

В случае $\pi=P$ получаем

Следствие 1. Пересечение любого множества разрешимых классов Фиттинга, нормальных в классе Фишера X является X -нормальным классом Фиттинга.

При $X=S$ специальным случаем следствия 1 является основной результат [1], который сформулируем как

Следствие 2 (Блессеноль, Гашюц [1]). Пересечение любого множества неединичных разрешимых нормальных классов Фиттинга является неединичным нормальным классом Фиттинга.

Список литературы

1. Blessenohl D., Gaschütz W. Über normale Schunk und Fittingklassen // Math. Z. 1970. Bd. 118. S. 1-8.

ПРИМЕНЕНИЕ МЕТОДОВ НЕПАРАМЕТРИЧЕСКОЙ СТАТИСТИКИ В УПРАВЛЕНИИ КАЧЕСТВОМ ПРОИЗВОДСТВЕННЫХ ПРОЦЕССОВ

З.Ю. Гордеева, А.И. Бочкин*

**Витебск, УО ФПБ ВФ «Международный университет «МИТСО»*

Витебск, ВГУ имени П.М. Машерова

Сегодня качество продукции – решающий фактор в условиях рыночной экономики. Оно во многом зависит от решений, принимаемых при управлении производством. В то же время при принятии управленческих решений, в том числе связанных с рисками, нередко традиционно используются интуитивный подход, без необходимого математико-статистического обоснования. Поэтому является актуальной задача повышения качества управленческих и технологических решений в условиях применения возможностей ЭВМ. Эра статистических таблиц миновала, статистика стала потенциально более доступной для пользователя, но дистанция между статистикой и нуждами практики еще сохраняется. Поэтому целями работы являются:

- создание простого и удобного пакета статистических программ, обеспечивающих принятие оптимальных и обоснованных решений;
- реализация новых подходов в статистике в связи с применением ЭВМ (непараметрические методы);
- анализ и предложения по прямому применению конкретных статистических критериев к практическим производственным ситуациям.

Материал и методы. Исходным материалом прикладной статистики вообще и обсуждаемого пакета программ являются данные наблюдений, взятые из реальных ситуаций производства, образования и управления процессами. Методы непараметрической статистики, реализованные в данном пакете программ, отобраны с целью перекрыть максимальное число проблемных ситуаций при минимальной подготовке пользователя.

Результаты и их обсуждение. В связи с происходящей компьютеризацией производства и управления лицу, принимающему решение (ЛПР) достаточно в общих чертах понимать логику и реализованные методы статистики в компьютерной программе, подготавливающей данные для его решения и уметь применять ее на уровне параметрического пользователя – менять параметры, описывающие ситуацию, требующую принятия (или непринятия) управляющего решения, уметь пользоваться методом численного эксперимента.

Многолетний опыт применения математической статистики показал недостаточность классических методов, основанных на нормальном законе распределения вероятностей. Эти методы ведут порой к вычислению средней температуры по больнице или среднего заработка уборщиков и миллионера. Стала очевидной необходимость развития и применения непараметрической статистики – свободной от вида распределений случайных величин. Ее основное понятие – не величина, а ранг – номер наблюдения в упорядоченной выборке. Заработок уборщика и миллионера отличается номером в выборке.

Сегодня непараметрическая статистика располагает группой методов и критериев, в общем достаточных для обеспечения оптимальных управленческих решений. Как показывает практика, адекватным средством реализации ее методов могут быть электронные таблицы типа EXCEL, обеспечивающие разведочный анализ имеющихся данных в стиле “что-если”, практическое моделирование на ЭВМ производственных процессов.

Рассмотрим ряд критериев и методов непараметрической статистики, полезных для управления качеством принимаемых решений.

Критерий знаков вычисляет отклонение наблюдений от равновероятного двузначного (+/-) распределения. Может быть полезен при анализе наличия систематической ошибки в настройке станка – устойчивый перевес допусков в размере изделия со знаком «+» или «-» от стандартного значения.

Критерий Бернулли позволяет быстро оценить согласованность доли брака в выборке с разрешенными нормами долей брака. Легко оценивается практически объем выборки, необходимый для обоснованного решения об приемке или отклонении партии изделий.

Ранговая корреляция по Кендаллу позволяет оценить наличие зависимости между величинами любых типов, например, между днем недели и производительностью труда, ценой однотипных изделий и числом покупок.

Критерий однородности по Манну-Уитни позволяет сравнивать выборки величин любого типа и сравнивать, например производительность рабочих, сортность продукции, доли брака и так далее.

Критерий сопряженности (точный) Фишера позволяет сравнивать частоты наблюдений любой природы, например, производительность труда курящих и некурящих рабочих и принимать решение о стимуляции здорового образа жизни. В случае нескольких пар частот наблюдений позволяет отловить немонотонные зависимости, например производительность труда в зависимости от возраста. Критерий не требует высоких частот наблюдений в отличие от критерия хи-квадрат.

Классическая лемма Маркова при известных границах наблюдений позволяет оценивать вероятность редких событий – пиковое возрастание заявок на склад, случайное скопление пассажиров на остановке.

Формула Гнеденко [1,с.160] позволяет оценить количество необходимого резервного оборудования для оперативной замены оборудования такого же типа с неабсолютной надежностью и принять обоснованное решение о мере риска в случае создания такого резерва.

Таким образом, непараметрическая статистика обеспечивает обработку данных в условиях неопределенности с минимальными предпосылками о характере искажения случайными факторами управляемых процессов, что ведет к снижению компоненты риска и оптимизацию принимаемых решений.

В результате разработан пакет прикладных программ для ЭВМ в среде Excel для пользователей с начальной подготовкой в области непараметрической статистики; реализован точный расчет доверительных вероятностей (в стиле крите-

рия Фишера); рассмотрены конкретные обучающие примеры применения программ, реализующих статистические критерии. Результатом работы является также адаптация методов непараметрической статистики к применению в управлении качеством конкретных производственных процессов.

Заключение. Реализованные в виде пакета программ для ЭВМ рассмотренные статистические методы предназначены для практического применения инженерами предприятий по качеству, в управлении процессами (в том числе в образовании), для информационного обеспечения оптимальных управленческих решений.

Список литературы

1. Гнеденко Б.В., Хинчин А.Я. Элементарное введение в теорию вероятностей. – М.: «Наука», 1970. – 168 с.

ПРИМЕНЕНИЕ КЛАСТЕРНОГО АНАЛИЗА В ПОСТРОЕНИИ ПОИСКОВЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

*С.А. Ермоченко
Витебск, ВГУ имени П.М. Машерова*

В настоящее время методы кластерного анализа в поисковых системах нашли очень широкое применение. Основной задачей поисковой системы является поиск тех документов среди достаточно большого массива имеющихся документов, которые более всего удовлетворяли бы вводимым пользователем ключевым словам. При этом кластерный анализ применяется для разбиения множества документов на такие подмножества, чтобы документы, наиболее удовлетворяющие критериям поиска, попадали в одно подмножество, или кластер [1].

Основным понятием, применяемым в кластерном анализе, является понятие метрики $\rho(d_n, d_m)$ как некоторого расстояния между двумя документами d_n и d_m . При этом такое расстояние должно быть тем меньше, чем более близки показатели релевантности документов к критерию поиска. Наиболее частым способом определения такого расстояния является сравнение весовых коэффициентов каждого документа по отношению к поисковым термам (ключевым словам). Конкретных способов вычисления таких коэффициентов существует несколько: весовой, вероятностный, семантический и т. д.

Однако для всех этих методик характерна одна общая черта: они предполагают однозначность поискового запроса (набора ключевых слов). Под однозначностью поискового запроса предполагается, что пользователю, составившему поисковый запрос, два документа, одинаково релевантных этому запросу, равно интересны. Но это не всегда так.

Целью данной работы является адаптация методов кластерного анализа для кластеризации результатов поиска, равнорелевантных поисковому запросу, по принадлежности документов к различным тематическим категориям. Такая кластеризация позволит представить результаты поиска не только в виде линейного списка, но и в виде иерархической структуры, в узлах которой располагаются различные категории и подкатегории.

Материал и методы. В качестве основного метода исследования используется сравнительно-сопоставительный анализ имеющихся методов кластерного анализа и их применимости к результатам поисковой выдачи.