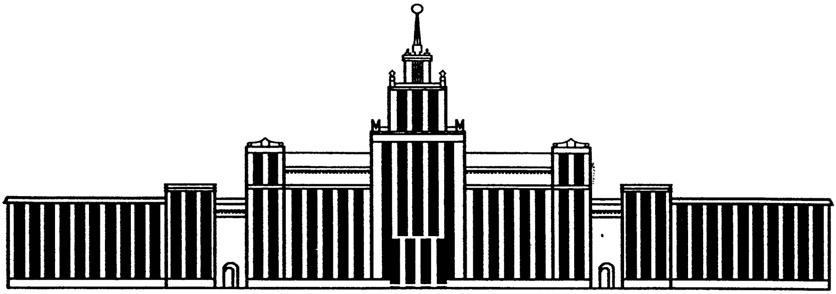
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ



ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

519.2(07)
П844

ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА

Методические указания

Челябинск
2017

Министерство образования и науки Российской Федерации
Южно-Уральский государственный университет
Кафедра «Технология и организация общественного питания»

519.2(07)
П844

ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА

Методические указания

Челябинск
Издательский центр ЮУрГУ
2017

УДК 519.242(075.8) + 664(075.8)
П844

Одобрено
учебно-методической комиссией
Института спорта, туризма и сервиса

Рецензент
Форенталь В.И. – ООО НПП «УЧТЕХ-ПРОФИ», зам. директора, к.т.н.

Планирование эксперимента: методические указания / сост.
П844 Л.С. Прохасько. – Челябинск: Издательский центр ЮУрГУ, 2017. – 33 с.

В методических указаниях содержатся задания для выполнения практических работ, приведено краткое теоретическое введение, рекомендации и порядок выполнения работ.

Практикум предназначен для студентов дневной и заочной форм обучения направления 19.04.04 «Технология продукции и организации общественного питания» (магистратура).

УДК 519.242(075.8) + 664(075.8)

Практическая работа 1

СПОСОБЫ ОБНАРУЖЕНИЯ И УСТРАНЕНИЯ ПОГРЕШНОСТЕЙ

Цель работы

Углубление теоретических знаний о грубых и систематических погрешностях, приобретение практических навыков исключения погрешностей из результатов измерений.

Общие сведения

Предметом измерения физической величины является определение ее количественных характеристик – веса, длины, площади, объема, скорости изменения и пр. Для определения абсолютного значения физической величины ее сравнивают с эталоном, который считается единицей величины. В процессе эксперимента экспериментатор сравнивает измеряемую величину не с основным эталоном, а с показаниями некоторого прибора – вторичным эталоном. Различают *прямое* и *косвенное* измерения.

Прямое измерение – искомое значение величины находят непосредственно с помощью измерительного прибора (например, измерение напряжения вольтметром, температуры – термометром и т.п.). Если прямые измерения невозможны, используют *косвенные* измерения – в них искомое значение величины находят на основании известной зависимости этой величины от других, допускающих прямое измерение.

Измерения выполняют однократно – *однократные измерения* и многократно – *многократные измерения*. Однократное измерение дает единственный результат, который принимают за окончательный результат измерения. Многократное измерение осуществляют повторением однократных измерений одной и той же физической величины, что приводит к получению набора данных. Окончательный результат многократного измерения получают из набора данных в виде среднего арифметического результатов всех отдельных измерений. Однако, как бы тщательно ни проводились измерения, любой результат этих измерений неизбежно содержит погрешность. Даже при многократных измерениях постоянной величины могут быть различные результаты измерений. Для определения понятия «погрешность» необходимо пояснить различие между такими понятиями, как истинное и действительное значение физической величины.

Истинное значение физической величины – это значение, идеальным образом отражающее свойство данного объекта как в количественном, так и в качественном отношении. На практике это понятие заменяют понятием «действительное значение». *Действительное значение физический величины* – значение, найденное экспериментально и настолько приближающееся к истинному, что для данной цели оно может быть использовано вместо него. Результат измерения всегда отличается от истинного значения измеряемой величины и представляет ее приближенное значение.

Погрешность результата измерения (сокращенно – погрешность измерения) – это количественная характеристика неоднозначности результата измерения, отклонение результата измерения от истинного значения измеряемой величины. Выявление погрешности измерения необходимо для определения окончательного результата измерения, которое нельзя рассматривать как истинное значение измеряемой физической из-за наличия погрешности.

Погрешность может быть выражена в единицах измеряемой величины x, в этом случае она обозначается Δx и носит название *абсолютной погрешности.* Однако абсолютная погрешность не отражает качества измерений, поэтому критерием качества измерения является отношение абсолютной погрешности к окончательному результату измерения:

$$\frac{\Delta x}{x} = \delta x. \tag{1}$$

Это отношение безразмерно. Величину δx называют *относительной погрешностью* и используют как в абсолютном, так и в процентном выражении.

Количество факторов, влияющих на точность измерения, достаточно велико, чем и объясняется большое количество видов погрешностей. По характеру изменения результатов при повторных измерениях, погрешности разделяются на систематические, случайные и грубые погрешности (промахи).

Систематическая погрешность измерения – составляющая погрешности измерения, остающаяся постоянной или закономерно изменяющаяся при повторных измерениях одной и той же величины.

Случайная погрешность измерения – составляющая погрешности измерения, изменяющаяся случайным образом при повторных измерениях одной и той же величины.

Грубая погрешность (промах) измерений – погрешность измерений, существенно превышающая ожидаемую при данных условиях или это погрешность результата отдельного измерения, входящего в ряд измерений, которая для данных условий резко отличается от остальных результатов этого ряда.

Грубые погрешности возникают вследствие неисправности измерительных приборов или ошибок в эксперименте, допущенных оператором по невнимательности. К ним можно отнести: – неверный отсчет по шкале измерительного прибора из-за неверного учета цены малых делений шкалы;

– неправильная запись результата наблюдений, значений отдельных мер использованного набора – ошибки оператора;

– хаотические изменения параметров напряжения, питающего средство измерения, например, его амплитуды или частоты.

Наиболее часто они допускаются неквалифицированным персоналом при неправильном обращении со средством измерения, неверным отсчетом показаний, ошибками при записи или вследствие внезапно возникшей посторонней причины.

Они сразу видны среди полученных результатов, так как полученные значения отличаются от остальных значений совокупности измерений. Если в процессе измерений удается найти причины, вызывающие существенные отличия, и после устранения этих причин повторные измерения не подтверждают подобных отличий, то такие измерения могут быть исключены из рассмотрения. Но необдуманное отбрасывание резко отличающихся от других результатов измерений может привести к существенному искажению характеристик измерений.

При однократных измерениях обнаружить промах не представляется возможным. Для уменьшения вероятности появления промахов измерения проводят 2–3 раза и за результат принимают среднее арифметическое полученных отсчетов. При многократных измерениях для обнаружения промахов используют статистические критерии.

Случайные величины и их характеристики

Основным типом погрешностей являются случайные погрешности. И так как они поддаются строгому математическому описанию, можно говорить о качестве измерений, в которых они присутствуют.

Случайная величина x полностью задается *плотностью вероятности* $\rho(x)$.

Среднее значение $\bar{x}$ измеряемой величины x указывает центр распределения, около которого группируются результаты отдельных измерений:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad (2)$$

где n – число измерений; x_i – результат i-го измерения.

Дисперсию вводят как средний квадрат отклонения отдельных результатов от среднего значения случайной величины:

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2. \qquad (3)$$

Коэффициент (n–1) в знаменателе формулы (3) появляется с целью получения несмещенной оценки для дисперсии, так как в связи с конечным количеством экспериментов вычисленное среднее значение $\bar{x}$ отличается от предельного, получаемого при $n \to \infty$.

Среднее квадратичное отклонение определяют как квадратный корень из дисперсии:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}. \qquad (4)$$

Эта величина характеризует разброс результатов отдельных измерений вокруг среднего значения, получаемого после обработки всех данных многократного измерения, причем точные значения σ и $\bar{x}$ являются предельными величинами, так как могут быть получены в случае достаточно большого количества проведенных измерений, в идеале – $n \to \infty$. При конечных n корректнее использовать термин *экспериментальная оценка*, который можно отнести и к среднему значению, и к дисперсии.

При обработке данных измерений в науке и технике обычно предполагают нормальный закон распределения случайных погрешностей измерений. Это предположение верно, так как показывает многочисленный опыт обработки данных, нормальный закон распределения случайных погрешностей всегда проявляется в случае, если суммарная погрешность есть результат неучтенного совместного воздействия множества причин, каждая из которых дает малый вклад в погрешность. При этом неважно, по какому закону распределен каждый из вкладов в отдельности.

Функциональное выражение для нормального распределения определяется формулой Гаусса:

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x_i - \bar{x})^2}{2\sigma^2}\right]. \qquad (5)$$

Вероятность того, что результат измерения попадет в интервал $[x_1; x_2]$, равна

$$P(x_1 \le x \le x_2) = \int_{x_1}^{x_2}\rho(x)dx. \qquad (6)$$

При увеличении границ промежутка в обе стороны до бесконечности интеграл от функции распределения $\int_{-\infty}^{+\infty} \rho(x)dx = 1$, то есть попадание результата измерения в диапазон $x \in (-\infty;\ +\infty)$ является достоверным событием.

Введем ε – величину отношения полуширины интервала Δx к среднему квадратичному отклонению σ:

$$\varepsilon = \frac{\Delta x}{\sigma}, \tag{7}$$

где Δx – произвольное отклонение x_i от средней величины $\bar{x}$.

В табл. 1 указана вероятность α в зависимости от ε:

$$\alpha = P(\bar{x} - \varepsilon\sigma \le x \le \bar{x} + \varepsilon\sigma). \tag{8}$$

Таблица 1

Доверительные интервалы ($\bar{x}$-Δx; $\bar{x}$+Δx) для доверительной вероятности α

α	0,680	0,900	0,950	0,990	0,997	0,999
ε	1,00	1,65	2,00	2,60	3,00	3,30

Анализ табл. 1 показывает, что результат измерения с вероятностью около 68 % попадает в интервал $(\bar{x} - \sigma;\ \bar{x} + \sigma)$, то есть каждое третье измерение оказывается вне этого интервала. При $(\bar{x} - 2\sigma;\ \bar{x} + 2\sigma)$ за пределами интервала окажется 5 % измерений. И для интервала $(\bar{x} - 3\sigma;\ \bar{x} + 3\sigma)$ – только один результат из трехсот, то есть интервал $(\bar{x} - 3\sigma;\ \bar{x} + 3\sigma)$ является почти достоверным, так как большинство отдельных результатов измерения случайной величины окажется сосредоточенным в интервале.

При обработке результатов эксперимента часто используют «правило 3σ» (критерий «трех сигм», правило «трех стандартов»), которое основано на указанном свойстве нормального распределения. Можно установить наличие промаха в результате отдельного измерения, а значит, отбросить его, если результат измерения более чем на 3σ отличается от измеренного среднего значения случайной величины.

Методы обнаружения и исключения грубых погрешностей

Вопрос о том, содержит ли результат наблюдений грубую погрешность, решается общими методами проверки статистических гипотез.

Проверяемая гипотеза состоит в утверждении, что результат наблюдения x_i не содержит грубой погрешности, т.е. является одним из значений измеряемой величины. Пользуясь определенными статистическими критериями, пытаются опровергнуть выдвинутую гипотезу. Если это удается, то результат наблюдений рассматривают как содержащий грубую погрешность и его исключают.

Для выявления грубых погрешностей задаются вероятностью q (*уровнем значимости*) того, что сомнительный результат действительно мог иметь место в данной совокупности результатов измерений. Обычно проверяют наибольшее и наименьшее значения результатов измерений. Для проверки гипотез используются следующие критерии.

Критерий «трех сигм» применяется для результатов измерений, распределенных по нормальному закону. Данный критерий надежен при числе измерений n > 20... 50. По этому критерию считается, что результат маловероятен и его можно считать промахом, если:

$$\left|\bar{x} - x_i\right| > 3\,\sigma, \tag{9}$$

где $\bar{x}$ – среднее арифметическое отдельных результатов измерений x_i; σ – среднее квадратичное отклонение. Величины $\bar{x}$ и σ вычисляют без учета экстремальных значений x_i.

Критерий Романовского применяется, если число измерений n < 20. При этом вычисляется отношение:

$$\frac{\left|\bar{x} - x_i\right|}{\sigma} = \beta, \tag{10}$$

которое сравнивается с критерием $\beta_{теор}$, выбранным по табл. 2. Величины $\bar{x}$ и σ вычисляют без учета экстремальных значений x_i.

Таблица 2

Теоретические значения $\beta_{\text{теор}}$ критерия Романовского

Вероятность	Число измерений						
q	$n = 4$	$n = 6$	$n = 8$	$n = 10$	$n = 12$	$n = 15$	$n = 20$
0,01	1,73	2,16	2,43	2,62	2,75	2,90	3,08
0,02	1,72	2,13	2,37	2,54	2,66	2,80	2,96
0,05	1,71	2,10	2,27	2,41	2,52	2,64	2,78
0,10	1,69	2,00	2,17	2,29	2,39	2,49	2,62

Если $\beta \geq \beta_{теор}$, то результат x_i считается промахом и отбрасывается.

Вариационный критерий Диксона удобный и достаточно мощный (с малыми вероятностями ошибок). При его применении полученные результаты наблюдений записывают в вариационный (упорядоченный по возрастанию) ряд $x_1, x_2, \ldots x_n$, причем ($x_1 < x_2 < \ldots < x_n$).

Критерий Диксона определяется как:

$$K_{Д} = \frac{x_n - x_{n-1}}{x_n - x_1}. \qquad (11)$$

Критическое условие: $K_{Д} > Z_q$. Значения Z_q приведены в табл. 3.

Таблица 3

Значения критерия Диксона

n	Z_q при q, равном			
	0,10	0,05	0,02	0,01
4	0,68	0,76	0,85	0,89
6	0,48	0,56	0,64	0,70
8	0,40	0,47	0,54	0,59
10	0,35	0,41	0,48	0,53

Критерии Шовине. Этот критерий применяется для отбрасывания грубых ошибок при нормальном распределении контролируемого параметра. По критерию Шовене отбрасывают только одно сомнительное значение. При использовании этого критерия рассчитывается вероятность получения значения, отклоняющегося от среднего значения выборки больше, чем сомнительное значение. Этот критерий может быть использован, если число измерений n < 10. В этом случае грубой ошибкой (промахом) считается результат x_i, если разность $|\bar{x} - x_i|$ превышает значения σ, определяемые в зависимости от числа измерений:

$$|\bar{x} - x_i| > \begin{cases} 1{,}6\,\sigma \text{ при } n = 3 \\ 1{,}7\,\sigma \text{ при } n = 6 \\ 1{,}9\,\sigma \text{ при } n = 8 \\ 2{,}0\,\sigma \text{ при } n = 10. \end{cases} \qquad (12)$$

Задания для выполнения

Задание 1. Для совокупности измерений (n=30), приведенных в табл. 4, используя критерий «трех сигм», проверить, являются ли следующие значения промахами:

– первый вариант – значение 0,959868;
– второй вариант – значение 47,40715;
– третий вариант – значение 98,05261;
– четвертый вариант – значение 78,60488;
– пятый вариант – значение 177,5341;
– шестой вариант – значение 299,4049;
– седьмой вариант – значение 368,2737;
– седьмой вариант – значение 89,65575.

Построить график частотного распределения – зависимость частоты появления значений (ось y) от измеряемой величины (ось x).

Если неравенство (9) не выполняется, то это означает, что подверженный сомнению результат измерений не является результатом грубой погрешности.

Таблица 4

Данные измерений для $n = 30$

1 вариант	2 вариант	3 вариант	4 вариант	5 вариант	6 вариант	7 вариант	8 вариант
0,288430	33,77514	86,55232	83,51146	157,1722	222,1259	330,8771	68,30683
0,628925	31,37394	64,01318	97,49077	135,4015	246,5495	326,899	58,63369
0,788781	15,37492	58,97855	82,44881	129,2953	264,5314	330,7367	50,01038
0,432691	14,60219	75,96667	86,91366	160,8783	245,0182	328,4433	58,51772
0,628224	18,36085	82,59163	85,40788	166,2679	265,6575	327,8603	65,93188
0,667074	17,35252	95,24064	94,52742	165,9053	241,1756	316,4266	52,33772
0,152257	45,80432	28,84487	82,57576	156,9402	281,1066	321,9825	54,82437
0,957244	15,48479	68,96164	94,58296	141,9085	265,5293	321,4133	69,30967
0,062868	**47,40715**	85,66240	92,98318	150,0845	238,6456	310,2451	68,83724
0,777490	30,52797	74,80667	95,87878	149,3928	273,1071	301,0956	64,85336
0,259346	39,17570	20,45106	98,83114	154,6909	259,4073	325,3304	52,37373
0,276223	25,19425	90,79043	87,27195	166,8923	228,0862	330,8023	55,00748
0,319926	33,26243	67,91894	83,40159	144,6693	**299,4049**	**368,2737**	65,80798
0,959868	24,64278	67,25150	**78,60488**	157,1648	223,0537	305,9786	64,44319
0,165593	21,82745	62,25532	92,36305	157,3315	280,2576	324,1874	64,01349
0,654134	30,02014	90,35371	96,54836	158,5186	251,2955	305,5315	63,12052
0,202429	32,29072	87,35832	86,93747	125,9481	278,9544	325,7256	69,60082
0,845637	31,48503	80,40254	92,97708	162,5098	287,5573	331,0312	56,98569
0,562883	26,25904	59,18647	95,09873	140,7678	255,3117	316,7547	**89,65575**
0,436934	33,74523	95,78387	82,41890	170,6746	276,4092	331,7881	51,26469
0,338328	29,19492	84,11451	86,19160	147,3516	274,8589	313,7028	65,30747
0,760155	39,10092	88,81680	87,30918	152,8416	249,8917	331,0327	63,39152
0,526261	40,68117	95,39933	87,73583	165,7009	290,7956	332,4671	68,02667
0,404370	41,86316	89,24955	81,86407	**177,5341**	283,6232	314,9419	68,89584
0,927793	42,07639	64,95529	83,00974	125,5940	263,9003	341,0489	57,65893
0,323344	43,58745	76,98111	89,71221	126,3192	254,6082	315,5461	60,17853
0,422315	33,35188	**98,05261**	86,99179	162,6539	280,3488	338,3663	67,43217
0,218970	27,92276	63,73882	89,58708	138,3923	293,0601	341,6822	67,51335
0,196387	29,33256	66,96036	88,72158	156,1248	256,6265	309,2517	52,90414
0,323564	39,17892	47,98111	89,88128	140,5667	276,3787	320,4505	56,34175

Задание 2. Используя критерий Романовского, проверить для ряда измерений, приведенных в табл. 5, являются ли следующие значения промахами:

– первый вариант (n = 10) – значение 68,76;

– второй вариант (n = 14) – значение 69,39;

– третий вариант (n = 7) – значение 96,19;

– четвертый вариант (n = 12) – значение 480,45;

– пятый вариант (n = 14) – значение 0,25.

Таблица 5

Данные измерений для $n = var$

Число измерений				
n = 10	n = 14	n = 7	n = 12	n = 14
68,76	57,44	80	521,84	8,79
59,82	58,58	81,27	548,38	9,42
64,25	51,59	84,39	526,75	8,78
56,07	**69,39**	91,16	505,34	4,32
61,41	57,68	93,27	520,4	4,35
64,61	62,81	94,33	511,76	2,39
62,45	62,27	**96,19**	535,97	9,15
61,35	52,15	–	549,88	4,84
53,37	69,36	–	589,49	8,42
63,24	52,41	–	508,69	4,53
–	63,49	–	555,74	9,71
–	66,61	–	**480,45**	3, 33
–	67,18	–	–	6,22
–	68,87	–	–	**0,25**

Задание 3. Было проведено десять измерений массы галетного печенья из каждой партии (вариант задания) – табл. 6. Используя вариационный критерий Диксона, определить для каждой партии «подозрительную» массу галетного печенья:

– первая партия (1 вариант) – значение 7,84;

– вторая партия (2 вариант) – значение 9,36;

– третья партия (3 вариант) – значение 10,19;

– четвертая партия (4 вариант) – значение 6,98;

– пятая партия (5 вариант) – значение 6,35.

Таблица 6

Масса галетного печенья

Число измерений n = 10				
1 вариант	2 вариант	3 вариант	4 вариант	5 вариант
8,03	8,35	**10,19**	8,35	8,04
8,42	8,52	8,44	6,91	8,05
8,13	8,63	8,35	8,63	8,12
8,44	8,64	8,52	8,64	8,93
8,16	8,03	8,63	8,33	8,64
7,84	8,42	8,64	8,35	8,44
8,47	8,13	8,35	8,52	**6,35**
8,07	**9,36**	8,52	8,63	8,52
8,19	8,19	8,63	8,64	8,63
8,36	9,30	8,64	**6,98**	8,64

Задание 4. При измерении линейного размера пяти деталей, взятых из одной партии, были получены следующие данные: 25,52 мм; 25,50 мм; 25,49 мм; 25,66 мм; 25,53 мм. Пользуясь критерием Шовине, проверить, является ли размер 25,66 мм промахом.

Содержание практической работы

1. Наименование и цель работы.
2. Основные теоретические положения, расчетные формулы.
3. Выполненные расчетные задания.
4. Выводы по работе.

Вопросы для подготовки к защите практической работы

1. Что такое абсолютная и относительная погрешности измерения?
2. Что такое приборная, модельная и случайная погрешности?
3. Что характеризуют средним значением и стандартным квадратичным отклонением?
4. Как определить присутствие в выборке грубых погрешностей при однократных и многократных измерениях?
5. Почему нормальное распределение чаще других встречается в экспериментах?
6. Какой смысл придают понятиям доверительной вероятности и доверительного интервала?
7. Какой критерий для определения наличия систематических погрешностей следует использовать при проведении многократных измерений, состоящих из нескольких серий?
8. Как применить критерий Романовского для исключения из выборки промахов?
9. Число измерений более 20. Какой выбрать критерий для получения более объективного результата по исключению грубых погрешностей?
10. Что представляет собой вероятность q, называемая уровнем значимости?
11. С помощью какого критерия можно выяснить источники возникновения систематических погрешностей?
12. Для обнаружения грубой погрешности использовали критерий Романовского. По какому критерию проверить правильность расчетов?
13. Можно ли устранить постоянную систематическую погрешность путем увеличения числа измерений?
14. В чем заключается цель данного практического занятия?
15. При числе данных опытов, равном 25, какой критерий лучше выбрать для обнаружения грубой погрешности?

Практическая работа 2

МЕТОДЫ ПРОГНОЗИРОВАНИЯ

Цель работы

Углубление теоретических знаний о методах прогнозирования, приобретение практических навыков в прогнозировании временных рядов.

Общие сведения

Современное общество постоянно испытывает необходимость в прогнозировании. Под *прогнозом* понимается научно обоснованное описание возможных состояний систем в будущем и сроков достижения этого состояния, а процесс разработки прогнозов называют *прогнозированием*.

Например, прогнозируют уровни безработицы, инфляции, промышленного производства, подоходного налога и пр. с целью выстраивания государственной политики. Прогнозируют потребность в оборудовании и персонале в различных отраслях промышленности в целом, и в отдельных производствах в частности. Прогнозируют объемы автомобильных, железнодорожных, авиаперевозок. В средних и высших учебных заведениях прогнозируют количество абитуриентов на следующий учебный год. При разработке краткосрочных и долгосрочных планов менеджеры прогнозируют объемы продаж, издержки производства, процентные ставки и т.д. Каким бы видом производства или бизнеса ни занималась фирма или компания, ей приходится планировать предпринимательскую деятельность на будущий период. В настоящее время также широко используют методы прогнозирования в науке. При изучении свойств сложных систем, в том числе при экспериментальных исследованиях, когда математически описать изучаемый процесс практически невозможно, но имеется некоторая характерная наблюдаемая величина, широко используют методы прогнозирования. Например, в химической кинетике (анализ динамики процессов; построение моделей и пр.), в сейсмологии (запись колебаний земной коры), в метеорологии (данные метеонаблюдений) и т.п.

В настоящее время разработано несколько различных методов прогноза. Одни из них более, другие менее точны в прогнозах. Принципиально существуют два общепринятых подхода к прогнозированию: *качественный* и *количественный* – представлены на рис. 1.

Методы качественного прогнозирования применяют в случае, если недоступны количественные данные. Понятно, эти методы носят субъективный характер и не всегда дают точный прогноз.

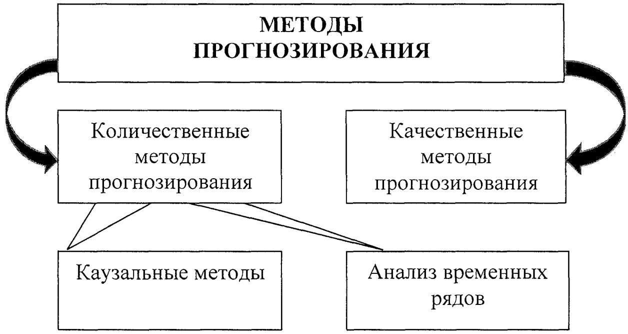


Рис. 1. Классификация методов прогнозирования

Если известны значительные количества реальных значений рассматриваемого показателя, а проецируемая в прошлом тенденция устойчива и стабильна, следует применять *методы количественного прогнозирования.* Задачей данного метода является предсказание по данным имеющихся наблюдений будущих значений измеряемых характеристик изучаемого объекта, то есть составление прогноза на некоторый отрезок времени вперед. Эти методы позволяют предсказать состояние объекта в будущем на основе данных о его прошлом при условии отсутствия вмешательства дополнительных факторов извне. По существу, *методы количественного прогнозирования* являются реконструкцией динамических систем.

Методы количественного прогнозирования разделяются на две категории: анализ временных рядов и методы анализа причинно-следственных зависимостей – так называемые каузальные методы – см. рис. 1.

Методы анализа причинно-следственных зависимостей или каузальные методы позволяют определить, какие факторы влияют на значения прогнозируемой переменной. К ним относятся методы множественного регрессионного анализа с запаздывающими переменными, эконометрическое моделирование, анализ лидирующих индикаторов, методы анализа диффузионных индексов и других экономических показателей.

Временный ряд (Time Series) представляет собой последовательность данных, описывающих объект в последовательные моменты времени, это динамический ряд, у которого в качестве признака используется время, то есть его можно назвать также хронологическим рядом. Временной ряд существенно отличается от простой выборки данных, так как при анализе учитывается взаимосвязь измерений со временем, а не только статистическое разнообразие и статистические характеристики выборки. Теоретические исследования, основанные на анализе временных рядов, дают мощный инструмент для понимания многих явлений, когда имеющихся данных для построения модели недостаточно.

Существует две основные цели *анализа временных рядов*:

– определение структуры ряда;

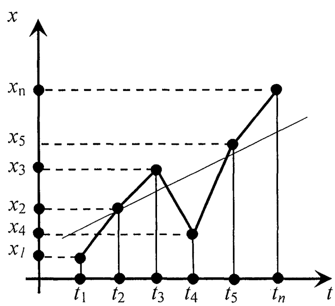
– прогнозирование.

Выявление структуры временного ряда необходимо для того, чтобы построить математическую модель того явления, которое является источником анализируемого временного ряда.

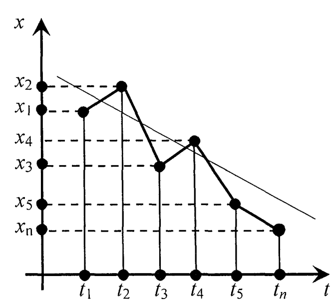
Прогнозирование временных рядов заключается в построении модели для предсказания будущих событий, основываясь на известных событиях прошлого, т.е. предсказания будущих данных до того, как они будут измерены. Прогноз будущих значений используется для эффективного принятия решений.

Элементы временного ряда называют *уровнями временного ряда.* Количество входящих в ряд уровней n называют его длиной $x(t)$, где $t = 1,\ldots, n$. Анализ временных рядов предполагает, что данные содержат систематическую составляющую и случайный шум (ошибку), который затрудняет обнаружение регулярных компонент. Компоненты систематической составляющей данных временного ряда отражают закономерность развития: 1 – *тренд* или долговременная тенденция развития; 2 – *сезонная компонента*; 3 –*циклическая компонента*; 4 – *остаточная компонента*.

Таким образом развитие реальных процессов можно рассматривать как совокупность некоторой устойчивой тенденции (тренда) и некоторой случайной составляющей, выраженной в колебании значений показателя вокруг линии тренда – рис. 2.



а) положительный тренд



а) отрицательный тренд

Рис. 2. Характер трендов временных рядов

Различают два вида временных рядов:

– моментные – значения рассматриваемого показателя $x_1, x_2, .., x_n$ отнесены к определенным моментам времени $t_1, t_2, \ldots, t_n$, причем $t_n > t_{n-1}$. – приведено в табл. 1, рис. 3 а);

– интервальные – характеризуют значения показателя за определенные интервалы времени или значения рассматриваемого показателя отнесены к определенным интервалам времени $[t_0 - t_1]$, $[t_1 - t_2], \ldots, [t_{n-1} - t_n]$ – приведено в табл. 2, рис. 3 б).

Таблица 1

Моментный временный ряд

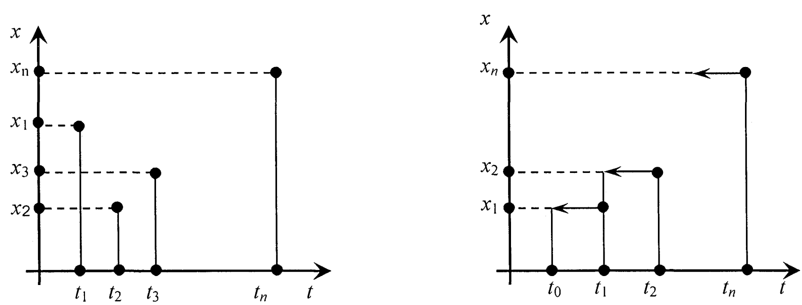
Момент времени	t_1	t_2	…	t_n
Значение показателя	x_1	x_2	…	x_n

Таблица 2

Интервальный временный ряд

Момент времени	$[t_0 - t_1]$	$[t_1 - t_2]$	…	$[t_{n-1} - t_n]$
Значение показателя	x_1	x_2	…	x_n

Производные ряды получаются из средних или относительных величин показателя.



а) моментный ряд

б) интервальный ряд

Рис. 3. Моментный и интервальный временные ряды

Если пренебречь сезонной, циклической и остаточной компонентами, то значения временного ряда можно представить как сумму двух составляющих – тренда и некоторой случайной составляющей:

ЗНАЧЕНИЕ X = ТРЕНД + СЛУЧАЙНАЯ СОСТАВЛЯЮЩАЯ

Более общая модель предполагает наличие вышеперечисленных составляющих, например, сезонной и циклической.

Нахождение тренда во временном ряду является своего рода искусством, его нельзя обнаружить автоматически или формально. Если характер изменения значений временного ряда является монотонным (устойчиво возрастает или убывает), то анализировать такой ряд несложно. Если же временный ряд содержит значительную ошибку, то обозначить тренд такого ряда затруднительно, поэтому необходимо предварительно выполнить сглаживание – процедуру локального усреднения данных, при которой несистематические компоненты взаимно погашают друг друга. На рис. 4 представлены наиболее применяемые методы сглаживания.

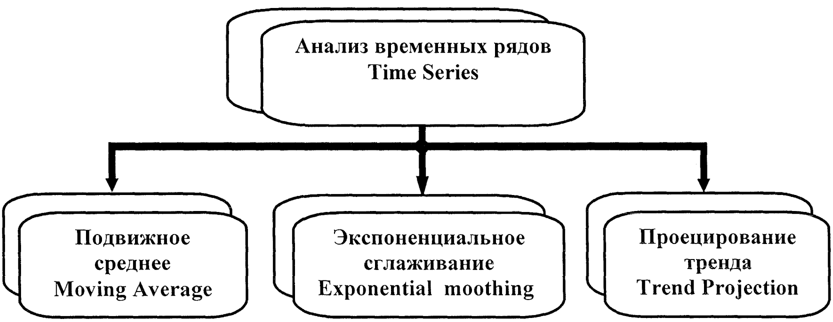


Рис. 4. Методы прогнозирования

Самый общий метод сглаживания – метод подвижного среднего. Этот метод разделяют на метод подвижного (скользящего) среднего и метод взвешенного (скользящего) среднего.

Метод подвижного среднего – расчет показателя на прогнозируемый момент времени осуществляется путем усреднения значений этого показателя за несколько предшествующих периодов времени. Расчетная формула имеет вид:

$$f_k = \frac{1}{n}\sum^{n} x_{k-1}, \tag{1}$$

где f_k – прогноз на момент времени t_k; x_{k-1} – значение показателя временного ряда в момент времени t_{k-1}; n – число периодов времени, предшествующих прогнозу.

Пример 1. Провести прогноз методом подвижного среднего нагрузки на телефонную сеть на период до понедельника следующей недели по имеющимся данным предшествующей недели, данные представлены в табл. 3.

Таблица 3

Количество звонков телефонной сети

День недели	Понедельник	Вторник	Среда	Четверг	Пятница	Суббота	Воскресенье
Порядковый номер дня t	1	2	3	4	5	6	7
Количество звонков, млн. x	4,4	3,8	3,2	4,7	4,1	2,9	2,3

Решение:

1) Рассчитаем прогнозируемое число звонков в соответствии с формулой (1), данные расчета сведем в табл. 4.

– прогнозируемое число звонков на четверг:

$$f_4 = \frac{4,4+3,8+3,2}{3} = 3,8.$$

– прогнозируемое число звонков на пятницу:

$$f_5 = \frac{3,8+3,2+4,7}{3} = 3,9.$$

– прогнозируемое число звонков на субботу:

$$f_6 = \frac{3,2+4,7+4,1}{3} = 4,0.$$

– прогнозируемое число звонков на воскресенье:

$$f_7 = \frac{4,7+4,1+2,9}{3} = 3,9.$$

– прогнозируемое число звонков на понедельник следующей недели:

$$f_8 = \frac{4,1+2,9+2,3}{3} = 3,1.$$

Данные расчета представим в табл. 4 и на рис. 5.

Таблица 4

Данные расчета по методу подвижного среднего

t	1	2	3	4	5	6	7	8
x	4,4	3,8	3,2	4,7	4,1	2,9	2,3	–
f	–	–	–	3,8	3,9	4,0	3,9	3,1

Метод взвешенного (скользящего) среднего – расчет показателя на прогнозируемый момент времени осуществляется с учетом весового показателя ξ_i каждого периода времени:

$$f_k = \frac{\sum^{n} \xi_{k-1} x_{k-1}}{\sum^{n} \xi_{k-1}} \qquad (2)$$

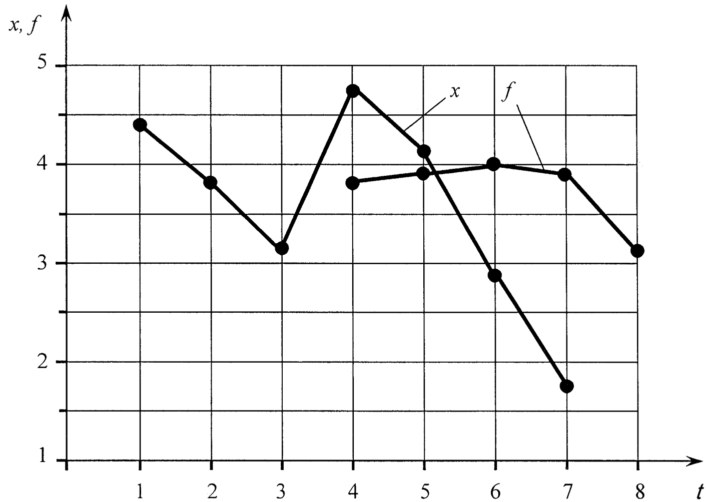


Рис. 5. График временного ряда x и прогноза f по методу подвижного среднего

Пример 2. Провести численный расчет прогноза методом взвешенного среднего в условиях предыдущего примера, приняв вес сегодняшнего показателя равен 0,6, вчерашнего – 0,3, позавчерашнего – 0,1 (не забываем – мы делаем прогноз в среду на четверг, пятницу и т.д. до понедельника следующей недели).

Решение:

По формуле (2) получим:

– прогнозируемое число звонков на четверг:

$$f_4 = \frac{0{,}1 \cdot 4{,}4 + 0{,}3 \cdot 3{,}8 + 0{,}6 \cdot 3{,}2}{0{,}1 + 0{,}3 + 0{,}6} = 3{,}5.$$

– прогнозируемое число звонков на пятницу:

$$f_5 = \frac{0{,}1 \cdot 3{,}8 + 0{,}3 \cdot 3{,}2 + 0{,}6 \cdot 4{,}7}{0{,}1 + 0{,}3 + 0{,}6} = 4{,}16.$$

– прогнозируемое число звонков на субботу:

$$f_6 = \frac{0{,}1 \cdot 3{,}2 + 0{,}3 \cdot 4{,}7 + 0{,}6 \cdot 4{,}1}{0{,}1 + 0{,}3 + 0{,}6} = 4{,}19.$$

– прогнозируемое число звонков на воскресенье:

$$f_7 = \frac{0,1 \cdot 4,7 + 0,3 \cdot 4,1 + 0,6 \cdot 2,9}{0,1 + 0,3 + 0,6} = 3,44.$$

– прогнозируемое число звонков на понедельник следующей недели:

$$f_8 = \frac{0,1 \cdot 4,1 + 0,3 \cdot 2,9 + 0,6 \cdot 2,3}{0,1 + 0,3 + 0,6} = 2,66.$$

Данные расчета представим в табл. 5 и на рис. 6.

Таблица 5

Данные расчета по методу взвешенного среднего

t	1	2	3	4	5	6	7	8
x	4,4	3,8	3,2	4,7	4,1	2,9	2,3	–
f	–	–	–	3,5	4,16	4,19	3,44	2,66

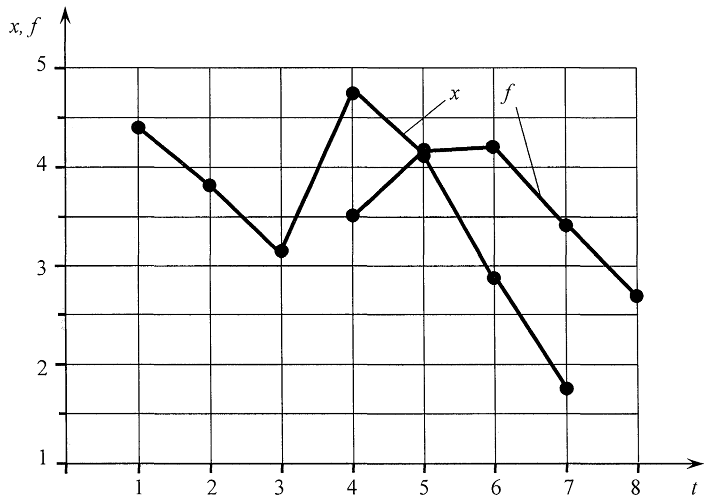


Рис. 6. График временного ряда $\boldsymbol{x}$ и прогноза f по методу взвешенного среднего

Метод экспоненциального сглаживания – расчет показателя на прогнозируемый момент времени осуществляется с учетом отклонения предыдущего прогноза от реального показателя по формуле:

$$f_k = f_{k-1} + \beta \cdot (x_{k-1} - f_{k-1}), \quad (3)$$

где β – постоянная сглаживания (0 < β < 1), коэффициент β выбирают с учетом практического опыта.

Пример 3. Провести прогноз методом экспоненциального сглаживания в условиях примера 1.

Решение:

Допустим, что в понедельник будет 4,3 млн. звонков. По формуле (3) рассчитаем прогнозируемые данные, приняв постоянную сглаживания $\beta = 0{,}2$. Отразим полученные результаты расчета в табл. 6 и на рис. 7.

Таблица 6

Данные расчета по методу взвешенного среднего

t	1	2	3	4	5	6	7	8
x	4,4	3,8	3,2	4,7	4,1	2,9	2,3	–
f	4,3	4,32	4,22	4,02	4,16	4,15	3,9	3,58

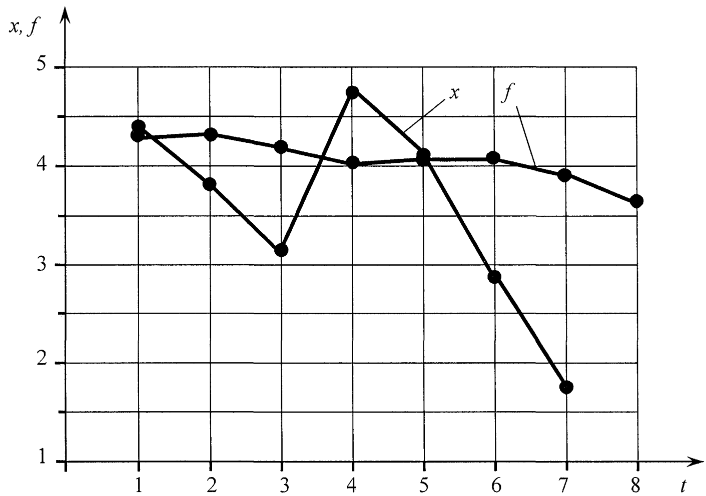


Рис. 7. График временного ряда x и прогноза f по методу экспоненциального сглаживания

Как видим по рис. 7, прогнозируемый тренд более сглажен, чем тренды, полученные по методу подвижного среднего – см. рис. 5 и рис. 6.

Метод проецирования тренда – в основе этого метода лежит идея построения прямой, которая наименее уклоняется от массива точек (t, x) заданного временного ряда, описываемого уравнением:

$$x = a \cdot t + b, \tag{4}$$

где a и $в$ – постоянные коэффициенты.

Расчет коэффициентов а и в ведется по методу наименьших квадратов, то есть решается система уравнений, которая имеет единственное решение:

$$\begin{cases} a \cdot \sum^{n} t_i + b \cdot n = \sum^{n} x_i \\ a \cdot \sum^{n} t_i^2 + b \sum^{n} t_i = \sum^{n} t_i \cdot x_i. \end{cases} \qquad (5)$$

Пример 4. Провести прогноз методом проецирования тренда в условиях примера 1.

Решение:

Нетрудно заметить, что метод проецирования тренда является прямым аналогом построения прямой регрессии, поэтому для проверки его надежности используют коэффициент линейной корреляции. Коэффициент линейной корреляции – параметр, который характеризует степень линейной взаимосвязи между двумя выборками. Он позволяет установить не только степень взаимосвязи, но и характер (прямая или обратная) этой взаимосвязи.

Используя данные табл. 3, рассчитаем промежуточные параметры, которые необходимы для определения коэффициентов a и b при решении системы уравнений (5), полученные результаты расчета представим в табл. 7.

Таблица 7

К расчету коэффициентов a и b

t_i	x_i	$t_i \cdot x_i$	t_i^2
1	4,4	4,4	1
2	3,8	7,6	4
3	3,2	9,6	9
4	4,7	18,8	16
5	4,1	20,5	25
6	2,9	17,4	36
7	2,3	16,1	49
Σ = 28	Σ = 25,4	Σ = 94,4	Σ = 140

Решим систему уравнений (5) и определим коэффициенты a и b:

$$28 \cdot a + 7 \cdot b = 25{,}4$$

$$140 \cdot a + 28 \cdot b = 94{,}4.$$

Таким образом, коэффициент, a = -0,26, коэффициент b = 4,66. По уравнению (4) – уравнению регрессии – можно осуществить прогноз количества звонков на любой момент времени, принадлежащий временному ряду:

$$x = -0{,}26 \cdot t + 4{,}66. \qquad (6)$$

Например, количество звонков на седьмой день составит:

$$f_7 = -0{,}26 \cdot 7 + 4{,}66 = 2{,}84.$$

Данные прогноза представим в табл. 8 и на рис. 8.

Таблица 8

Данные расчета по методу проецирования тренда

t	1	2	3	4	5	6	7	8
x	4,4	3,8	3,2	4,7	4,1	2,9	2,3	–
f	4,4	4,14	3,88	3,62	3,36	3,10	2,84	2,58

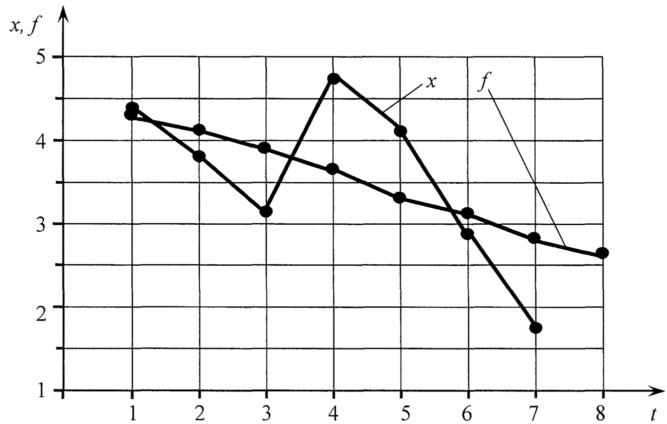


Рис. 8. График временного ряда x и прогноза f по методу проецирования тренда

Коэффициент линейной корреляции рассчитывается по формуле:

$$r_{xy} = \frac{\sum\left(x_i - \bar{x}\right)\cdot\left(y_i - \bar{y}\right)}{\sqrt{\sum\left(x_i - \bar{x}\right)^2 \cdot \sum\left(y_i - \bar{y}\right)^2}} \qquad (7)$$

где x_i — значения, принимаемые в выборке X, y_i — значения, принимаемые в выборке Y; $\bar{x}$ — средняя по X, $\bar{y}$ — средняя по Y.

Коэффициент корреляции изменяется от –1 до 1. Когда при расчете получается величина большая +1 или меньшая –1, это означает, что произошла ошибка в вычислениях. При значении 0 линейной зависимости между двумя выборками нет.

В табл. 9 приведены данные, учитывающие тесноту связи в прямой или обратной зависимости от величины коэффициента корреляции.

Таблица 9

Теснота связи и величина коэффициента корреляции

Коэффициент корреляции r_{xy}	Теснота связи
$\pm$ 0,91-1,0	Очень сильная
$\pm$ 0,81-0,9	Весьма сильная
$\pm$ 0,65-0,8	Сильная
$\pm$ 0,45-0,64	Умеренная
$\pm$ 0,25-0,44	Слабая
До $\pm$ 0,25	Очень слабая
«+» – прямая зависимость «–» – обратная зависимость	

Задание. Используя метод подвижного среднего (метод подвижного скользящего среднего и метод взвешенного среднего), метод экспоненциального сглаживания и метод проецирования тренда, осуществить прогноз данных, представленных в табл. 10. Результаты расчета представить в виде таблицы и на графике: на график временного ряда x нанести четыре графика прогнозов f в соответствии с четырьмя методами.

Таблица 10

Данные для расчета

1 вариант	Период времени t, час	8.00	9.00	10.00	11.00	12.00	13.00	14.00	15.00	16.00	17.00
	Объем продаж ценных бумаг x, млн.	4,5	7,5	9,7	15,8	17,1	8,4	10,3	12,2	6,2	5,5
2 вариант	Период Времени t, день	1	2	3	4	5	6	7	8	9	10
	Количество испытаний, x	4	7	6	9	12	10	8	7	13	4
3 вариант	Период времени t, день	1	2	3	4	5	6	7	8	9	10
	Количество опытов, x	12	10	8	17	10	12	16	8	7	15
4 вариант	Период времени t, месяц	1	2	3	4	5	6	7	8	9	10
	Количество авиарейсов, x	158	180	125	175	168	278	350	456	330	295
5 вариант	Период времени t, год	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
	Количество урожая, x, млн. т	45,6	65,8	55,9	53,4	59,5	60,2	51,9	69,4	78,5	79,3

Содержание практической работы

1. Наименование и цель работы.
2. Основные теоретические положения, расчетные формулы.
3. Выполненные расчетные задания.
4. Выводы по работе.

Вопросы для подготовки к защите практической работы

1. Что называют временным рядом?
2. Статистические методы прогнозирования и анализ временных рядов.
3. Цели анализа временных рядов.
4. Метод подвижного среднего при анализе временных рядов.
5. Метод экспоненциального сглаживания.
6. Метод проецирования тренда.
7. Казуальные и качественные методы прогнозирования.
8. Приведите примеры прогнозирования.
9. Какая идея лежит в основе метода проецирования тренда?
10. С какой целью применяют сглаживание временных рядов?
11. Какой метод дает более сглаженный прогнозируемый тренд?
12. Как определить весовой показатель периода времени в методе взвешенного (скользящего) среднего?
13. Назовите виды временных рядов.
14. Проиллюстрируйте моментный временный ряд.
15. Проиллюстрируйте интервальный временный ряд.

Практическая работа 3

ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА. ОБРАБОТКА ДАННЫХ

Цель работы

Углубление теоретических знаний о методах планирования эксперимента, приобретение практических навыков в планировании эксперимента с минимальным значением вариабельности при обработке данных.

Общие сведения

Магистрант в рамках освоения магистерской программы неизбежно сталкивается с проведением многочисленных экспериментов при выполнении лабораторных и научно-исследовательских работ, курсового проектирования. Не менее важно, чем правильно провести эксперимент, правильно обработать, интерпретировать результаты эксперимента.

Эксперимент как научное исследование – это форма, в которой и посредством которой наука существует и развивается.

Эксперимент предполагает осуществление тщательной подготовки перед его проведением или планирование эксперимента.

Под планированием эксперимента понимается процедура выбора числа опытов и условий их проведения, необходимых для решения поставленной задачи с требуемой точностью.

В основе современной науки о пище заложено множество наукоемких дисциплин (биотехнология, химия пищи, нанотехнологии и пр.), научные исследования в каждой из которой отличает широкой спектр вариабельности характеристик, параметров, свойств и пр., характерный для естественнонаучных дисциплин. И это является одной из причин возникающих трудностей при интерпретации результатов. Поэтому необходимо научиться правильно истолковывать данные и, что не менее важно, – правильно выбирать схемы эксперимента с минимальным значением вариабельности. Поэтому цель планирования эксперимента заключается в создании схемы, которая необходима для получения как можно большей информации при наименьших затратах для выполнения исследования. План, разработанный для исследования, характеризуется рядом свойств, которые приведены на рис. 1.

Задачи, для решения которых может использоваться планирование эксперимента, отличаются большим разнообразием. Это: исследование количественных данных диаграмм «состав – свойство»; поиск оптимальных условий различных процессов, систем; построение интерполяционных формул обработки данных; определение оптимальных методов решения теоретических задач и пр.

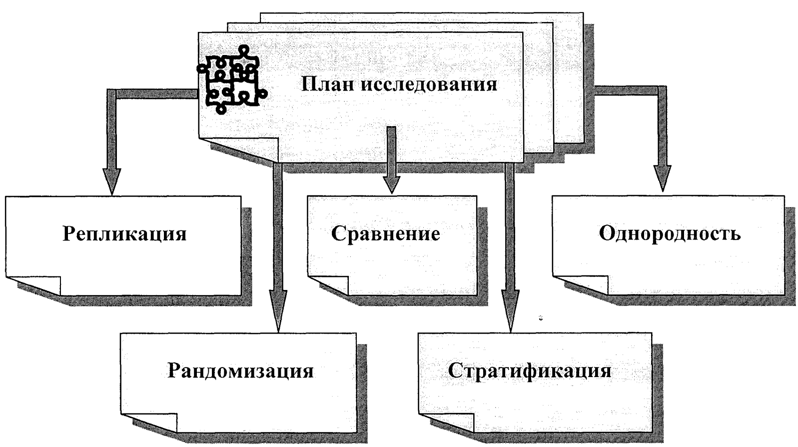


Рис. 1. Характеристики экспериментального плана научного эксперимента

Репликация или повторяемость – необходимый компонент постановки эксперимента. Результат каждого отдельного измерения содержит долю неопределенности, возникшей под влиянием множества случайных факторов. Поэтому необходимо провести несколько испытаний, чтобы определить возможный источник изменчивости.

Сравнение. Если невозможно точное определение результата измерения, то в научном эксперименте объект сравнивается либо с неким заданным стандартом, либо с контрольным объектом.

Однородность. Число измерений не должно быть слишком велико, чтобы не утратилась однородность процесса.

***Рандомизация* или порядок исследования.** Это процесс, используемый для группировки элементов таким образом, чтобы каждый элемент исследования наравне с остальными имел равную вероятность попасть в контрольную группу для исследования. Рандомизация необходима для анализа результатов исследования с применением статистических методов обработки результатов.

***Стратификация* или** распределение экспериментальных единиц в относительно однородные группы (блоки, слои). **Стратификация** позволяет минимизировать эффект известных неслучайных источников вариабельности. Внутри каждого блока ошибку эксперимента предполагают меньшей относительно варианта со случайным отбором для эксперимента такого же количества объектов.

Любое исследование начинается с выбора цели, которая определяет структуру научного исследования, его стратегию (косвенно, не всегда), а также выводы результатов исследования.

После выбора цели работы определяют зависимые переменные – это переменные, которые будут измеряться при проведении исследования.

Так как есть зависимые переменные, то должны быть еще и независимые переменные (факторы). Именно с факторами исследователь имеет дело при проведении исследований.

Обозначим $y_1, y_2, ..., y_n$ зависимые переменные, величина которых зависит от факторов (независимых переменных $x_1, x_2, ..., x_k$).

Функциональную зависимость можно записать в виде:

$$y = f(x_1, x_2, ..., x_k). \qquad (1)$$

Факторы, определяющие процесс, изменяются в соответствии с определенными алгоритмами. Для дальнейшей обработки данных их представляют в виде математической модели, этапы построения которой показаны на рис. 2.

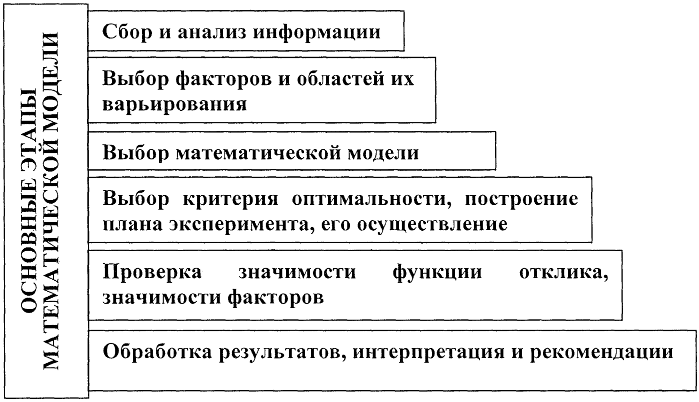


Рис. 2. Этапы построения математической модели

Функциональную зависимость (1) одной величины y от нескольких других x_i называют многофакторным экспериментом (МФЭ).

Каждый фактор может принимать одно из нескольких значений. Такие значения называют *уровнями фактора*. Если фактор способен принимать бесконечное число значений, то выбирается несколько дискретных уровней, количество которых зависит от конкретики задачи.

Есть несколько требований, предъявляемых к объекту изучения. Во-первых, *степень воспроизводимости* результатов опытов с данным объектом. С этой целью проводят эксперимент, затем повторяют его через неравные промежутки времени и сравнивают результаты. Если разброс значений не превышает требуемой точности, то объект исследования удовлетворяет требованию воспроизводимости.

Другое требование к объекту – это *управляемость*. Управляемым считается объект, на котором можно провести активный эксперимент или эксперимент, на течение которого можно влиять.

Существует также ряд требований к факторам:

– управляемости – фактор должен изменяться по требуемому закону или оставаться постоянным во время проведения опыта;

– совместимости – необходимо осуществление любых комбинаций факторов в пределах области их варьирования, все комбинации факторов осуществимы и безопасны;

– независимости друг от друга.

Фиксированный набор уровней факторов однозначно определяет возможные функциональные зависимости (количество возможных состояний) в соответствии с уравнением (1). Чтобы вычислить количество возможных состояний, надо число уровней факторов q возвести в степень количества факторов n:

$$N = q^n \tag{2}$$

Эксперимент, число опытов которого равно числу возможных сочетаний уровня плана, называется *полным факторным экспериментом* (ПФЭ).

Минимальное число уровней, обычно применяемое на первой стадии работы, равно 2. Очевидно, что для двухфакторной функции отклика число опытов ПФЭ составляет $N = 2^2 = 4$, для трехфакторной функции число опытов $N = 2^3 = 8$ и т.д, то есть наблюдается увеличение количества опытов. С увеличением числа уровней повышается чувствительность эксперимента, но одновременно возрастает число опытов в геометрической прогрессии. Поэтому уже на этапе планирования необходимо решить вопрос об оптимизации процесса исследования для решения поставленной задачи.

При планировании экспериментов чаще всего применяются планы 1-ого и 2-ого порядков. Планы более высоких порядков применяются редко из-за их сложности.

Планы 1-ого порядка – это планы, которые позволяют провести эксперимент для отыскания уравнения регрессии, содержащее только первые степени факторов и их произведения.

Планы 2-ого порядка – это планы, которые позволяют провести эксперимент для отыскания уравнения регрессии, содержащие вторые степени факторов.

Рассмотрим в качестве примера симметричный двухуровневый план двухфакторной функции, для которого число возможных сочетаний равно $N = 2^n$, где $n = 2$ – число факторов. Верхний и нижний уровни обозначают в кодированных координатах через +1 и –1. Условия эксперимента представим в виде таблицы – матрицы планирования, где строки соответствуют различным опытам, а столбцы — значениям факторов. Пример матрицы планирования для ПФЭ 2^2 приведен в табл. 1.

Таблица 1

Матрица планирования эксперимента

Номер опыта	x_1	x_2	y
1	–1	–1	y_1
2	+1	+1	y_2
3	–1	+1	y_3
4	+1	–1	y_4

Геометрическая интерпретация ПФЭ плана 2^2 задается координатами вершин квадрата – представлена на рис. 3.

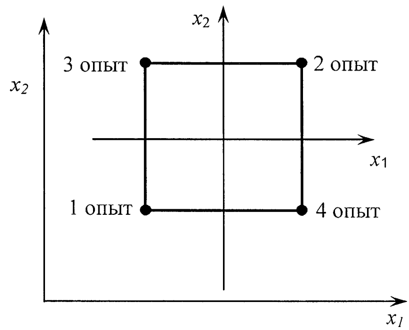


Рис. 3. Геометрическая интерпретация полного факторного эксперимента 2^2

Геометрическая интерпретация для ПФЭ 2^3 – представлена на рис. 4.

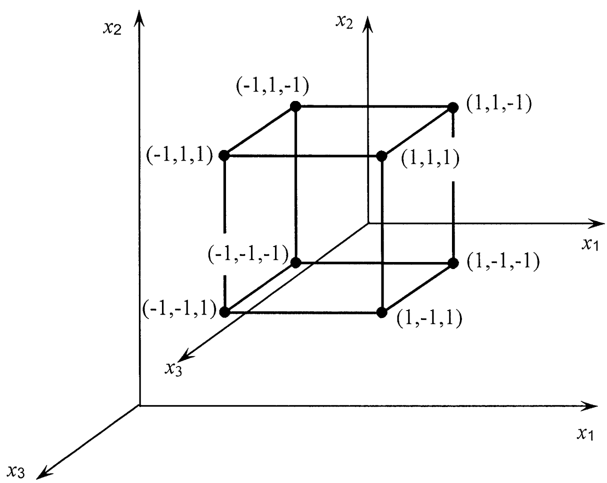


Рис. 4. Геометрическая интерпретация полного факторного эксперимента 2^3

При планировании эксперимента принимают, что простыми являются алгебраические полиномы. Наиболее часто применяются приведенные ниже полиномы.

Полином первой степени:

$$y = b_0 + \sum_{1}^{k} b_i x_i + \sum_{1}^{k} b_{ij} x_i x_j \tag{3}$$

Таким образом уравнение регрессии первого порядка для $k = 3$ имеет вид:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 . \tag{4}$$

Полином второй степени:

$$y = b_0 + \sum_1^k b_i x_i + \sum_1^k b_{ij} x_i x_j + \sum_1^k b_{ii} x_i^2 . \tag{5}$$

Здесь y – значения критерия; b – коэффициенты регрессии: b_i – линейные коэффициенты; b_{ij} – коэффициенты двойного взаимодействия; x_i – кодированные значения факторов (+1; –1).

Каждый столбец в матрице планирования называют вектор-столбцом, а каждую строку – вектор-строкой. Таким образом, в табл. 1 два вектора-столбца независимых переменных и один вектор-столбец параметра оптимизации.

Для оценки коэффициентов b применяют метод наименьших квадратов. Приведем формулы для расчета коэффициентов:

$$b_1 = \frac{\sum_{j=1}^{n} x_{1j} y_j}{n}; \; b_2 = \frac{\sum_{j=1}^{n} x_{2j} y_j}{n}; \; b_0 = \frac{\sum_{j=1}^{n} x_{0j} y_j}{n}; \; b_{12} = \frac{\sum_{j=1}^{n} (x_1 \cdot x_2)_j y_j}{n}. \tag{6}$$

Приведем пример расчета. Пусть таблично задана матрица планирования с учетом взаимодействия факторов x_1 и x_2 – табл. 2.

Таблица 2

Матрица планирования эксперимента 2^2

Номер опыта	x_1	x_2	$x_1 \cdot x_2$	y
1	+1	+1	+1	y_1
2	–1	+1	–1	y_2
3	+1	–1	–1	y_3
4	–1	–1	+1	y_4

Рассчитаем коэффициенты регрессии для данной матрицы планирования:

$$b_0 = \frac{y_1 + y_2 + y_3 + y_4}{4}; \; b_1 = \frac{+y_1 - y_2 + y_3 - y_4}{4};$$

$$b_2 = \frac{+y_1 + y_2 - y_3 - y_4}{4}; \; b_{12} = \frac{+y_1 - y_2 - y_3 + y_4}{4}.$$

Задание для выполнения

При проведении эксперимента в каждом опыте измерялись значения трех (n = 3) независимых переменных (x_1, x_2, x_3), которые определяли результат каждого опыта y. Причем y измеряли также по три раза в каждом опыте для увеличения степени достоверности. В соответствии с теорией планирования эксперимента, когда число опытов равно числу возможных сочетаний уровня плана, было проведено восемь опытов. Данные опыта приведены в табл. 3.

Таблица 3

Результаты непосредственных измерений

№ опыта	x_1	x_2	x_3	y_1	y_2	y_3
1	11,50	0,055	2,02	122	108	142
2	12,35	0,065	1,98	125	109	143
3	10,48	0,046	1,89	115	115	141
4	9,68	0,063	2,05	114	114	138
5	11,45	0,059	2,08	116	115	139
6	12,36	0,057	2,03	122	116	138
7	10,12	0,062	2,14	116	110	141
8	9,98	0,067	1,93	123	108	137

Интервалы варьирования факторов представлены в табл. 4.

Таблица 4

Интервалы варьирования факторов

Фактор	Уровень –1	Уровень +1
x_1	$< 10{,}5$	$> 10{,}5$
x_2	$> 0{,}060$	$< 0{,}060$
x_3	$< 2{,}01$	$> 2{,}01$
Отклик		
y		
Тип плана	Полный факторный (варьирование факторов на 2 уровнях)	
Количество повторных опытов	3	

Заполнив матрицу планирования (табл. 5), определите коэффициенты линейных членов модели и эффекты парных взаимодействий (заполните табл. 6, запишите регрессионную модель в виде:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 .$$

Таблица 5

Матрица планирования и результаты непосредственных измерений

№	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	y_1	y_2	y_3	$\overline{y}$
1								122	108	142	
2								125	109	143	
3								115	115	141	
4								114	114	138	
5								116	115	139	
6								122	116	138	
7								116	110	141	
8								123	108	137	

Таблица 6

Результаты статистического анализа эксперимента

Коэффициенты	Оценка	Вывод
b_0		
b_1		
b_2		
b_3		
b_{12}		
b_{13}		
b_{23}		

ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА

Методические указания

Составитель **Прохасько** Любовь Савельевна

Техн. редактор *А.В. Миних*

Издательский центр Южно-Уральского государственного университета

Подписано в печать 11.12.2017. Формат 60×84 1/16. Печать цифровая.
Усл. печ. л. 2,09. Тираж 50 экз. Заказ 411/96.

Отпечатано в типографии Издательского центра ЮУрГУ.
454080, г. Челябинск, проспект Ленина, 76.