

УДК 57. 087

КОМПЛЕКСНЫЙ ПОДХОД К ОЦЕНКЕ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

^{1,2}Мадай Д.Ю., ²Вебер В.Р., ²Гривков А.С., ²Барт В.А., ²Мадай О.Д., ¹Гурин А.В.,
¹Никитина Е.А., ¹Безуглая Т.О.

¹Санкт-Петербургский государственный университет, Санкт-Петербург,
e-mail: dmitrij.madai@yandex.ru;

²Новгородский государственный университет им. Ярослава Мудрого, Великий Новгород

В настоящее время требуется комплексный подход оценки результатов исследования, который поможет экспериментаторам осуществлять объективнее и достоверные выводы по результатам собственных исследований. На базе статистической обработки тестового массива данных предполагается проиллюстрировать возможности следующих методов систематизации наблюдений: гистограммный анализ, корреляционные плеяды, главные компоненты, кривые дожития. Смысл таких исследований состоит в том, что при этом происходит повышение качественного понимания реального материала до уровня, достаточного для объективных практических выводов [5]. Следует отметить, что существенная сложность статистического анализа данных состоит в разнообразии аспектов, неоднородно влияющих на результат. Можно выделить четыре основных аспекта проблемы: научный, экспериментальный, организационный и методологический [4]. Научный аспект определяет уровень объективности (точности) исследований. Здесь самое важное, чтобы степень подготовленности и проведения эксперимента была согласована с уровнем систематизации его результатов.

Ключевые слова: научный, экспериментальный, организационный и методологический аспекты.

COMPLEX APPROACH TO ESTIMATE EXPERIMENTAL STUDY

^{1,2}Madai D.I., ²Veber V.R., ²Grivkov A.S., ²Bart V.A., ²Madai O.D., ¹Gurin A.V.,
¹Nikitina E.A., ¹Bezugaia T.O.

¹Saint Petersburg State University Saint Petersburg, e-mail: dmitrij.madai@yandex.ru;

²Novgorod State University after Yaroslav-the-Wise, Veliky Novgorod

Nowadays it is necessary to use complex approach to estimate experimental studies. This approach helps the researchers to make reliable and authoritative conclusions from their own investigations. Based on statistical analysis it is supposed to illustrate the following methods of test results classification: histogram analysis, correlation analysis, prime component, survivor curve. It should be mentioned that the statistical analytical complexity consists of a variety of considerations heterogeneously influenced on the results. Four main aspects of the problem are: scientific, experimental, executive and methodological. Scientific aspect determines the level of research reliability. It is the efficiency of the experiment that should be coordinated with the experimental results.

Keywords: scientific, experimental, executive and methodological aspects

Для улучшения лечебного эффекта препаратов необходимо повышать деятельность специфических и неспецифических звеньев защиты, усиливать гибкость ответных компенсаторно-приспособительных реакций в рамках универсальных постагрессивных адаптационных реакций (ПАР), например, с помощью особых физических способов, химических и биологических средств направленного действия на их этиологическое и патогенетическое звено [11].

С учетом изложенного, разработка показаний, теоретическое и патофизиологическое обоснование комплексного применения биологически активных препаратов при лечении ГВЗ приобретает особую актуальность, причем, решение этой важной проблемы имеет не только большое медицинское, но и социально-экономическое значение при решении практических задач охраны здоровья.

Отсутствие объективных критериев оценки показателей применения биологически активных препаратов сильно затруд-

няет адекватную коррегирующую терапию ПАР [5, 11].

Использование же современных методов статистической обработки позволит распределить больных по характеру и интенсивности проявления патологического процесса на группы с «благоприятным» и «неблагоприятным» вариантом течения заболевания. В свою очередь, такой «синдромальный» анализ проявления процесса дает возможность определить в структуре изучаемых реакций уровень долевого участия этих симптомов. Исследование корреляционных (кооперативных) связей симптомов позволит определить конкретное требуемое звено приложения для препаратов, применяемых для патогенетической медикаментозной коррекции патологического процесса.

1. Характеристика основного эксперимента и полученного тестового массива данных. Изучение характера течения постагрессивных адаптационных реакций у больных с гнойно-воспалительными

заболеваниями, гнойными осложнениями ран основывалось на клиническом анализе результатов лечения 142 больных. Из них 1-ю группу составил 21 больной с гнойными ранами, 2-ю группу – 18 пострадавших с сочетанными травмами, 3-ю группу – 28 больных с гнойными послеоперационными ранами, 4-ю группу – 41 больной с травматическим остеомиелитом и 5-ю – 34 больных с флегмонами. Все больные находились на лечении в клиниках военно-полевой хирургии ВмедА, НИИ скорой помощи им. И.И. Джанелидзе и 8 отделения НОКБ 2006-2012 гг.

В данном методическом руководстве тестовый массив состоит из 26 больных с флегмонами ЧЛО: из них 16 больных с неблагоприятной и 10 – с благоприятной формами течения болезни. Все больные наблюдались в 4-х временных точках: на 1-е, 7-е, 14-е и 21-е сутки.

Рассмотрим клиническую значимость ряда основных показателей, достаточно полно характеризующих, по нашему мнению, уровень резистентности организма, а также их динамику у больных с флегмонами челюстно-лицевой области, гнойными послеоперационными ранами.

Как мы полагаем, эти показатели определяют важные морфологические и биохимические механизмы постагрессивных реакций, причем их функции генетически детерминированы. Следовательно, предлагаемое изучение динамики изменения этих показателей в процессе лечения, а также исследование изменения биологических реакций организма на контакт с медикаментами, влияние этих реакций на течение воспаления или раневой болезни даст объективную информацию о динамике патологического процесса и эффективность его коррекции.

В приведенном ниже списке показателей (признаков) заболевания указан код, номер признака и его физиологические уровни значимости¹

Ig A / (1) Иммуноглобулины А ($1.42-1.98 \times 0.1^6$ г/л).

РТМЛ / (2) Реакция торможения миграции лейкоцитов (40-75%). настоящей работе мы изучали РТМЛ с конкавалином А – специфическим митогеном в отношении Т-лимфоцитов.

Л / (3) Лейкоциты ($4-8,8 \times 10^9$ /л).

М / (4) Моноциты ($0.290-0.6 \times 10^9$ /л).

Лф / (5) Лимфоциты ($1.2-3.0 \times 10^9$ /л) – основа как клеточного, так и гуморального иммунитета – предшественник антителоо-

бразующей клетки и носитель иммунологической памяти (Петров Р.В. с соавт., 1982).

ЛИИ / (6) ЛИИ – лейкоцитарный индекс интоксикации по Я.Я. Кальф-Калифу (0.5-1.0) – показатель тяжести патологического процесса и прогноз заболевания

$$\text{ЛИИ} = \frac{(C+2П+3Ю+4Ми)(ПлК+1)}{(Мон+Лимф)(Эоз+1)},$$

где С – сегментоядерные нейтрофилы; П – палочкоядерные; Ю – юные; Ми – миелоциты; Плк – плазматические клетки; Мон – моноциты; Лимф – лимфоциты; Эоз – эозинофилы.

Повышение ЛИИ до 3.0 единиц указывает преимущественно на ограниченный характер процесса, а его повышение до 4.0 и более – на значительную эндогенную интоксикацию.

РЭ / (7) Активность раневых энзимов (более 111.5 МЕ/л в 1 фазе раневого процесса, менее 80 МЕ/л – во 2 фазе) – косвенный признак регенерационной активности ран [11].

КФА / (8) Коэффициент функциональной активности моноцитов ($1,7 \pm 0.09$) – отношение активных и неактивных форм моноцитов.

ЩФ / (9) Активность щелочной фосфатазы в сыворотке крови (278830 нмоль/(с×л)). ЩФ (К.Ф.3.1.3.1.).

АМ / (10) Амплитуда моды R-R интервала на кардиоинтервалометрии (сек) – показатель преобладания воздействия симпатической и парасимпатической иннервации на работу сердца у больных и пострадавших с постагрессивными адаптационными реакциями. Данный показатель мы не учитывали, так как он был статистически незначим.

SH/SS / (11) Коэффициент отношения общего содержания окисленных и восстановленных групп цельной крови.

Цель эксперимента с массивом диктует необходимость знакомства с элементарной техникой методологий последовательного анализа выборок и индивидуумов.

Первичная обработка данных. Статистический вывод – это тонкая грань между нашими предположениями о структуре объекта исследования и реальными наблюдениями. Цель статистического анализа данных – повышение качественного уровня понимания предмета исследования. Этот процесс подчиняется определенным законам. Один из них (принцип иерархичности уровней точности) гласит: «Менее точные уровни исследования оказывают большее влияние на результат» [5]. Поэтому первичная обработка данных является самой

¹Гистограммы выгодно отличаются от диаграммного (графического) представления вариационного ряда тем, что они имеют управляющий параметр – длину интервала разбиения.

важной и сложной частью статистического исследования. К ней приходится возвращаться на всех последующих этапах обработки данных. Какие проблемы стоят на этом этапе?

Типы наблюдений. Природа признаков. Прежде всего, следует разобраться, с каким типом наблюдений мы имеем дело. Полезна следующая четверка исходных понятий статистики – индивиды, признаки, значения признаков и параметр. Индивиды – те объекты, с которых в процессе наблюдений снимаются некие их характеристики (показатели), называемые признаками. Значениями признаков являются: числа, для количественных показателей (например, 175 см. – значение признака «рост»), и градации для качественных признаков. Например, для такого качественного показателя, как тип патологического процесса имеются следующие фонологические аспекты (градации): «флегмона», «сочетанные травмы» и т.д.

Классические методы статистики развивались в основном для количественных признаков. Обработка данных качественных признаков представляет собой наиболее сложную задачу статистики. Здесь можно выделить три главных направления:

1. Шкалирование;
2. Использование информационных статистик;
3. Дисперсионный анализ.

Первое и третье направления связаны с идеей подбора для градаций каких-то числовых значений. Применение различных шкал требует исследования клинической значимости признака. В дисперсионном анализе значения качественного признака (фактора) кодируются средними показателями по градациям из значений одновременно наблюдаемого количественного признака (отклика). Эти методы мы рассматривать не будем.

Информационными статистиками называются функции от выборочных наблюдений, зависящие только от частот.

Важным частным случаем качественных признаков является ситуация, когда их градации можно упорядочить. В этом случае значениям можно приписать числа (ранги), имеющие смысл порядкового номера градации с учетом повторностей ее появления в выборке. В настоящее время ранговая статистика бурно развивается как часть непараметрических методов. Поясним процедуру ранжирования (т.е. приписывания рангов) на примере.

Для признака «состояние больного» упорядоченными значениями могут быть: Т – «тяжелое», С – «средней степени тяже-

сти», У – «удовлетворительное». Если, например, наблюдения 10-ти больных показали следующие их состояния:

(С, Т, С, У, С, У, С, У, С, С),

то, упорядочивая их, имеем:

Т, С, С, С, С, С, С, У, У, У,

1 2 3 4 5 6 7 8 9 10

и сопоставляя каждому значению, его номер по порядку, усредняя номера для одинаковых значений, получаем в качестве значений признака их ранги (значению Т соответствует в данной выборке ранг 1, С – ранг 4,5, У – ранг 9). Таким образом, наша выборка после ранжирования примет вид

4,5 ; 1 ; 4,5 ; 9 ; 4,5 ; 9 ; 4,5 ; 9 ; 4,5 ; 4,5).

С ней можно обращаться как с выборкой из количественного признака: искать числовые характеристики для рангов, строить ранговые критерии и т.д.

Параметр – показатель неопределенности статистического эксперимента в теоретических конструкциях. Параметрические модели в статистике содержательно более богаты, чем непараметрические, ибо в них присутствует предыдущий опыт исследования проблемы.

По своей структуре можно выделить следующие типы наблюдений: матрицы, процессы, более сложные конфигурации (например, латинские квадраты или блок-схемы в дисперсионном анализе и т.д.). В отдельный тип (назовем его прочерки или артефакты) следует отнести наблюдения, в которых имеются пропуски значений признаков по техническим или каким-либо другим причинам.

Матрицы наблюдений – это прямоугольные таблицы значений x_{ij} признаков (номера признаков i , изменяют от 1 до k) на индивидах (с номерами j , меняющимися от 1 до n). В электронной таблице исходных данных компьютерной системы STATISTICA признакам соответствует название «Variables – Переменные», а индивидам – «Cases – Случай». Предполагается, что индивиды независимы и образуют повторную выборку объема n . В случае одномерного количественного признака ($k=1$) повторная выборка представляет собой набор из n чисел: (x_1, x_2, \dots, x_n) – его значений, полученных в результате n независимых повторений одного и того же эксперимента. В реальности получение повторной выборки часто является непростой проблемой.

Признаки, наоборот, зависимы между собой. Более того, зависимость между при-

знаками и является чаще всего предметом исследования. Исследование взаимосвязей между указанными признаками лежит в основе статистического решения определенных выше задач. Например, в приводимой в табл. 1 матрице наблюдений указаны значения 9 признаков в четырех временных точках (столбцы) для 26 индивидов (строки). Эта таблица является матрицей наблюдений для тестового массива, в ней $n=26$ и $k=9$. Приведенные значения получены из историй болезней 26 больных.

Наблюдения типа процессов представляют собой, как правило, динамику изме-

нения во времени рассматриваемого признака. Значения наблюдаемого признака в разные моменты времени зависимы между собой и, следовательно, не могут рассматриваться как повторности. Но если для каждого (повторного) больного из данной группы (выборки) имеется свое изображение динамики изменения признака, то мы говорим о повторных реализациях процесса (или реализации процесса с повторностями). Для наглядной визуализации данного процесса на рис. 1. представлена динамика уровней лимфоцитов на 1, 7, 14 и 21-е сутки.

Тестовый массив данных. Клинико-лабораторные показатели

1-е сутки наблюдения										7-е сутки									
№	Ig1	PTM1	A1	M1	ЛФ1	РЭ1	КФА1	ЩФ1	SS/SH1	№	Ig2	PTM2	A2	M2	ЛФ2	РЭ2	КФА2	ЩФ2	SS/SH2
1	1,2	52,3	24	0,2	1,4	278	1,3	1003	0,36	1	1,3	52,3	17	0,3	1,5	145	1,5	1150	1,31
2	1,2	52,2	24	0,1	1,5	278	1,3	1002	0,35	2	1,2	52,4	17	0,3	1,4	144	1,5	1151	1,32
3	1,2	52,3	24	0,1	1,4	279	1,4	1003	0,35	3	1,3	52,4	17	0,3	1,5	145	1,4	1150	1,33
4	1,2	52,2	24	0,2	1,4	278	1,6	1002	0,36	4	1,3	52,3	17	0,4	1,5	145	1,5	1151	1,33
5	1,2	52,2	26	0,3	1,6	269	1,5	1006	0,35	5	1,3	53,1	17	0,4	1,4	145	1,6	1150	1,33
6	1,2	53,1	26	0,3	1,6	270	1,5	1006	0,35	6	1,3	53,2	18	0,4	1,6	145	1,6	1152	1,35
7	1,2	53,1	27	0,3	1,6	266	1,5	1005	0,36	7	1,3	53,4	18	0,4	1,5	145	1,5	1152	1,35
8	1,2	53,2	27	0,3	1,6	266	1,4	1006	0,35	8	1,3	53,4	18	0,3	1,5	146	1,5	1152	1,35
9	1,2	52,1	24	0,2	1,5	278	1,5	1004	0,35	9	1,2	52,2	17	0,3	1,4	144	1,6	1151	1,33
10	1,2	52,2	26	0,3	1,6	270	1,5	1005	0,36	10	1,2	53,1	17	0,4	1,4	145	1,6	1151	1,33
11	1,2	53,2	27	0,3	1,6	266	1,4	1006	0,35	11	1,3	53,4	18	0,3	1,5	146	1,5	1152	1,35
12	1,2	53,2	27	0,3	1,6	266	1,4	1005	0,34	12	1,3	53,4	18	0,4	1,5	145	1,5	1151	1,34
13	1,2	53,4	26	0,4	1,7	267	1,5	1005	0,35	13	1,3	53,5	18	0,7	1,5	146	1,5	1151	1,35
14	1,2	53,4	25	0,4	1,7	266	1,5	1005	0,35	14	1,3	53,5	18	0,3	1,5	146	1,5	1151	1,33
15	1,2	54,4	25	0,4	1,7	266	1,5	1005	0,34	15	1,2	53,4	18	0,3	1,5	146	1,5	1151	1,33
16	1,2	55,4	25	0,4	1,7	266	1,4	1005	0,35	16	1,3	53,4	18	0,3	1,6	146	1,5	1151	1,34
17	1,2	54,8	25	0,4	1,7	266	1,4	1005	0,35	17	1,3	53,5	18	0,3	1,6	146	1,5	1151	1,35
18	1,2	52,4	23	0,1	1,4	278	1,3	1001	0,38	18	1,3	51,7	18	0,4	1,6	143	1,5	1155	1,33
19	1,2	52,1	24	0,1	1,4	279	1,4	1002	0,36	19	1,3	52,1	17	0,4	1,5	145	1,4	1149	1,3
20	1,2	52,4	24	0,1	1,5	278	1,3	1002	0,35	20	1,2	52,4	17	0,4	1,5	144	1,4	1151	1,3
21	1,2	52,3	24	0,1	1,4	279	1,4	1003	0,35	21	1,2	52,4	17	0,3	1,5	145	1,4	1150	1,33
22	1,2	52,1	24	0,2	1,5	278	1,5	1004	0,35	22	1,3	52,2	17	0,3	1,4	144	1,6	1151	1,33
23	1,2	52,2	26	0,3	1,6	270	1,5	1005	0,36	23	1,3	53,1	17	0,4	1,4	145	1,6	1151	1,33
24	1,2	52,3	24	0,1	1,4	279	1,4	1003	0,35	24	1,3	52,4	17	0,3	1,5	145	1,4	1150	1,33
25	1,2	52,2	24	0,2	1,4	278	1,6	1002	0,36	25	1,3	52,3	17	0,4	1,5	145	1,5	1151	1,33
26	1,2	52,2	26	0,3	1,6	270	1,5	1005	0,36	26	1,3	53,1	17	0,4	1,4	145	1,6	1151	1,33

14-е сутки									21-е сутки										
№	Ig3	PTM3	A3	M3	ЛФ3	РЭ3	КФА3	ЩФ3	SS/SH3	№	Ig4	PTM4	A4	M4	ЛФ4	РЭ4	КФА4	ЩФ4	SS/SH4
1	3	54,4	13	0,5	1,5	128	1,5	1022	1,44	1	3,2	66,7	11	0,7	2,6	52,1	1,5	538	2,77
2	3,1	54,5	13	0,5	1,6	128	1,4	1021	1,45	2	3,2	67,1	10	0,8	2,6	52,2	1,5	537	2,76
3	3	54,5	13	0,6	1,8	128	1,5	1022	1,45	3	3,2	66,8	10	0,8	2,7	52,2	1,5	538	2,77
4	3	54,4	13	0,5	1,7	127	1,4	1023	1,44	4	3,2	66,7	10	0,7	2,6	52,1	1,6	537	2,76
5	3,1	56,5	14	0,7	1,5	130	1,5	1025	1,45	5	3,2	65,6	10	0,7	2,5	52,5	1,6	534	2,75
6	3,1	56,4	15	0,7	1,5	129	1,5	1025	1,46	6	3,3	65,4	11	0,8	2,5	55,3	1,6	536	2,75
7	3,1	55,5	15	0,9	1,5	129	1,4	1026	1,45	7	3,3	65,3	12	0,7	2,4	52,5	1,5	535	2,76
8	3	55,4	15	0,6	1,5	129	1,5	1025	1,46	8	3,3	65,7	12	0,7	2,4	52,5	1,5	535	2,75
9	3	54,4	13	0,6	1,6	129	1,5	1025	1,46	9	3,2	66,6	10	0,7	2,5	52,5	1,6	537	2,75
10	3,1	56,4	14	0,7	1,5	129	1,5	1025	1,46	10	3,2	65,6	10	0,7	2,5	52,5	1,6	537	2,76
11	3	55,4	15	0,6	1,5	129	1,5	1025	1,46	11	3,3	65,7	12	0,7	2,4	52,5	1,5	535	2,75
12	3	55,5	15	0,6	1,5	130	1,5	1025	1,45	12	3,3	65,5	12	0,7	2,4	52,5	1,5	535	2,75
13	3	56,6	15	0,7	1,6	128	1,5	1025	1,46	13	3,2	65,6	12	0,7	2,5	52,4	1,5	535	2,75
14	3,1	56,5	15	0,7	1,6	127	1,5	1025	1,45	14	3,2	65,6	12	0,7	2,5	52,4	1,5	535	2,69
15	3,1	55,5	15	0,7	1,6	127	1,5	1025	1,45	15	3,2	65,5	12	0,6	2,5	52,7	1,5	535	2,65
16	3,1	54,6	14	0,7	1,6	127	1,5	1025	1,45	16	3,2	66,5	12	0,6	2,5	52,7	1,5	535	2,65
17	3,1	54,7	14	0,6	1,6	127	1,5	1025	1,45	17	3,3	65,7	13	0,6	2,5	52,7	1,5	535	2,65
18	3,1	54,4	13	0,5	1,8	128	1,6	1020	1,44	18	3,2	67,3	10	0,9	2,5	51,2	1,5	541	2,79
19	3	54,7	13	0,5	1,9	129	1,5	1020	1,43	19	3,4	66,9	11	0,8	2,7	52	1,6	538	2,77
20	3	54,5	13	0,6	1,6	129	1,5	1023	1,44	20	3,2	66,8	11	0,8	2,7	52	1,6	538	2,78
21	3	54,5	13	0,6	1,8	128	1,5	1022	1,45	21	3,2	66,8	10	0,8	2,7	52,2	1,5	538	2,77
22	3	54,4	13	0,6	1,6	129	1,5	1025	1,46	22	3,2	66,6	10	0,7	2,5	52,5	1,6	537	2,75
23	3,1	56,4	14	0,7	1,5	129	1,5	1025	1,46	23	3,2	65,6	10	0,7	2,5	52,5	1,6	537	2,76
24	3	54,5	13	0,6	1,8	128	1,5	1022	1,45	24	3,2	66,8	10	0,8	2,7	52,2	1,5	538	2,77
25	3	54,4	13	0,5	1,7	127	1,4	1023	1,44	25	3,2	66,7	10	0,7	2,6	52,1	1,6	537	2,76
26	3,1	56,4	14	0,7	1,5	129	1,5	1025	1,46	26	3,2	65,6	10	0,7	2,5	52,5	1,6	537	2,76

--- Неблагоприятное течение (I гр.); -- Благоприятное течение (II гр.);

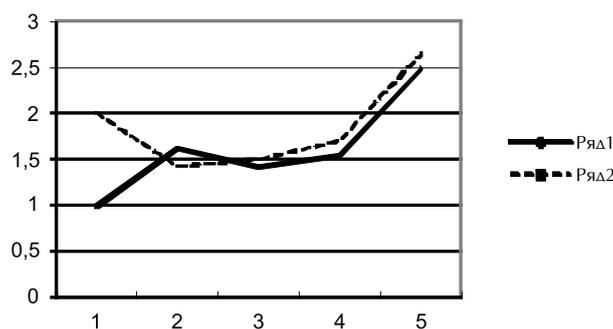


Рис. 1. Динамика уровней лимфоцитов в 1-е (1), 7-е (2), 14-е (3) и 21-е (4) сутки у больных 1-й группы с неблагоприятным течением ПАР (1-й ряд) и 2-й группы с благоприятным течением ПАР (2-й ряд)

Если в матрице наблюдений имеются прочерки технического плана, то есть не связанные с существом дела (потеряна часть данных или они не наблюдались по техническим причинам), то не следует вставлять вместо пропусков искусственные значения признаков. Следует либо удалить соответствующие индивиды или признаки, либо, если это невозможно (например, мал объем оставшейся выборки) предсказывать пропущенные значения признака, по остальным признакам с учетом такой замены в выводах. Борьба с артефактами является непростым вопросом [4]. В системе STATISTICA есть специальные команды для различных стратегий решения этой проблемы. Например, замена прочерков на средние значения.

Гистограммный анализ. Наглядное представление об исходной выборке обычно задается в виде различных диаграмм распределений частот значений признака в выборке. Отметим важную терминологию. Неодинаковые значения признака, упорядоченные по возрастанию, называются вариантами. В выборке вариантам соответствуют частоты (повторности вариант). Представление выборки в виде набора вариант с набором их частот называется вариационным рядом. Набор выборочных частот для всех значений признака называется выборочным распределением признака. Полный набор вероятностей, с которыми признак может принимать свои значения, называется генеральным распределением признака. Все характеристики распределений могут рассматриваться на выборочном и генеральном уровнях.

Генеральный уровень соответствует более объективному анализу данных. Поэтому часто целью статистического анализа является исследование генеральных свойств признака на основе выборочных наблюдений. Функции от выборочных наблюдений,

рассматриваемые на генеральном уровне называются статистиками¹.

Среди диаграмм для выборочных распределений наиболее часто используются гистограммы, которые строятся следующим образом. Область значений признака разбивается на интервалы одинаковой длины и над каждым интервалом строится прямоугольник, площадь которого равна числу выборочных наблюдений, попавших в данный интервал. Наиболее часто встречающиеся варианты называются модами. На гистограмме модам соответствуют прямоугольники с максимальными высотами. Гистограммный анализ позволяет ответить на очень важные вопросы об однородности и точности выборочных наблюдений.

На результаты наблюдений существенное влияние может оказывать прибор, используемый в эксперименте (роль прибора может играть и человек). Обычно влияние прибора характеризуют систематической и случайной ошибкой. Но можно уменьшить влияние таких шумов (навести контрастность) и на уровне всего распределения в гистограммном анализе при достаточном числе наблюдений. Для этого следует сравнивать гистограммы при различных значениях длины интервала разбиения (процедура «стягивания окна»). Моды, которые сохраняются при такой процедуре, называются устойчивыми. Длина интервала разбиения, при которой все устойчивые моды сохранены, а случайные моды исчезли, определяет истинную (внутреннюю) точность данных наблюдений².

¹Это принятый рабочий термин; не следует его путать с названием статистики как науки.

²Гистограммы выгодно отличаются от диаграммного (графического) представления вариационного ряда тем, что они имеют управляющий параметр – длину интервала разбиения.

На рис. 2 видно, что разбиение на 26 интервалов (случай А) слишком мелкое – много случайных мод. Случай В слишком груб. Точность данных лучше всего соответствует случаю Б, где наблюдаются две устойчивые моды. Материал неоднороден по этому признаку. В нем смешаны две различные совокупности исходных данных, отражающих характер течения патологического процесса. Однородная выборка характеризуется симметричной гистограммой с одной

устойчивой модой и может интерпретироваться как наблюдение постоянного значения (константы), зашумленного случайными ошибками прибора. В этой ситуации, в случае непрерывного признака, чаще всего ей соответствует генеральное распределение, называемое нормальным. Оно полностью определяется двумя параметрами: генеральным средним (оно называется математическим ожиданием) и (генеральной) дисперсией.

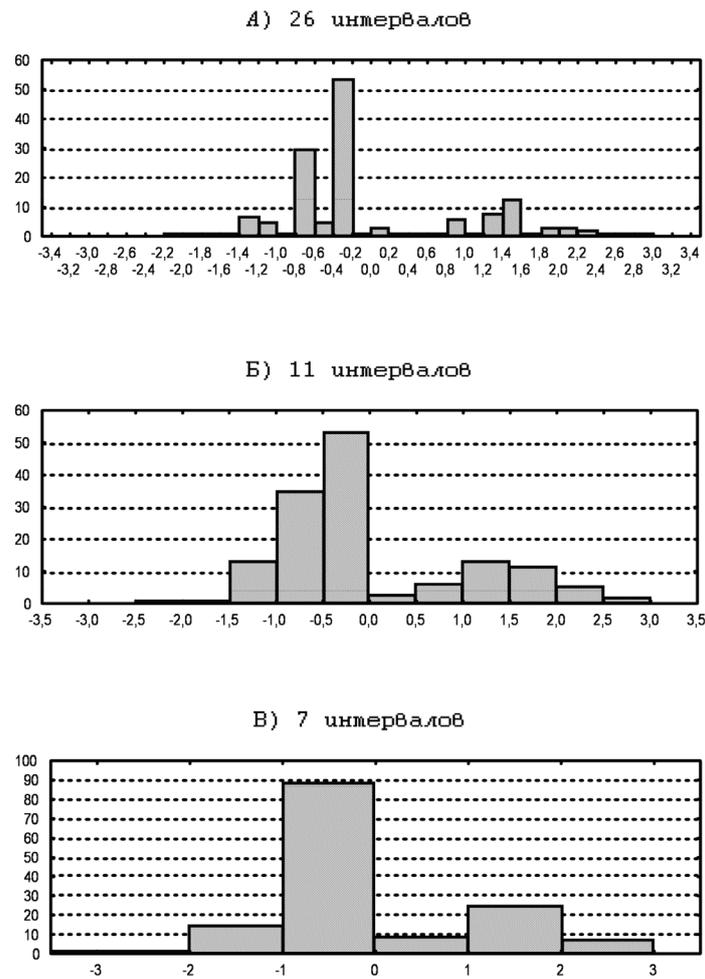


Рис. 2. Процедура «стягивания окна» на примере общего массива больных. В качестве признака выбран 2-й фактор на 1-е сутки наблюдения

Поли-modalность гистограммы – показатель существенной неоднородности выборки. Для правильности дальнейших выводов важно понять естественную природу такой неоднородности. В реальных меди-

цинских данных гистограмма, не противоречащая предположению о нормальном распределении, часто оказывается несимметричной, но, как правило, унимодальной (рис. 3).

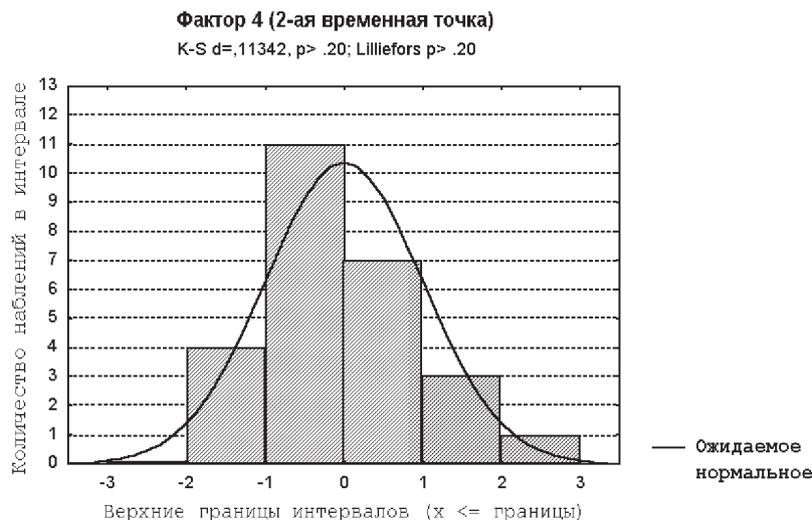


Рис. 3. Пример гистограммы близкой к показателю однородной выборки. В качестве признака выбран 4-й фактор во второй временной точке тестового массива

Выводы

Организационные и экспериментальные аспекты развиваются по своим законам, которые постоянно приходится учитывать. Наибольшие трудности представляет методология, то есть согласование традиционного языка описания эксперимента с принятой терминологией в методах обработки его результатов. Лучшее всего иметь постоянный контакт с профессионалом – математиком, но многое можно сделать и самостоятельно, учитывая современный прогресс в компьютерной технике вычислений и сервиса представления данных. Помощь в этом и является конкретной целью данных методических указаний. Основным техническим инструментом в систематизации наблюдений является система статистических программ для персональных компьютеров STATISTICA.

Список литературы

1. Барт А.Г. Анализ медико-биологических систем: метод частично обратных функций. – СПб.: изд-во. С.-Петерб. ун-та.– 2003. – 280 с.
2. Бондаренко Б.Б., Барт В.А., Демченко Е.А., Барт А.Г., Заславский М.Л., Актуальные аспекты методологии клинико-статистического анализа // Клиническая и экспериментальная кардиология: Сб. ст. / под ред. чл.-корр. РАМН Е.В. Шляхто, 2005. С. 136-148.
3. Вебер В.Р., Мадай Д.Ю., Барт А.Г. Биометрический анализ: Руководство для студентов ВУЗов и слушателей 1-го и 6-го факультетов. – СПб., 2003. – 104 с., ил.
4. Ван-дер Варден Б., Математическая статистика. – М.: ИЛ, 1960.
5. Мадай Д.Ю. Комплексное использование полисахаридов для коррекции постагрессивных адаптационных реакций (ПАР) у больных с гнойно-воспалительными заболеваниями и травмами челюстно-лицевой области: Автореф. дис. ... д-ра мед. наук. – СПб.: Воен.-мед. акад., 1998. – 44 с.
6. Тьюки Дж., Анализ результатов наблюдений. – М.: Мир, 1981. – 693 с.
7. Тюрин Ю.Н., Макаров А.А., Статистический анализ данных на компьютере. – М.: «ИНФРА-М», 1998. – 528 с.