

Кобылянский Б.Б.

Учебно-научный профессионально-педагогический институт
Украинской инженерно-педагогической академии

Мнухин А.Г.

Запорожская государственная инженерная академия

НЕПАРАМЕТРИЧЕСКАЯ СТАТИСТИКА ДЛЯ ОЦЕНКИ ОСОБЕННОСТЕЙ ГОРНОГО ПРОИЗВОДСТВА

На современном этапе технического развития все комплексные системы промышленного производства (в частности, угольной промышленности) требуют автоматизации всего процесса, включающего работу добычного и транспортного оборудования, что нельзя выполнить без детального математического описания всех специфических явлений, характеризующих конкретное производство. Поэтому впервые для решения задачи управления сложными технологическими, многофакторными системами предлагается использование непараметрической статистики, которая в отечественной специальной литературе не нашла должного отражения. И тем не менее, именно такие математические подходы позволяют выполнить корректный прогноз рассматриваемых явлений за пределами имеющихся экспериментальных данных с достаточной для решения большинства практических задач точностью, особенно в тех случаях, когда число самих данных относительно невелико.

Ключевые слова: угольная промышленность, оценка, непараметрическая статистика.

Постановка проблемы. На современном этапе предприятия угольной промышленности оснащены автоматизированными линиями и комплексами, внедряются современные технологические комплексы и участки, гибкие производственные системы. Все это способствует решению двух взаимосвязанных задач, таких как: выпуск более качественной продукции и повышение безопасности производственного процесса. Охрана труда, будучи системой специальных знаний, предназначена в качестве средства обеспечения безопасности технологических процессов и производства [1]. Совершенствованию системы управления охраной труда различных предприятий постоянно уделяется внимание ученых Украины [3–6]. Одним из основных способов оценки техногенного риска на предприятиях повышенной опасности, в частности угольных, является исследование уровня возможного травматизма.

Изложение основного материала. Известно, что высокой аварийности и травматизму на шахтах Украины, особенно на Донбассе, способствуют сложнейшие горно-геологические условия разработки угольных пластов, устаревший шахтный фонд и горно-шахтное оборудование, а также целый ряд других факторов, которые нельзя

не учитывать при анализе травматизма. Известен также такой факт, что изучение различных оди-ночных факторов и их групп требуют различных статистических подходов и математических методов для их практической реализации.

Число шахт с особо опасными условиями труда представлено в табл. 1, из которой видно, что в угольной промышленности Украины 89 % шахт газовые, 60 % – опасные по взрывчатости угольной пыли, 45,2 % – опасные по газодинамическим явлениям (внезапные выбросы угля, породы и газа, горные удары) и 22 % – по самовозгоранию угля.

Вместе с тем анализ причин травматизма и аварийности свидетельствует о значительном удельном весе среди причин этих происшествий человеческого фактора, особенно аварий с групповым травматизмом. Это подтверждают материалы расследования аварий 1 и 2 категорий с групповыми несчастными случаями. Данные материалов расследования аварий и травматизма свидетельствуют о том, что среди их причин значительную роль играет нарушение правил техники безопасности. Для установления и группировки основных причин и тенденции этих явлений необходимо провести более глубокое изучение условий,

Таблиця 1

Число шахт с особо опасными условиями труда

Регион	Общее число шахт с особо опасными условиями труда	Количество шахт – технических единиц по видам опасности										Взрывчатая угольная пыль	Самовозгорание угля
		Выделение метана						Горные удары	Взрывчатая угольная пыль	Самовозгорание угля			
		Всего шахт	в т.ч. по категориям			Внезапные выбросы угля и газа							
			I	II	III								
Донецкая обл.	105	47	10	5	3	29	53	13	65	27			
Луганская обл.	70	34	5	2	4	23	20	-	34	11			
Другие области	15	15	4	4	4	3	-	-	15	5			
Украина	190	96	19	11	11	55	73	13	114	43			
Удельный вес шахт, %	100,0	50,6	10,0	5,8	5,8	29,8	38,4	6,8	60,0	22,6			

Таблиця 2

Среднее годовое число аварий и простоев на шахтах Украины

Причины аварий и простоев	Число случаев		%	Потери	
	Всего			Рабочего времени, ч	Добычи, тыс. т.
Аварии и простои эксплуатационные в т.ч. завалы (обрушения) в лавах	29440		26,2	208859	3745
Отказы горно-шахтного оборудования в т.ч. забойного оборудования	6394			43803	832
Аварии и простои на главных подъемных установках	34800		31,0	178757	4023
Аварии (простои) на шахтном транспорте	28062			145436	3339
Аварии (простои) на конвейерном транспорте	548		0,4	5114	111
Остановки по предписанию инспекции и участка вентиляции и техники безопасности (ВТБ)	19005		16,9	97028	2471
Прочие опасности, всего	12871			59006	1934
Прочие опасности, всего	4162		3,8	44113	969
в т.ч. необеспеченность крепежным лесом	24314		21,7	208887	3335
Остановки вентиляторов главного проветривания	1817			9384	212
Загазованность выработок	211			7033	142
Загрязнение (подтопление) выработок	143			2321	24
Всего	432			6991	197
Всего	112269		100,0	742758	14654

в которых традиционные методы элементарной статистики оказываются недостаточно эффективными.

Химическая и горная промышленность, военные науки, планирование сколько-нибудь серьезных процессов и явлений – везде используются методы математической статистики высокого уровня, базирующиеся на использовании теории распределений, и, в первую очередь, нормальных и квазинормальных распределений. Однако становится очевидным, что по мере усложнения характера протекания рассматриваемых процессов, т. е. при вынужденном использовании других несимметричных распределений, в частности Стьюдента или Колмогорова, для описания массивов экспериментальных данных или их выборок традиционных методов статистики становится уже недостаточно. Даже при полном их соответствии решению ряда задач горной электротехники [2], использование их для прогноза состояния эргатических систем, например, предприятий угольной промышленности, они не позволяют дать корректную оценку рассматриваемому явлению и выполнить прогноз поведения системы в условиях, отличных от рассматриваемых.

В мире ежегодно травмируется до 200 тыс. шахтеров, в том числе 6-10 тыс. человек со смертельным исходом. Наибольшее число травм происходит на шахтах Китайской народной республики, где ежедневно в результате несчастных случаев на угольных шахтах гибнет более 10 шахтеров, а в 1995 году число погибших в этой стране превышало 10 тыс. 400 горняков.

За последние 30 лет даже на относительно малоаварийных польских шахтах произошло 15 взрывов метано-воздушной смеси, при этом пострадало 426 человек, в том числе 219 – смертельно [7]. В этот период имели место 17 катастрофических взрывов газа и пыли на шахтах Российской Федерации, Югославии (4), Японии (5), ЮАР (3), а также в ФРГ, Турции, Перу, Тайване, Мозамбике и др. [7].

Вместе с тем в ведущих угледобывающих странах в послевоенные годы число аварий и катастроф значительно снизилось. Обобщенные данные по соответствующей динамике смертельного травматизма в первой десятке угледобывающих стран представлены в табл. 2, вследствие чего для соответствующего анализа упомянутых данных, характеризуемых малым объемом, требуются иные математические методы, и, в частности, непараметрической статистики.

Для того чтобы правильно понять идеи *непараметрической* статистики (термин был впервые введен Wolfowitz, 1942), следует познакомиться с идеями *параметрической* статистики. Поэтому изначально следует ознакомиться с понятием статистической значимости критерия, основанного на распределении определенной статистики (вы можете просмотреть эту главу, прежде чем продолжить чтение). Говоря кратко, если вы знаете распределение наблюдаемой переменной, то можете предсказать, как в повторных выборках равного объема будет «вести себя» используемая статистика – т.е. каким образом она будет распределена. Пусть, например, имеется 100 случайных выборок, из одной популяции по 100 взрослых человек в каждой. Вычислим среднюю квалификацию (возраст или стаж) субъектов в каждой выборке, т.е. построим выборочное среднее. Тогда распределение выборочных средних можно хорошо аппроксимировать нормальным распределением (более точно, t -распределением Стьюдента с 99 степенями свободы). Теперь представьте, что случайным образом извлечена еще одна выборка из работников некоего предприятия, где, по вашим представлениям, проживают люди с возрастом выше среднего. Если средний возраст людей в этой выборке попадает в верхнюю 95% критическую область t распределения, то можно сделать обоснованный вывод, что работники этого объединения, действительно, в среднем более высокую квалификацию (чем в целом по промышленности), т.е. что это действительно наиболее хорошо функционирующее предприятие.

Возникает вопрос: «Действительно ли большинство переменных имеют нормальное распределение? В рассмотренном примере использовался тот факт, что в повторных выборках равного объема средние значения (возраста людей) будут иметь t -распределение (с определенным средним и дисперсией). Однако это верно, если рассматриваемая переменная имеет нормальное распределение, т.е. что распределение людей определенной квалификации нормально распределено (рис. 1).

Одним из факторов, ограничивающих применения критериев, основанных на предположении нормальности, является объем выборки. До тех пор пока выборка достаточно большая (например, 100 или больше наблюдений), можно считать, что выборочное распределение нормально, даже если вы не уверены, что распределение переменной в популяции является нормальным. Тем не менее, если выборка мала, эти критерии следует использовать только при наличии уверенности,

что переменная действительно имеет нормальное распределение. Однако нет способа проверить это предположение на малой выборке.

Использование критериев, основанных на предположении нормальности, кроме того, ограничено шкалой измерений. Такие статистические методы, как t-критерий, регрессия и т. д. предполагают, что исходные данные непрерывны. Однако имеются ситуации, когда данные, скорее, просто ранжированы (измерены в порядковой шкале), чем измерены точно.

Типичный пример дают группы данных: первую позицию занимает группа с максимальным числом рабочих конкретной специальности, вторую позицию занимает группа с максимальным числом рабочих среди оставшихся групп (среди групп, из которых удалена первая группа) и т. д. Зная рейтинги, а рабочих одной из групп больше числа рабочих другой, но насколько больше, сказать уже нельзя. Представьте, вы имеете 5 групп: А, В, С, D, E, которые располагаются на 5 первых мест. Пусть в текущем месяце мы имели следующую расстановку: А, В, С, D, E, а в предыдущем месяце: D, E, А, В, С. Спрашивается, произошли существенные изменения в рейтингах групп или нет? В данной ситуации, очевидно, мы не можем использовать t-критерий, чтобы сравнить эти две группы данных, и переходим в область специфических вероятностных вычислений (а любой статистический критерий содержит в себе вероятностную калькуляцию!). Мы рассуждаем примерно следующим образом: насколько велика вероятность того, что отличие в двух расстановках групп вызвано чисто случайными причинами или это отличие слишком велико и не может быть объяснено за счет чистой случайности. В этих рассуждениях мы используем лишь ранги или перестановки групп и никак не используем кон-

кретный вид распределения числа объектов изучения на них.

Для анализа малых выборок и для данных, измеренных в бедных шкалах, применяют непараметрические методы, и в частности, краткий обзор непараметрических процедур. По существу, для каждого параметрического критерия имеется, по крайней мере, одна непараметрическая альтернатива.

В целом подход к статистическим критериям в анализе данных должен быть прагматическим и не отягощен лишними теоретическими рассуждениями. Имея в своем распоряжении компьютер с системой STATISTICA, вы легко примените к своим данным несколько критериев. Зная о некоторых подводных камнях методов, вы путем простого экспериментирования выберете верное решение. Развитие сюжета довольно естественно: если нужно сравнить значения двух переменных, то вы используете t-критерий. Однако следует помнить, что он основан на предположении нормальности и равенстве дисперсий в каждой группе. Освобождение от этих предположений приводит к непараметрическим тестам, которые особенно полезны для малых выборок.

Далее имеются две ситуации, связанные с исходными данными: зависимые и независимые выборки, в которых применяется t-критерий для зависимых и независимых выборок соответственно.

Развитие t-критерия приводит к дисперсионному анализу, который используется, когда число сравниваемых групп (шахт, объединений) больше двух. Соответствующее развитие непараметрических процедур приводит к непараметрическому дисперсионному анализу, правда, существенно более бедному, чем классический дисперсионный анализ.

Для оценки зависимости, или, выражаясь несколько высокопарно, степени тесноты связи, вычисляют коэффициент корреляции Пирсона. Строго говоря, его применение имеет ограничения, связанные, например, с типом шкалы, в которой измерены данные, и нелинейностью зависимости, поэтому в качестве альтернативы используются также непараметрические, или так называемые ранговые, коэффициенты корреляции, применяемые, например, для ранжированных данных. Если данные измерены в номинальной шкале, то их естественно представлять в таблицах сопряженности, в которых используется критерий хи-квадрат Пирсона с различными вариациями и поправками на точность.

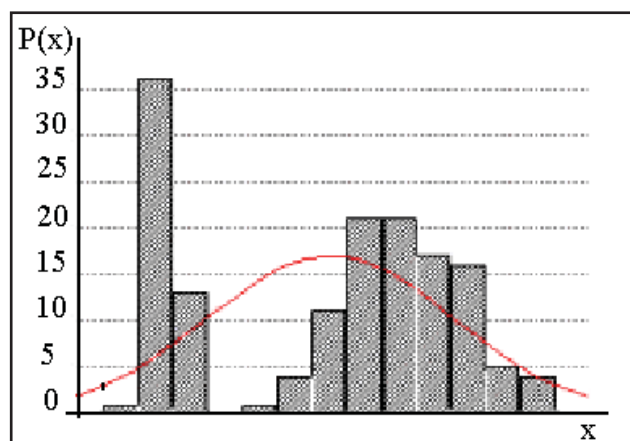


Рис. 1. Сопоставление законов распределения

Итак, по существу имеется всего несколько типов критериев и процедур, которые нужно знать и уметь использовать в зависимости от специфики данных. Вам нужно определить, какой критерий следует применять в конкретной ситуации.

Непараметрические методы наиболее приемлемы, когда объем выборок мал. Если данных достаточно (например, $n > 100$), часто не имеет смысла использовать непараметрическую статистику. Если размер выборки очень мал (например, $n = 10$ или меньше), то уровни значимости для тех непараметрических критериев, которые используют нормальное приближение, можно рассматривать только как грубые оценки.

Для того чтобы оценить зависимость между двумя переменными, обычно вычисляют коэффициент корреляции Пирсона. Непараметрическими аналогами коэффициента корреляции Пирсона являются коэффициенты ранговой корреляции Спирмена R , статистика Кендалла и коэффициент Гамма.

Коэффициент ранговой корреляции оценивает величину зависимости между переменными, измеренными в порядковых шкалах, т. е. между порядковыми переменными.

Прозрачный способ построения парных коэффициентов корреляции из обобщенного коэффициента корреляции предложил Daniels. Обобщенный коэффициент корреляции определяется формулой:

$$\Gamma = \frac{\sum a_{ij} b_{ij}}{\sqrt{(\sum a_{ij}^2)(\sum b_{ij}^2)}}, \quad (1)$$

где $a_{ij} = a(X_i, X_j)$, $b_{ij} = b(Y_i, Y_j)$ — некоторые функции пар наблюдений X и Y соответственно, суммирование ведется по всем парам i, j .

Заметим, что при $a_{ij} = X_j - X_i$, $b_{ij} = Y_j - Y_i$ получаем обычный коэффициент корреляции Пирсона. Если переменные ранжированы, то мы работаем с рангами. Упорядочим значения X_i по возрастанию, то есть построим вариационный ряд этих величин. Номер величины X_i в этом ряде называется ее рангом и обозначается R_i .

Затем упорядочим значения Y_i в порядке возрастания. Номер величины Y_{ii} в этом ряде называется ее рангом и обозначается S_i . Коэффициент ранговой корреляции Спирмена вычисляется как обобщенный коэффициент парной корреляции с заменой наблюдений их рангами. Формально для обобщенного коэффициента корреляции нужно положить $a_{ij} = R_j - R_i$, $b_{ij} = S_j - S_i$.

Коэффициент Кендалла вычисляется, если в формуле для обобщенного коэффициента поло-

жить $a_{ij} = 1$ при $R_i < R_j$ и $a_{ij} = -1$ при $R_i > R_j$. Величины b_{ij} задаются аналогичными соотношениями с заменой рангов R_{ij} ранги S_i наблюдений Y . Итак, мы ясно видим, что идея всех корреляций возникает из одного и того же источника.

В том случае, когда имеется более двух переменных, то используют коэффициент конкордации Кендалла. Например, он применяется для оценки согласованности мнений независимых экспертов, например, баллов, выставленных одному и тому же явлению или фактору.

Если имеются две категориальные переменные, то для оценки степени зависимости используют стандартные статистики и соответствующие критерии для таблиц сопряженности: хи-квадрат, точный критерий Фишера и др.

Нелегко дать простой и однозначный совет, касающийся использования этих процедур. Каждая имеет свои достоинства и свои недостатки. Например, двухвыборочный критерий Колмогорова-Смирнова чувствителен не только к различию в положении двух распределений, но также и к форме распределения. Фактически он чувствителен к любому отклонению от гипотезы однородности, но не указывает, с каким именно отклонением мы имеем дело.

В общем, если результат исследования является важным и наблюдений немного, то всегда целесообразно испытать непараметрические тесты. Возможно, результаты тестирования (разными тестами) будут различны. В таком случае следует попытаться понять, почему разные тесты дали разные результаты. С другой стороны, непараметрические тесты имеют меньшую мощность, чем их параметрические конкуренты, и если важно обнаружить даже слабые эффекты, следует провести многократные испытания и особенно внимательно выбирать статистику критерия.

Аварии на угольных шахтах относятся, естественно, к случайным событиям. Поэтому для того, чтобы иметь возможность объективно проводить сравнение отдельных шахт, групп шахт (объединений, холдинговых компаний) или изучать отрасль в целом по аварийности, необходимо определить статистически подтвержденные показатели аварийности. Изучая соответствующие литературные источники, например [8], возможно сделать вывод о необходимости использования для анализа данных о горном производстве различных статистических распределений, кроме нормального, в частности логарифмически-нормального, экспоненциального, Вейбулла, гамма

Таблиця 3

Результаты определения закона распределения аварий на ряде крупных шахт Украины

№ п/п	Наименование закона распределения вероятности аварий	Параметры распределения		Число степеней свободы	Оценка достоверности			Уровень значимости	Оценка соответствия
		Наименование	Значение		Фактический	Критерий Пирсона	Теоретический при уровне значимости		
1.	Нормальный	Среднее	413,707	10	148,651	0,05	0,01	0	Не соотв.
		Стандартное отклонение				18,3	23,2		
2.	Логарифмически нормальный	Среднее	451,614	12	24,762	21,0	26,2	7,30 · 10 ⁻³	Не соотв.
		Стандартное отклонение	541,425						
3.	Экспоненциальный	Среднее	413,700	13	45,670	22,4	27,7	1,62 · 10 ⁻³	Не соотв.
		Форма (α)	1,267						
4.	Вейбулла	Масштаб (β)	446,861 · 10 ⁻³	10	16,782	18,3	23,2	0,073	Соотв.
		Форма (α)	1,415						
5.	Гамма-распределение	Масштаб (β)	8,420 · 10 ⁻³	10	15,62	18,3	23,2	0,111	Соотв.

распределения, а, возможно, и некоторых других, в том числе Стьюдента, Пирсона и др.

Для проверки высказанных выше положений данные по аварийности в горной промышленности по 60 шахтам обрабатывались посредством компьютерных технологий. Результаты расчетов проведены в табл. 3, из которой следует, что дальнейшая обработка данных и последующий анализ могут быть реализованы только методами непараметрической статистики.

Классическая статистика хи-квадрат Пирсона замечательна тем, что ее распределение приближается к распределению хи-квадрат, для которого имеются подробные таблицы. Процентные точки распределения хи-квадрат могут быть также эффективно вычислены в системе STATISTICA.

Свойство критерия хи-квадрат (точность аппроксимации распределения статистики распределением хи-квадрат) для таблиц 2×2 с малыми ожидаемыми частотами может быть улучшено за счет уменьшения абсолютного значения разностей между ожидаемыми и наблюдаемыми частотами на величину 0,5 перед возведением в квадрат. Эта непрерывность обычно применяется, когда ячейки содержат только малые частоты и некоторые ожидаемые частоты становятся меньше 5 (или даже меньше 10).

Если сумма частот небольшая, то лучше использовать точный критерий Фишера вместо критерия хи-квадрат. Известны рекомендации Кокрена для таблиц 2×2 : если сумма всех частот в таблице меньше 20, то следует использовать точный критерий Фишера. Если сумма частот больше 40, то можно применять критерий хи-квадрат с поправкой на непрерывность.

Так как в данных обычно имеются ячейки с малыми частотами (2 и 3), то для улучшения точности критерия хи-квадрат используем указанную поправку. Поскольку нас интересует односторонняя альтернатива, мы делим уровень $p = 0,0012$ пополам и получаем 0,0006.

Нелегко дать однозначное предложение, касающееся использования непараметрических процедур. Каждая непараметрическая процедура в модуле имеет свои достоинства и свои недостатки. Например, двухвыборочный критерий Колмогорова-Смирнова чувствителен не только к различию в положении двух распределений, например, к различиям средних, но также чувствителен и к форме распределения. Критерий Вилкоксона парных сравнений предполагает, что можно ранжировать различия между сравниваемыми наблюдениями. Если это не так, лучше

использовать критерий знаков. В общем, если результат исследования является важным, то всегда целесообразно применить различные непараметрические тесты. Возможно, результаты проверки (разными тестами) будут различны. В таком случае следует попытаться понять, почему разные тесты дали разные результаты. С другой стороны, непараметрические тесты имеют меньшую статистическую мощность (менее чувствительны), чем их параметрические конкуренты, и если важно обнаружить даже слабые отклонения (например, является ли данная пищевая добавка опасной для людей), следует особенно внимательно выбирать статистику критерия.

Выводы. В целом непараметрические методы наиболее приемлемы, когда объем выборок мал. Если данных много (например, $n > 100$), то не имеет смысла использовать непараметрические статистики. Главное здесь состоит в том,

что когда выборки становятся очень большими, то выборочные средние подчиняются нормальному закону, даже если исходная переменная не является нормальной или измерена с погрешностью. Таким образом, параметрические методы, являющиеся более чувствительными (имеют большую статистическую мощность), всегда подходят для больших выборок. Большинство критериев значимости многих непараметрических статистик, описанных далее, основываются на асимптотической теории (больших выборок) поэтому соответствующие тесты часто не выполняются, если размер выборки становится слишком малым.

Таким образом, из изложенного вытекает, что для обработки объемов выборок в сотни данных, характерных для угольной промышленности Украины, наиболее подходят методы непараметрической статистики.

Список литературы:

1. Минько В.М. Охрана труда в машиностроении: учеб. для студ. учреждений сред. проф. образования. М.: Академия, 2012. 256 с.
2. Типове положення про порядок проведення навчання і перевірки знань з питань охорони праці: НПАОП 0.00-4.12-05. Х.: Форт, 2005. 40 с.
3. Ступницька Н.В. Підвищення ефективності планування заходів запобігання виробничому травматизму на підприємствах машинобудування: автореф. дис. ... канд. техн. наук : 05.26.01 «Охорона праці» / ДУ «Львівська Політехніка». Л., 1999. 22 с.
4. Кружилко О.Є. Удосконалення комплексної оцінки стану охорони праці на підприємствах: автореф. дис. ... канд. техн. наук: 05.26.01 «Охорона праці». К., 2001. 20 с.
5. Гунченко О.М. Вдосконалення системи управління охороною праці на машинобудівних підприємствах: автореф. дис. ... канд. техн. наук: 05.26.01 «Охорона праці». Луганськ, 2007. 20 с.
6. Проблеми стану і необхідності вдосконалення системи управління охороною праці в галузі машинобудування / М.А. Касьянов, В.О. Медяник, О.М. Гунченко, Д.А. Вишневецький. Вісник Східноукраїнського національного університету ім. В. Даля. Луганськ: СНУ ім. В. Даля, 2008. № 6(124), ч. 2. С. 3–9.
7. Урбанчик Ю. Малая механизация в горнодобывающей промышленности в свете техники безопасности и производительности труда: Пер. с польск. Катовице: Механизация, 2001. С. 7.
8. Левкин Н.Е. Предотвращение аварий и травматизма в угольных шахтах Украины. Макеевка: МакНИИ, 2002. 392 с.

НЕПАРАМЕТРИЧНА СТАТИСТИКА ДЛЯ ОЦІНКИ ОСОБЛИВОСТЕЙ ГІРНИЧОГО ВИРОБНИЦТВА

На сучасному етапі технічного розвитку всі комплексні системи промислового виробництва (зокрема, вугільної промисловості) вимагають автоматизації всього процесу, включаючи роботу видобувного та транспортного устаткування, що не можна виконати без детального математичного опису всіх специфічних явищ, що характеризують конкретне виробництво.

Тому вперше для рішення задачі управління складними технологічними багатофакторними системами пропонується використання непараметричної статистики, які у вітчизняній спеціальній літературі не знайшли належного відображення. І тим не менше, саме такі математичні підходи дозволяють виконати коректний прогноз розглянутих явищ за межами наявних експериментальних даних ыз достатньою для вирішення більшості практичних завдань точністю, особливо в тих випадках, коли число самих даних відносно невелике.

Ключові слова: вугільна промисловість, оцінка, непараметрична статистика.

**NONPARAMETRIC STATISTICS FOR ASSESSMENT
OF MINERAL PRODUCTION FEATURES**

At the present stage of technical development all the complex system of industrial production and particularly the coal industry, requires automating the entire process including the work of mining and transport equipment, which can not be done without a detailed mathematical description of specific phenomena characterizing the specific production. Therefore, for the first time to solve the problem of managing complex technology, multifactorial systems, provided the use of non-parametric statistics, which are in the national literature is not adequately reflected. Nevertheless, it is such mathematical approaches allow to perform a correct forecast of the phenomena in question beyond the available experimental data sufficient to solve the most practical problems with accuracy, especially in those cases where the number of the data is relatively small.

Key words: coal industry, assessment, nonparametric statistics.