При построении фигурных диаграмм статистические данные изображаются в виде рисунков-символов, которые в наилучшей степени отражают сущность изображаемого явления (⌂, Δ, ۞ и т.п.). Здесь каждому знаку-символу условно придаётся определенное числовое значение, и, путем последовательного их расположения на поле графика,

формируются соответствующие «полосы». Величина отображаемого показателя определяется количеством стандартных знаков в каждой «полосе». Иногда для каждого сравниваемого показателя строят по одному условному знаку-символу, но разной величины – пропорциональной величине изображаемого показателя. Фигурные диаграммы более выразительны, наглядны, легко воспринимаются, и поэтому их часто применяют в рекламах.

Круговые, квадратные и прямоугольные диаграммы – это диаграммы, принцип построения которых состоит в том, что сравниваемые величины изображаются в виде правильных геометрических фигур, которые строятся так, чтобы площади их соотносились между собой как количества, этими фигурами отображаемые. Таким образом, эти диаграммы выражают величину изображаемого показателя размером своей площади.

При построении круговых или квадратных диаграмм используют геометрические фигуры: круг и квадрат. Известно, что площадь круга равна πr^2 (r – радиус круга, π – постоянная величина, приблизительно равная 3,14), а площадь квадрата равна квадрату его сторон. Для построения этого типа диаграмм необходимо сначала путём несложных арифметических действий найти радиус круга или сторону квадрата. Затем на базе полученных данных и в соответствии с принятым масштабом – строить график. При этом квадраты или круги необходимо расположить на одинаковом друг от друга расстоянии, а в каждой фигуре (или над ней) указать числовое значение, которое она изображает.

К рассматриваемому виду диаграмм относится графическое изображение, полученное путем построения один в другом кругов или квадратов.

Прямоугольные диаграммы строятся для показателей, получаемых путём умножения двух других. Тогда стороны прямоугольника будут

отражать эти два сомножителя, а его площадь – величину результативного показателя.

Диаграммы динамики характеризуют развитие явления во времени. Здесь чаще всего используются линейные диаграммы, а также – столбиковые. Реже используются полосовые, фигурные, квадратные, круговые, прямоугольные.

Линейные диаграммы строят в прямоугольной системе координат, причём на оси абсцисс откладываются показатели времени, а на оси ординат – значения изучаемого показателя. По отметкам точек обеих осей координат определяется положение каждого уровня на поле графика. Последовательно соединяя точки отрезками линий, получают статистическую кривую, характеризующую развитие изучаемого явления во времени.

Столбиковые, полосовые, фигурные, квадратные, круговые, прямоугольные диаграммы динамики строятся аналогично диаграммам сравнения; отличие заключается в том, что здесь сравниваются не разные явления, а одно и то же явление, но в разные периоды времени.

Диаграммы контроля выполнения плана используется при решении задач контроля выполнения плана, а также договорных обязательств, нормативов и т.п. С этой целью обычно строятся линейные диаграммы, подобные тому, которые были использованы при построении диаграмм динамики. Особенностью этого вида диаграмм является то, что на одном поле графика наносятся две линии: одна отражает накопительное с течением времени плановое задание, а вторая – фактическое исполнение этого плана.

Структурные диаграммы используются для характеристики состава статистической совокупности. Здесь чаще всего строят секторные диаграммы, реже – полосовые и столбиковые.

В секторных диаграммах площадь круга принимается за величину всей изучаемой статистической совокупности, а площади отдельных секторов отображают долю (удельный вес) её составных частей. Поскольку площади секторов пропорциональны их центральным углам, то при построении секторной диаграммы сумма всех углов (360°) распределяется пропорционально удельным весам отдельных частей изучаемой статистической совокупности. При процентном выражении состава изучаемой статистической совокупности исходят из отношения: 100% = 360°, тогда 1% = 3,6°.

Столбиковые и полосовые диаграммы структуры. Здесь при изображении состава статистической совокупности столбик (полоса) соответствует 100% (или единице). Если столбиковые (полосовые) диаграммы преследуют также цель сравнения структур различных объектов или одного и того же объекта, но в разные периоды времени, то размеры всех столбиков (полос) должны быть одинаковы; в этих диаграммах могут изменяться только составляющие столбика (полосы).

Графики вариационных рядов. Различного рода графики вариационных рядов строят с целью более подробного анализа статистических данных.

Полигон (многоугольник) – графическое изображение дискретного ряда распределения. На поле графика в системе прямоугольных координат наносятся точки, абсциссами которых являются варианты ряда распределения, а ординатами – частоты. Ломаная линия, соединяющая эти точки, и есть полигон.

Гистограмма (строение) – графическое изображение интервального ряда распределения, при этом интервальный ряд распределения изображается столбиковой диаграммой, в которой основания столбиков – это интервалы значений варьирующего признака, а высоты столбиков – частоты. В гистограмме между столбиками нет зазоров.

Графики зависимости варьирующих признаков строятся с целью визуального определения зависимостимежду двумя варьирующими признаками. Для этого строится поле корреляции. Поле корреляции – графическое изображение в системе прямоугольных координат точек, абсциссами которых являются значения факторного признака, а ординатами – значения результативного признака.

При использовании классических методов для построения оптимальных статистических таблиц, оптимальное решение общей задачи получается в виде выпуклой комбинации оптимальных допустимых решений подзадач. Другими словами, решения подзадач не дают в явном виде допустимое оптимальное решение общей задачи [1]. Этот принцип требует многократного обмена информацией, так как эта процедура обладает медленной сходимостью к оптимуму, и значит, она практически трудно реализуема.

Для частичного упрощения классического алгоритма решения задачи можно использовать обобщенный симплекс – метод. При этом решается задача, где допустимым множеством является линейная комбинация точек из Парето множества. Здесь перед использовании полученной информации происходит ее обобщение и в результате ускоряется сходимость алгоритма и уменьшается кратность обмена информацией. Для таких моделей характерны следующие недостатки. Во-первых размерность задачи оптимизации принятия решений очень большая. Во вторых, эта модель не подходит для решения задач с качественными критериями. В качественных критериях сравнение ведется с помощью отношений «больше», «меньше», «равно», «много», «мало» и т.д. Кроме этого указанная модель имеет еще и следующие недостатки:

- модель плохо адаптируется к изменению внешней ситуации;
- модель очень большая, что приводит к невысокой оперативности, а это уменьшает эффект от использования компьютера;

- модель не приспособлена для учета качественных факторов;
- модель не учитывает опыт, знания и интуицию лица, принимающего решение.

В данной статье нами предлагается модель, где с помощью теории нечетких множеств [2], получаем свободный от указанных недостатков алгоритм. Полученные в работе результаты анализируются на примере лингвистической таблицы статистики.

Постановка задачи: Определить такие соотношения показателей опроса, которые обеспечили бы минимизацию времени в принятии решения.

Задача принятия решения представляет собой задачу линейного программирования. Целевая функция и ограничения задачи оптимизации представляются в виде:

$$F_0(\overline{f_1}(x_1),..,\ \overline{f_N}(x_N)) \rightarrow \min \quad (1)$$

$$F_k(\overline{f_1}(x_1),..,\ \overline{f_N}(x_N)) \geq B_k \quad (2)$$

$$x_i \in X \quad (3)$$

где f_i – качественные показатели, X – множество полученных информационных данных, B_k - ограничения на качественные показатели. Здесь задача векторной оптимизации имеет следующий вид:

$$\sum_{i=1}^{n} x_i f_i(y) \rightarrow \max \quad (4)$$

где коэффициенты функций f_i, являются размытыми и харатеризуются функциями принадлежности

$$\mu_x(x_{ij}) : X^n \rightarrow [0,1] \quad (5)$$

Рассматриваемая задача решается методом последовательных приближений.

Для простоты все нечеткие множества в данной работе характеризованы единой формой функции принадлежности

$$\mu(x) = e^{-k|x-c|} \quad (6)$$

где x – элемент множества, k и c параметры функции принадлежности. Параметры k и c вычисляются для каждой конкретной нечеткой переменной.

Метод решения. Нечеткий алгоритм состоит из следующих этапов:

1. Решение задачи оптимизации при нечетких весовых коэффициентах. В результате решения задачи определяются компромиссные значения показателей статистических таблиц.
2. С помощью лингвистического протокола определение лингвистического аналога полученных значений.
3. Решение задачи линейного программирования относительно весовых коэффициентов (элементов таблицы) оптимизации.
4. Преобразование полученных от решения задачи оптимизации результатов в лингвистический аналог.
5. Определение значений показателей, обеспечивающих требования лица, принимающего решения.

Применение теории нечетких множеств для решения задачи построения статистических таблиц объясняется тем, что человек принимает решение с помощью лингвистических высказываний без всяких затруднений.

Апробация нечеткого алгоритма на примере построения статистических таблиц с помощью предлагаемого метода [3] очень приемлемы для практики. Отметим, что до сих пор применяемые для решения этой проблемы методы и подходы не отвечают поставленным требованиям. Это связано тем, что в начальных исследованиях приходим к таким выводам:

- в процессе построения статистических карт (статистические карты - контурные карты, на которых статистические показатели изображаются с помощью графических образов) полученные

одинаковые данные по параметрам формы, интенсивности, частоты и времени разные. Например, картограммы плотности населения. (Картограмма – контурная карта, на которой отдельные территории в зависимости от величины изображаемого показателя обозначаются с помощью: штриховки, расцветки (фоновые картограммы); точек (точечные картограммы). Картодиаграмма – сочетание контурной карты с фигурной диаграммой. Здесь геометрические символы, которые используются, располагаются не в один ряд (как в фигурных диаграммах), а размещаются по всей карте);

– с точки зрения анализа основная проблема, это нахождение элементов таблицы в очень маленьком диапазоне. Например, кумулята (кумулятивная кривая) строится по данным о накопленных частотах, которые откладываются на оси ординат, при этом на оси абсцисс откладываются значения исследуемого признака интервального ряда распределения. Нижней границе первого интервала соответствует частота, равная нулю, а верхней границе – частота данного интервала. Верхней границе 2-го интервала соответствует накопленная частота этого интервала и т.д. Огива – строится подобно кумуляте, только в ней названия осей меняются местами. Кривая концентрации (кривая Лоренца) – строится для изучения степени неравномерности распределения определённого суммарного показателя между отдельными группами вариационного ряда. Кривая Лоренца начинается с нуля, а заканчивается в точке с координатами (100%;100%). При равномерном распределении кривая концентрации совпадёт с прямой распределения – прямой, которая делит угол пополам. Для неравномерного распределения характерно приближение кривой концентрации к оси абцисс;
– особенности статистических таблиц как и при выборе элементов, так и при цифровой обработке, требуют расширенного применения

статистических методов и подходов. Это связано с тем, что при одинаковых условиях в зависимости от применения соответствующего метода получаем разные результаты. Например, при увеличении числа наблюдений за одной и той же статистической совокупностью увеличивается число групп интервального ряда распределения. Это приводит к уменьшению величины интервала. Тогда ломаная линия, связывающая центры вершин столбиков (тоже – полигон) превращается в плавную кривую, называемую кривой распределения;

– из-за неадекватности существующих моделей в процессе принятие решения присутствуют нечеткости. Поэтому в условиях «недостаточной» статистики целесообразно использовать нечеткий алгоритм диагностики.

Успех нечеткого алгоритма во многом зависит от того, будут ли заданные вопросы достаточно естественными для лица, принимающего решения. Лицу, принимающему решение, предъявляем два варианта (V1 и V2) и просим их сравнить, используя следующий набор лингвистических выражений:

– V1 или V2 лучше сказать не могу;
– V1 или V2 хуже сказать не могу;
– V1 несколько лучше V2;
– V1 несколько хуже V2;
– V1 заметно лучше V2;
– V1 заметно хуже V2;
– V1 намного лучше V2;
– V1 намного хуже V2.

Диалог продолжается до тех пор, пока предъявляемые для сравнения варианты не будут эквивалентны с точки зрения ЛПР.

Выводы. Используя диалог лингвистических переменных, ЛПР делает свои высказывания более естественными, потому что в словах информации намного больше, чем в цифрах.

Это дает возможность расширить диапазон. Другими словами, в зависимости от ситуации, когда имеется не один, а несколько значений, можно принимать более точное решение.

Увеличивается количество вариаций, т.е. для разных пациентов по одной и той же патологии можно выбрать разные комбинации.

Литература

1. Лэсдон Л.С. Оптимизация больших систем. М: Наука. 1975.
2. Заде Л. Понятие лингвистических переменной и ее применение к принятию приближенных решений. М. : Мир, 1976.
3. Аббаскулиев А.С. Числовые характеристики случайных величин с нечеткими значениями., Международный научный журнал «Интернаука». URL: http://www.inter-nauka.com/issues/2017/15/3079
4. Аббаскулиев А.С. Об одном алгоритме вычисления характеристик случайных чисел с нечеткими значениями. Международный научный журнал «Интернаука». URL: http://www.inter-nauka.com/issues/2018/5/45