**Математическая статистика. Начало**

Для изучения темы желательно знать [**азы теории вероятности**](http://www.mathprofi.ru/teorija_verojatnostei.html), в частности, [**случайные величины**](http://www.mathprofi.ru/sluchainaya_velichina.html) – многие понятия и формулы будут очень и очень схожи.

***Что такое математическая статистика*?**

Её часто называют то наукой, то разделом математики. И это правда :) Математическая статистика, буду краток, ***изучает методы сбора и обработки статистической информации для получения научных и практических выводов.***

Статистическая – это та, которую можно выразить числами.

 Эта информация появляется в результате исследования массовых *(обычно)* явлений, которые носят случайный характер.

Причём, информация может носить как количественный характер (например, размеры чего-либо), так и качественную природу Немедленный пример. Что главное орудие физика? Секундомер:

Пример 1

Студент Константин выполняет лабораторную работу по определению коэффициента вязкости жидкости методом Стокса.

…

Экспериментальная часть этой работы состоит в том, что в высокий цилиндрический сосуд с жидкостью сбрасывается достаточно маленький и тяжёлый шарик, после чего замеряется время его погружения.

Время погружения шарика зависит от множества случайных факторов: прямоты рук экспериментатора, погрешности измерения времени, хаотичного движения молекул жидкости и т.д.,. Поэтому эксперимент  целесообразно провести 5-10 раз (как оно обычно и требуется).

Предположим, что в результате 5 опытов получены следующие результаты (в секундах):


Что произошло? Студент Костя собрал ***первичные*** *(ещё не обработанные)* статистические данные. Они *эмпирические* (взяты непосредственно из опыта), носят случайный характер И массовый. Ну а как нет? Все однокурсники только и занимаются тем, что бросают в сосуды шарики, да и мало ли на планете похожих шариков, которые тонут в похожей жидкости.

Ну а мы потихоньку погружаемся в терминологию:

- полученные экспериментальные значения называются ***вариантами***, а их совокупность – ***вариационным рядом.***

Почему так? Потому что полученные значения *варьируются* под воздействием случайных факторов.

.

Далее.

Далее Константин должен обработать полученные данные.

Во-первых, посмотреть, а нет ли среди полученных значений *варианты*, которая сильно отличается от всех остальных? Наличие такого значения сигнализирует о том, что соответствующий опыт проведён неудачно и его следует исключить из рассмотрения.

Нет, все значения достаточно близки друг к другу, и теперь напрашивается вычислить  среднюю величину – разделить сумму значений на их  количество:
 секунды.

Это значение называют ***простой средней*** или, как многие знают, ***средним арифметическим***. Его стандартно обозначают с чёрточкой наверху.

***Справка на всякий случай:*** *математический значок означает суммирование, а переменная  играет роль «счётчика»; в данном случае  изменяется от 1 до 5*.

Всё.

Статические данные обработаны, осталось сделать выводы. А именно, с помощью значения  вычислить коэффициент вязкости жидкости …

 Пример 2

Студенческая группа сдала коллоквиум по [**матанализу**](http://www.mathprofi.ru/predely_po_koshi.html) со следующими результатами:


Требуется определить среднюю успеваемость группы

Как нетрудно понять, роль ***вариант***  здесь играют полученные оценки, а   – это соответствующие ***частоты*** – количество студентов, которые получили ту или иную оценку. Подсчитаем общую численность группы:
 человек и, привыкаем к терминам, исследуемое множество называют ***статистической совокупностью***, а количество его элементов – ***объёмом***совокупности.

Теперь обратим внимание на следующую вещь: двоечников и отличников у нас мало, а нормальных студентов :) много. И возникает вопрос: как вычислить «справедливую» среднюю оценку по всей совокупности?

Решение напрашивается – с помощью так называемой ***средневзвешенной средней:***
 – средняя успеваемость по группе. …да, суровые у меня сегодня примеры :) Давайте проанализируем их **принципиальные отличия**:

1) В первом примере проводится статистическое исследование *количественной* величины (времени), а во втором «оцифровывается» и анализируется *качественный* признак (успеваемость).

2) В первом случае исследуемая величина *непрерывна*, и, строго говоря, все полученные значения **различны** (отличаются хоть какими-то миллисекундами). Во втором случае варианты *дискретны*, т.е. представляют собой отдельно взятые изолированные значения. Следует заметить, что они не обязаны быть целыми, так, например, можно ввести в рассмотрение оценки 2,5; 3,5 и 4,5. И у дискретной величины, как правило, есть *неоднократно* встречающиеся (одинаковые) варианты, так, например, «пятёрка» встретилась 3 раза.

3) В первом примере речь идёт о ***выборке*** значений. Что это значит? Это значит, что шарик можно сбрасывать в воду гораздо бОльшее и теоретически вообще бесконечное количество раз. Таким образом, проведённые 5 опытов есть, по сути, *выборка*, которую называют ***выборочной совокупностью***. При этом соответствующее среднее значение принято называть ***выборочной средней***.

Второй пример отличен тем, что в нём исследуется ВСЯ совокупность, и поэтому её называют И я обязательно приму соответствующие меры!

, а соответствующее среднее значение – ***генеральной средней***. Но такая ситуация редкость. Редко когда удаётся исследовать всю совокупность.

.

**И в заключение вводного урока систематизируем самое важное:**

**Математическая статистика** – это наука, изучающая методы сбора и обработки статистической информации для получения научных и практических выводов.

**Основным методом** матстатистики является **выборочный метод**, его суть состоит в исследовании представительной *выборочной совокупности* – для достоверной характеристики совокупности генеральной. Данный метод экономит временные, трудовые и материальные затраты, поскольку исследование всей совокупности зачастую затруднено или невозможно.