

ПЗ №1: Общие сведения об экспериментальных исследованиях

Аннотация: В лекции рассматриваются понятия эксперимента, математической модели исследования, выборочной оценки, статистической гипотезы. Даются основные понятия, относящиеся к математической модели, представление о корреляционной связи и её статистическое изучение.

Ключевые слова: [Доверительная](#)

[вероятность](#) , [Add](#), [create](#), [Data](#), [editor](#), [input](#), [logic](#), [math](#), [net](#), [OPEN](#), [programmable logic device](#), [rules](#), [surface](#), [target](#), [XML](#), [агрегация среднего значения](#), [в диапазоне от x до y](#), [генератор](#), [двойной щелчок](#), [жизненный цикл](#), [зависимость между признаками](#), [имя переменной](#), [ключевое слово](#), [логический](#), [маршрутам](#), [надежность](#), [область определения](#), [память](#), [работам](#), [связь](#), [таблица](#), [универсального множества](#), [фазификация фактических данных](#), [частотность](#), [шкала](#), [элементарные приемы](#), [язык программирования](#)

Введение

На начальном этапе любой научной работы мы проводим эксперименты. Результаты таких экспериментов требуют всевозможных статистических и логических обработок, которые связаны подчас со сложными математическими выкладками и программными решениями.

Для решения задач обработки данных используются различные статистические методы: проверка гипотез, оценивание параметров и числовых характеристик случайных величин и процессов, корреляционный и дисперсионный анализ. Чтобы решение задачи было проведено качественно, необходимо провести предварительную обработку данных, для того чтобы не возвращаться повторно к решению той или иной задачи после получения результатов на последующем этапе обработки. Чтобы понять, с помощью, какой программы, или каким методом проводить обработку данных, необходимо помнить, что при наблюдении и проведении эксперимента встречаются ошибки грубые, систематические и случайные. В зависимости от точности и сложности эксперимента выбираются и методы обработки данных. Если эксперимент не предполагает особой точности и сложности можно выбрать простую программу и провести статистическую обработку, если же необходима высокая точность измерения, то и программу необходимо выбирать более сложную. Другими словами технология обработки экспериментальных данных зависит от того результата, к которому Вы стремитесь.

В настоящее время жизнь человека настолько насыщена различного рода информацией. Мы изучаем те или иные объекты исследуем их, а затем пытаемся объяснить с помощью математических формул и графиков, что мы получили.

В широком смысле слова эксперимент - это средство получения новых знаний об объектах реального мира, в какой-либо предметной области.

Кроме того, при проведении эксперимента мы стараемся автоматизировать его обработку, тем самым *деятельность* человека переходит из реального мира в электронный формат.

Многочисленная литература по обработке экспериментальных данных сводится к статистическому расчету этих данных с помощью математической модели. Более того, часто вводятся новые термины, порой не определяемые или имеющие разные названия в различных источниках. Много отдельных вопросов по обработке данных опубликованы в научных журналах, сложным математическим языком и со сложными математическими выкладками.

Следует отметить, что сейчас есть много всевозможных программ, которые не требуют таких академических знаний математики, для обработки полученных результатов вполне хватит знаний высшей школы и знания работы с программным продуктом.

Эксперимент, как таковой ничего не значит, если его правильно не обработать. После проведения эксперимента начинается самая сложная часть, это обработка его. Для того чтобы эксперимент считался законченным его необходимо просчитать. Как и каким способом сделать расчет? А затем представить его

результаты, и поможет данная книга. Кроме того, в нее включены задания по лабораторным работам и варианты практических *работ* по исследованию того или иного объекта.

Общие сведения об экспериментальных исследованиях

Цель лекции: Уяснить необходимость экспериментальных исследований наряду с теоретическими выводами. Ознакомиться с основными требованиями, которые предъявляются к экспериментальным исследованиям и обработке экспериментов с помощью математической модели. Рассмотрение требований, которые выполняются при проведении статистических наблюдений.

Эксперимент является важнейшим средством получения новых знаний не только в области естественных и технических наук, но и в экономике, социологии, политике, психологии, литературоведении и в других отраслях. Экспериментальные исследования дают критерии оценки обоснованности и приемлемости на практике любых теорий и теоретических предположений. Одним из основных этапов любого эксперимента является обработка экспериментальных данных. Она необходима для получения ответа на вопрос: "Достоверны ли полученные опытные данные в пределах требуемой точности или допусков?".

В зависимости от целей эксперимента (исследование, управление, *контроль*) могут быть использованы различные методы анализа объекта или явления. Это необходимо для прогнозирования состояния в разных условиях функционирования, оптимизации отдельных параметров, а также для решения каких-то других специфичных задач. Особенно важна тщательная математическая обработка результатов экспериментов, подтверждающая теоретические выводы и построения по диссертациям на соискание ученых степеней.

Применение различных методов обработки экспериментальных данных, критериев достоверности и адекватности моделей, изучаемым процессам или явлениям, оценка точности и надежности результатов эксперимента требует знания основных положений теории вероятностей и математической статистики, умелого использования принципов и приемов программирования. Кроме того, в связи с усложнением алгоритмов обработки данных необходимы глубокие знания основных вычислительных методов.

Конечной целью любой обработки экспериментальных данных является выдвижение гипотез о классе и структуре математической модели исследуемого явления, *определение* состава и объема дополнительных измерений, выбор возможных методов последующей статистической обработки и *анализ* выполнения основных предпосылок, лежащих в их основе.

Остановимся на определении математической модели (М.м.).

Математическая модель – это приближённое описание какого-либо класса явлений внешнего мира, выраженное с помощью математической символики. М. м. — мощный метод познания внешнего мира, а также прогнозирования и управления. *Анализ* М. м. позволяет проникнуть в сущность изучаемых явлений. Процесс математического моделирования, то есть изучения явления с помощью М. м., можно подразделить на 4 этапа.

Первый этап — формулирование законов, связывающих основные объекты модели. Этот этап требует широкого знания фактов, относящихся к изучаемым явлениям, и глубокого проникновения в их взаимосвязи. Эта стадия завершается записью в математических терминах сформулированных качеств, представлений о связях между объектами модели.

Второй этап — исследование математических задач, к которым приводят М. м. Основным вопросом здесь является решение *прямой* задачи, то есть получение в результате анализа модели выходных данных для дальнейшего их сопоставления с результатами наблюдений изучаемых явлений. На этом этапе важную роль приобретают математический аппарат, необходимый для анализа М. м., и вычислительная техника.

Третий этап — выяснение того, удовлетворяет ли принятая гипотетическая модель критерию практики, то есть выяснение вопроса о том, согласуются ли результаты наблюдений с теоретическими следствиями модели в пределах точности наблюдений. Если модель была вполне определена — все параметры её были заданы, — то *определение* уклонений теоретических следствий от наблюдений даёт решения *прямой* задачи с последующей оценкой уклонений. Если уклонения выходят за пределы точности наблюдений, то модель не может быть принята.

Четвёртый этап — последующий *анализ* модели в связи с накоплением данных об изучаемых явлениях и модернизация модели. В процессе развития науки и техники, данные об изучаемых явлениях, всё более и более уточняются, и наступает момент, когда выводы, получаемые на основании существующей М. м., не соответствуют нашим знаниям о явлении. В таком случае возникает необходимость построения, более совершенной математической модели.

Метод математического моделирования, сводящий исследование явлений внешнего мира к математическим задачам, занимает ведущее *место* среди других методов исследования, особенно в связи с появлением ЭВМ. Он позволяет проектировать новые технические средства, работающие в оптимальных режимах, для решения сложных задач науки и техники; проектировать новые явления. М. м. проявили себя как важное средство управления. Они применяются в самых различных областях знания, стали необходимым аппаратом в области экономического планирования и являются важным элементом автоматизированных систем управления.

Для разработки современной М.м. необходимо решить некоторые частные задачи, среди которых можно выделить следующие:

1. Анализ, выбраковка и восстановление аномальных (сбитых) или пропущенных измерений. Эта задача связана с тем, что исходная экспериментальная информация обычно неоднородна по качеству. В основной массе результатов прямых измерений, получаемых с возможно малыми погрешностями, в экспериментальных данных часто имеются грубые ошибки, вызванные разными причинами. К ним могут быть отнесены просчеты экспериментатора, сбои вычислительной техники, аномалии в работе измерительных приборов и т. д. Без глубокого анализа качества данных, устранения или хотя бы существенного уменьшения влияния аномальных данных на результаты последующей обработки можно сделать ложные выводы об изучаемом объекте или явлении.
2. Экспериментальная проверка законов распределения экспериментальных данных, оценка параметров и числовых характеристик наблюдаемых случайных величин или процессов. Выбор методов последующей обработки, направленной на построение и проверку адекватности выбранной модели исследуемому явлению, существенно зависит от закона распределения наблюдаемых величин. Получаемые при решении задачи выводы о природе экспериментальных данных могут быть как общими (независимость измерений, их равнозначность, характер погрешностей и др.), так и содержать детальную информацию о статистических свойствах данных (вид закона распределения, его параметры). Решение центральной задачи предварительной обработки не является чисто математическим, а требует также и содержательного анализа изучаемого процесса, схемы и методики проведения эксперимента.
3. Группировка исходной информации при большом объеме экспериментальных данных. При этом должны быть учтены особенности их законов распределения, которые выявлены на предыдущем этапе обработки.
4. Объединение нескольких групп измерений, полученных, возможно, в различное время или в различных условиях, для совместной обработки.
5. Выявление статистических связей и взаимовлияния различных измеряемых факторов и результирующих переменных, последовательных измерений одних и тех же величин. Решение этой задачи позволяет отобрать те, переменные, которые оказывают наиболее сильное влияние на результирующий признак. Выделенные факторы используются для дальнейшей обработки, в частности, методами регрессионного анализа. Анализ корреляционных связей делает возможным выдвижение гипотез о структуре взаимосвязи переменных и, в конечном итоге, о структуре модели объекта исследований.

В ходе предварительной обработки, кроме указанных выше задач, часто решают и другие, имеющие частный характер: *отображение*, преобразование и унификацию типа наблюдений, визуализацию многомерных данных и др.

Следует отметить, что в зависимости от конечных целей исследования, сложности изучаемого явления и уровня априорной информации о нем объем задач, выполняемых в ходе предварительной обработки, может существенно изменяться. То же самое можно сказать и о соотношении целей и задач, которые решаются при предварительной обработке и на последующих этапах статистического анализа, направленных на построение модели явления.

Так, например, если целью эксперимента является изменение значения неизвестной, но заведомо постоянной величины путем прямых многократных измерений с помощью средства измерений с известными

характеристиками погрешностей, то полная обработка результатов измерения ограничивается простейшей предварительной обработкой данных (оценкой математического ожидания). В то же время, если измеряемая величина является переменной, а закон распределения погрешностей измерительного прибора неизвестен, то для решения конечной задачи потребуется проведение, как предварительной обработки данных, так и применение статистических методов исследования физических зависимостей.

Для решения задач предварительной обработки используются различные *статистические методы*: проверка гипотез, *оценивание* параметров и числовых характеристик случайных величин и процессов, корреляционный и дисперсионный *анализ*. Для предварительной обработки, оказывающей, как следует из сказанного, первостепенное влияние на качество решения конечных задач исследования, характерно итерационное решение основных задач, когда повторно возвращаются к решению той или иной задачи после получения результатов на последующем этапе обработки.

При обработке числовых массивов результатов эксперимента, как случайных величин, на практике применяют следующие выборочные оценки:

математическое ожидание

$$M_x = \frac{1}{N} \sum_{i=1}^N x_i = \int_{-\infty}^{+\infty} x f(x) dx \quad (1.1)$$

дисперсия

$$D_x = \frac{1}{N-1} \sum_{i=1}^N (x_i - M_x)^2 = \int_{-\infty}^{+\infty} (x_i - M_x)^2 f(x) dx \quad (1.2)$$

коэффициент асимметрии

$$A = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - M_x)^3}{\sigma^3} \quad (1.3)$$

коэффициент эксцесса

$$E = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - M_x)^4}{\sigma^4} - 3 \quad (1.4)$$

где x_i значение результата в i - ом опыте; N - число результатов в массиве; $\sigma_x = \pm \sqrt{D_x}$ - среднеквадратичное отклонение.

Производная оценка от величины математического ожидания и дисперсии является *коэффициент вариации*, определяемый в процентах по формуле:

$$V = \frac{\sigma_x}{M_x} \cdot 100 \quad (1.5)$$

Дисперсия, среднее квадратичное отклонение и *коэффициент вариации* являются количественными характеристиками, оценки рассеивания значений результатов эксперимента как случайной величины и применяются при изучении различных действий со случайным исходом. *Коэффициент асимметрии* и

коэффициент эксцесса являются характеристиками высшего порядка. Первый характеризует "скошенность распределения", а второй – степень его "островершинности"

Вычисленные по экспериментально наблюдаемым случайным величинам и случайным функциям статистические характеристики несут информацию не обо всей генеральной совокупности, которая в общем случае бесконечна, а лишь о некоторой ее части — выборке, элементы которой измерены с определенными ошибками. В связи с этим в результате эксперимента получают лишь некоторые оценки параметров генеральной совокупности.

Следовательно, и любая выборочная оценка — это *случайная величина*, *точность* определения которой и возможные при этом ошибки необходимо контролировать. Следует также иметь в виду, что вычисленные моменты распределения являются точечными оценками выборочных величин, так как каждый из них оценивает параметры генеральной совокупности с помощью единственного числа. Они позволяют судить о значении вычисленной статистической характеристики в данной точке и ничего не говорят о возможных пределах варьирования самой оценки.

К вычисляемым в результате эксперимента оценкам случайных величин предъявляются три основных требования: состоятельности, несмещенности и эффективности. Полагают, что оценка состоятельна, если с ростом объема выборки она стремится по вероятности к истинному значению, несмещена, если ее математическое ожидание стремится к истинному значению, и эффективна, когда оценка обладает наименьшим рассеянием по сравнению с любыми другими оценками. Из двух оценок эффективнее та, которая обладает меньшей дисперсией, т. е. значения которой рассеиваются в более узком интервале.

На уровень рассеивания оценок значительное влияние оказывают ошибки, имеющие *место* при эксперименте.

При выборочном наблюдении встречаются ошибки трех видов: грубые, систематические и случайные.

Грубые ошибки, отличающиеся большим отклонением от центра группирования выборки, отсеиваются на этапе первичного анализа материалов.

Точность измерений любой физической величины характеризуется, абсолютной $\Delta x = x - \bar{x}$ и относительной $\frac{\Delta x}{\bar{x}} \cdot 100$ ошибками (здесь \bar{x} - истинное значение), которые, в свою очередь, состоят из суммы систематических δ и случайных ε ошибок

Систематические ошибки δ постоянны при определении каждого члена выборки и зависят от технического уровня измерительной аппаратуры и техники эксперимента. Эти ошибки можно свести к минимуму периодической тарировкой приборов с помощью более совершенных и повышением точности метода определения исследуемых переменных. Случайные ошибки обусловлены влиянием большого количества факторов. Их появление неодинаково и случайно от измерения к измерению и не может быть предварительно учтено из-за их зависимости от изменения условий измерений и изменчивости самих измеряемых величин. Однако при достаточно большом количестве экспериментов суммарное значение случайных ошибок, изменяющихся примерно одинаково в положительную и отрицательную сторону, приближается к нулю. Случайные ошибки в подавляющем большинстве подчиняются нормальному закону распределения с *математическим ожиданием*, равным "0".

В практике исследований систематические и случайные ошибки близки друг к другу и совместно определяют ошибку измерений. При оценке точности измерений рекомендуется учитывать **суммарную ошибку**

$$\varepsilon_{\Sigma} = \delta + \varepsilon \frac{\sigma_{\Sigma}}{\sqrt{n}} \quad (1.6)$$

где σ_{Σ} — среднеквадратическое отклонение случайной величины при числе измерений n .

Для величин, определяемых косвенно — методом расчета по другим измеренным случайным величинам, оценка погрешностей осуществляется вычислением статистических оценок по соответствующим функциональным зависимостям.

Выборочные характеристики M_x, σ_x определяемые на основе ограниченного числа наблюдений, могут приближаться к истинным значениям характеристик генеральной совокупности M_x^0, σ_x^0 лишь с определенной точностью:

$$M_x^0 = M_x + \varepsilon, \sigma_x^0 = \sigma_x + \sigma_\varepsilon \quad (1.7)$$

Точность выборочного наблюдения (эксперимента) может задаваться в единицах измерения исследуемой величины, в единицах выборочного значения σ_x и в процентах исследуемой величины или характеристики.

Систематическая ошибка, будучи постоянной, при этом может не учитываться.

Вероятность того, что истинное значение характеристик генеральной совокупности находится в отмеченных пределах, равна

$$P(M_x - \varepsilon < M_x^0 < M_x + \varepsilon), P(\sigma_x - \sigma_\varepsilon < \sigma_x^0 < \sigma_x + \sigma_\varepsilon) \quad (1.8)$$

и называется **надежностью данной оценки**.

Так как математическое ожидание любой выборки само является случайной величиной, то полезно установить такой **интервал**, в котором с заданной степенью достоверности будет заключено значение оцениваемого параметра.

Интервал $M_x \pm \varepsilon$, в который в общем случае может быть произвольным $[a_i, b_i]$, называется доверительными границами, а соответствующая **вероятность** - доверительной вероятностью или, как часто говорят, надежностью. **Доверительную вероятность** для удобства обозначают как

$$(1 - \alpha) = P(a_i < M_x \leq b_i) \quad (1.9)$$

Соответственно α есть **вероятность** ошибки, которая на кривой распределения изображается в виде двух половинок $\alpha/2$.

Вероятность ошибки характеризует долю риска в оценке истинного значения оцениваемой величины и часто называется уровнем значимости. Для удобства величину доверительного интервала устанавливают в долях среднеквадратического отклонения $\pm z\sigma_x$. Тогда доверительную **вероятность** определяют, как **площадь**, ограниченную кривой нормального распределения на интервале $\pm z\sigma_x$. Используя формулу стандартного нормального распределения (9)

$$z = \frac{x - M_x}{\sigma_x} \quad (1.10)$$

при $M_x = 0$ и $\sigma_x = 1$

доверительную вероятность, согласно (8), записывают в таком виде:

$$P(\bar{x} - z \frac{\sigma_x}{\sqrt{n}} \leq M_x \leq \bar{x} + z \frac{\sigma_x}{\sqrt{n}}) \quad (1.11)$$

, где \bar{x} - оценка среднего значения генеральной совокупности

доверительный интервал для дисперсии

$$\left[\frac{n\sigma_x^2}{x^2\alpha/2} \leq D_x \leq \frac{n\sigma_x^2}{x^2(1-\alpha)/2} \right] \quad (1.12)$$

Определяют доверительный *интервал* в такой последовательности:

вычисляют *параметр* выборки $\bar{x} \approx M_x$, выбирают

доверительную *вероятность*, $(1 - \alpha)$ определяют соответствующее выбранному

значению $(1 - \alpha)$ число из таблицы табулированных значений стандартного нормального

распределения; вычисляют *доверительный интервал* $a \leq M_x < b$.

С увеличением количества замеров *достоверность* эксперимента возрастает, а доверительный *интервал* уменьшается. *Таблица* используется в том случае, когда о дисперсии исследуемой величины нельзя составить определенного мнения. Если же на основании априорных сведений или предварительных опытов (среднее квадратичное отклонение), известно, то по формуле случайной выборочной ошибки, равной половине длины доверительного интервала

$$\varepsilon_{0,5} = z \frac{\sigma_x}{\sqrt{n}} \quad (1.13)$$

определяют необходимое число замеров, гарантирующее требуемую *надежность*

$$n = \frac{z^2 \sigma_x^2}{\varepsilon^2} \quad (1.14)$$

Точность и *надежность* оценки выборочных характеристик не следует смешивать с δ , которую часто вычисляют по такой формуле:

$$\Delta u = \frac{v}{\sqrt{n}} \delta \quad (1.15)$$

где V — *коэффициент вариации* выборочного наблюдения, % (в случае оценки точности для

сельскохозяйственных машин и процессов считается достаточным, если $\Delta u \leq 3 \dots 5\%$).

Кроме установления доверительных интервалов, задачи оценки случайных величин включают *анализ* законов распределения изучаемых величин, проверку принадлежности двух выборок к одной генеральной совокупности, сравнение средних дисперсий для различных выборок и др.

Рабочим инструментом статистического анализа при решении отмеченных задач оценки являются *статистические гипотезы*. **Статистическими гипотезами** именуется суждения, применяемые при различных видах анализа, касающихся, по существу, выяснения свойств некоторой генеральной совокупности случайных величин. *Гипотеза* в статистике трактуется как предположение о распределении случайных величин.

Гипотеза, отклонения от которой приписываются данному случаю, называется нулевой и обозначается H_0 .

Альтернативная или противоположная *гипотеза* называется конкурирующей и обозначается H_1 . Гипотезы проверяют при помощи специально подобранной случайной величины, распределение которой известно или может быть установлено при малом объеме выборки.

Остановимся на рассмотрении требований (или правил), которые должны выполняться при проведении специальных статистических наблюдений. Эти правила сформулировал еще в XIX веке известный бельгийский статистик А. Кетле.

Первое правило: Программа статистических наблюдений должна включать только те вопросы, на которые необходимо получить ответы, исходя из целей статистических наблюдений. Исходя из этого правила, из наблюдений надо исключить все показатели, которые предполагается получить на всякий случай.

Второе правило : в программу наблюдений не стоит включать вопросы, на которые не удастся получить ответы удовлетворительного качества.

Третье правило : в программу наблюдений не должны включаться вопросы, которые могут вызвать недоверие обследуемых субъектов (единиц совокупности) относительно целей проведения статистического исследования. При организации наблюдения всегда следует помнить о воздействии, которое оказывает наблюдение на изучаемый *объект* (единиц совокупности).

Выполнение этих правил достигается путем рассмотрения (еще до наблюдения) всех стадий статистического исследования - от целей и методов сбора, до способа сводки и группировки, а так же анализа. Только в этом случае можно быть уверенным, что *программа* наблюдений определена правильно. Иначе неизбежны излишества в программе наблюдений, или отсутствия в ней некоторых вопросов, без ответа на которые цели исследования не могут быть выполнены.

Корреляционная связь и её статистическое изучение.

Все явления объективного мира взаимосвязаны и взаимообусловлены. Связи между явлениями и признаками отличаются разнообразием. Основное внимание исследователей, как правило, приковано к причинно-следственным связям. При изучении таких связей одни признаки (процессы, явления) выступают в качестве факторов (независимых), обуславливающих изменение других признаков (зависимых, результативных).

Зависимость между признаками – факторами (факторными признаками) и признаками, которые являются результатом влияния этих факторов (результативными признаками) может быть функциональной или корреляционной.

Функциональные связи характеризуются полным соответствием между изменением факторного признака и изменением результативного признака, т.е. каждому значению признака-фактора соответствует строго определенное *значение* результативного признака.

Например, компонентная связь и взаимосвязь индексов.

В корреляционной связи между изменением факторного и результативного признака нет такого полного соответствия, воздействие факторов проявляется лишь в среднем при массовом наблюдении фактических данных. При изучении причинно-следственных связей решающее *слово* должно принадлежать теории (сущности) изучаемого явления. Теоретический анализ должен показать о наличии или возможности связи между данными признаками, какие факторы влияют на формирование и изменение данного результативного признака.

При исследовании корреляционных зависимостей решается широкий круг вопросов:

- предварительный анализ свойств изучаемой совокупности;
- установление факта наличия связи, определение ее направления и формы;
- измерение степени тесноты связи между признаками;
- нахождение аналитического (математического) выражения связи или построение регрессионной модели;
- оценка адекватности модели, ее интерпретации и практическое использование.

При проведении корреляционно-регрессионного анализа, должны соблюдаться определенные требования:

- необходимо обеспечить качественную однородность изучаемой совокупности. Например, при изучении зависимостей между технико-экономическими показателями, необходимо отбирать не только предприятия одной отрасли, но также предприятия, выпускающие однородную и однотипную продукцию, имеющие одинаковый характер техпроцессов и тип используемого оборудования;
- однородность изучаемой совокупности следует дополнить количественными характеристиками совокупности (с применением показателей вариации, проверки гипотезы о принадлежности "выделяющихся" или аномальных значений признака исследуемой совокупности);

- учет требований закона больших чисел: исходные данные должны быть массивными, представительными;
- включаемые в исследование признаки-факторы должны быть основными (оказывать, решающее влияние на уровень результативного признака) и быть независимыми друг от друга (не должны дублировать друг друга);
- при практическом применении результатов изучения корреляционной связи следует иметь в виду, что все основные положения теории корреляции и регрессии разрабатывались из предложения о нормальном характере распределения рассматриваемых признаков.

Поэтому целесообразно проверить правомерность такого подхода (изучить формы распределения признаков).

Изложенное указывает на то, что корреляционно-регрессионный *анализ* применяется преимущественно только к признакам, которые имеют количественное *выражение*.

При статистическом изучении корреляционной связи между признаками исходным материалом являются данные об индивидуальных значениях этих признаков в изучаемой статистической совокупности.

Статистическая наука в настоящее время располагает большим набором приемов (методов) выявления корреляционной связи. Одни приемы можно отнести к элементарным (простейшим), другие предусматривают использование специального сложного математического аппарата.

К *элементарным приемам (методам)* выявления наличия корреляционной связи относятся: параллельное сопоставление рядов значений факторного и результативного признаков, графическое изображение фактических данных с помощью поля корреляции, построение групповой и корреляционной таблиц, факторные (аналитические) группировки и исчисление групповых средних.

К *сложным методам изучения взаимосвязей* относятся балансовые таблицы, дисперсионный *анализ*, методы теории корреляции и регрессии, методы многомерного анализа, методы распознавания образов, метод главных компонентов и др.

При отсутствии ярко выраженной причинной связи между факторным и результативным признаками наличие и характер связи можно установить при помощи метода параллельных рядов: в одной таблице приводятся упорядоченные значения факторного признака, который обычно обозначается символом X , и соответствующие им значения результативного признака, который обычно обозначается символом Y .

Наличие и характер связи определяется по степени согласованности вариации данных рядов.

В тех случаях, когда возрастание величины факторного признака влечет за собой возрастание величины результативного признака, говорят о возможном наличии *прямой* корреляционной связи. Если же с увеличением факторного признака величина результативного признака имеет тенденцию к уменьшению, то можно предполагать обратную *связь* между этими признаками.

Краткие итоги

Рассмотрены понятия эксперимента и математической модели. Конечной целью любой обработки экспериментальных данных является выдвижение гипотез о классе и структуре математической модели исследуемого явления.

Математическая модель – это приближённое описание какого-либо класса явлений внешнего мира, выраженное с помощью математической символики. Процесс математического моделирования, то есть изучения явления с помощью М. м., можно подразделить на 4 этапа.

Дано *представление* о статистическом анализе. Рабочим инструментом статистического анализа при решении отмеченных задач оценки являются статистические гипотезы. Статистическими гипотезами именуется суждения, применяемые при различных видах анализа, касающихся, по существу, выяснения свойств некоторой генеральной совокупности случайных величин.

Рассмотрены основные три правила, которые должны выполняться при проведении специальных статистических наблюдений.

Дано понятие корреляционной связи и её статистическое изучение.

Вопросы для самопроверки

1. Что такое экспериментальные исследования и с какой целью они проводятся?
2. Дайте определение математической модели.
3. Охарактеризуйте этапы математического моделирования.
4. Что является рабочим инструментом статистического анализа?
5. Дайте определение статистической гипотезы?
6. Какие ошибки встречаются при выборочном наблюдении?
7. Перечислите основные правила, которые должны выполняться при проведении статистических наблюдений.
8. Что такое корреляционная связь?
9. Охарактеризуйте методы выявления корреляционной связи.