Доклад

по предмету «Математические методы

в психологии» на тему:

«Переменные и их измерение.

Символы, данные и операции.

Обозначение сигма (Σ)»

подготовила студентка

2003 г.

План

1. Переменные и их измерение.
2. Символы, данные и операции.
3. Обозначение сигма (Σ)

**Переменные и их измерение**

Переменными являются характеристики людей или вещей, например, вес, возраст, время реакции, беглость чтения, количество детей, число студентов. Интуиция и опыт подсказывают нам, что некоторые из этих переменных *непрерывны* (то есть, измерения их могут дать любое значение внутри некоторой области), таковы вес, возраст и время реакции. Мы твердо знаем, что некоторые переменные *дискретны* (то есть их измерения могут давать только отдельные значения), например количество детей. Наиболее известны те дискретные переменные, которые измеряются в результате счета. "Количество детей" может быть 0, 1, 2, 3 ... Естественно, эта переменная не может принимать промежуточные значения, скажем 1,75.

С другой стороны, мы уверены, что если бы только у нас были соответствующие приборы, средства и время, то можно было бы измерять непрерывные переменные с желаемой точностью. Измеряя время в состязаниях по бегу, мы предпочитаем останавливаться на определении десятых секунды. Но хотя сообщается, что расстояние 90 м было преодолено за 10,4 сек, более точные хронометры могли бы показать, что рекордное время равно 10,416 сек. Но даже это время не точно; просто оно верно до тысячных долей секунды. *Настоящего*, или *точного*, измерения переменной никогда нельзя достигнуть, так как измерение всегда должно где-то оборвать *точное значение* (под точным значением, или меткой, не надо понимать "истинную" или совершенно устойчивую метку, которой не бывает. Реальная метка может быть нестабильной во времени). В силу этого точное значение переменной - это *косвенное значение*. Оно является результатом процесса измерения. Мы не рассчитываем на совпадение косвенного и фактического значений переменной, но первое задает пределы для последнего. Например, если рост человека, измеренный с точностью до сантиметра, составляет 157 см, то его действительный рост в это время и в этих условиях находится между 156,5 и 157,5 см.

Измерение любой непрерывной переменной должно сопровождаться определением точности процесса измерения. Скорости хронометрируются с точностью до десятой доли секунды; рост может быть измерен с точностью до сантиметра; возраст - с точностью до дня. *Чувствительность* процесса измерения задается минимальной единицей цифровой шкалы, которая фиксируется. Таким образом, чувствительностью в трех вышеприведенных примерах были соответственно десятые доли секунды, сантиметры и дни.

Часто мы хотим задать те границы любого найденного значения, в которых находится точное значение. Например, каковы минимальные и максимальные действительные значения рост; которые соответствуют установленному росту 147 см, если измерение осуществляется с точностью до сантиметра? *Пределы, для точного значения в окрестности любого найденного значения устанавливаются путем прибавления и вычитания половины чувствительности измерительного процесса от найденного значения.* Таким образом, человек с установленным ростом 147 см имеет действительный рост в интервале между 147 см – (1 см/2) = 146,5 см и 147 см + (1 см/2) = 147,5 см.

Следующие примеры должны внести ясность в эту процедуру (определение пределов для точного значения в такой форме требует многих оговорок и может рассматриваться лишь как сугубо ориентировочное).

|  |  |  |  |
| --- | --- | --- | --- |
| Переменная | Чувствительность измерения | Результат измерения | Пределы точного знаения |
| Вес | кг | 59 кг | 58,5-59,5 кг |
| Возраст | год | 25 лет | 24 года 6 мес – 25 лет 6 мес |
| Время реакции | 1/100 сек | 0,53 сек | 0,525-0,535 сек |
| Время пробга | 2/10 сек | 5,6 сек | 5,5-5,7 сек |

Оценки в педагогических и психологических тестах часто получаются в результате подсчета числа правильных ответов, которые дает испытуемый. Джон ответил правильно на 45 из 90 вопросов по разделу речи школьного теста на определение способностей. Поэтому соответствующая цифровая метка для переменной "речевые способности школьника" составляет 45. Поскольку мы вообще рассматриваем переменные, лежащие в основе педагогических и психологических тестов, чаще всего как непрерывные, чувствительностью этих измеряющих средств является единица или одна точка шкалы (если тесты составлены из отдельных пунктов). Таким образом, точная метка Джона в тесте лежит между метками 44,5 и 45,5. Если сначала это покажется вам странным, вспомните, что измеряется непрерывная переменная "речевые способности школьника", а не дискретная "число точных ответов".

Практический вопрос, который чуть позже возникнет, касается отношения к результатам измерений при выполнении вычислений. Если 10 измерений *IQ* имеют одинаковое значение 105, а пределы точного значения образуют диапазон от 104,5 до 105,5, то 10 меток обычно считаются равномерно распределенными в интервале, ограниченном пределами точного значения. Если кому-нибудь для какой-либо цели нужно узнать, сколько меток превышает 105,2, надо учесть 3 метки из 10 в интервале от 104,5 до 105,5 (см. рис.). Это удобное допущение принимается при определении средних величин и расчете меток, ниже которых лежат заданные проценты испытуемых.

Приходится, к сожалению, признать, что понятие пределов точных значений дискретных измерении является одним из рабочих орудий статистика. Хотя и бессмысленно говорить, что точное число студентов, которых обучает данный преподаватель, лежит между 33,5 и 34,5 - очевидно, что их 34, - это все же иногда делается при выполнении расчетов.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 105,0  104,5  105,2  105,5 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |  |
|  | |  |  |  |  | |  | |  |  | |

**Символы, данные и операции**

Если мы хотим указать на множество чисел вообще, не записывая конкретно каждое из них, мы выбираем любую произвольную величину, например Х*i* (читается "X с индексом *i*'"). X заменяет число; *i*, называемое подстрочным *индексом*, указывает, какое число имеет *i*-й номер. Когда индекс зафиксирован, скажем, на значении 4, то Х4 означает определенное число: четвертый член некоторой группы. X1 обозначает одно число, Х2- другое, причем 1 и 2 являются только обозначениями или наименованиями: мы не можем заключить на основе индекса, что больше – Х1 или Х2. Мы можем обозначить 4,3; 2,1; 6,7 и 3,5 через Х1, Х2, Х3 и X4. Конечно, мы могли бы обозначить 4,3 через X2 вместо X1, как мы сделали. X1- просто первое число в нашей последовательности из n чисел, а Хn - последнее.

Если мы имеем группу из n чисел (числом n может быть 2, 3, 100 или любое другое число), то можем обозначить ее элементы с помощью символов X1, Х2, ..., Хп. Вообще *i*-е измерение (X) есть X*i*, где *i* может быть каким-нибудь одним из индексов 1, 2, ..., n.

Данные могли бы быть расположены в таблице со строками и столбцами. Каждый элемент такого расположения можно описать, если мы знаем группу (столбец), в которой он находится, и его положение в этой группе (строку):

|  |  |  |  |
| --- | --- | --- | --- |
| Порядок внутри группы | Номер группы | | |
| 1 | 2 | 3 |
| Первый (1) | Х11 = 4,0 | Х12 = 6,5 | Х13 = 4,4 |
| Второй (2) | Х21 = 2,3 | Х22 = 2,1 | Х23 = 5,3 |

Когда мы пишем Х12, то имеем в виду первый элемент второй группы, 6,5. AM заменяет второе число в третьей группе, 5,3. Когда же мы пишем Х*ij*, то мы можем обозначать каждое из этих 6 чисел, придавая *i* значение 1 или 2, а *j*- 1, 2 или 3.

Допустим, вы собирались провести эксперимент, в котором 12 человек читали бы одну брошюру, а 10 человек - другую. Вполне возможно, что вам захочется говорить о числах, которые получатся в результате этого эксперимента, раньше, чем они будут получены. Вместо того чтобы сказать: "Я собираюсь сравнить третий номер в первой группе со вторым номером. во второй группе", вы можете сказать: "Я думаю сравнить Х31 с Х22". Символы должны стать полезным и стенографически экономным средством.

Данные можно классифицировать применительно к любому количеству характеристик.

**Обозначение сигма (Σ)**

Анализ большинства данных включает, между прочим, сложение, вычитание, умножение и деление чисел. Поскольку мы хотим поговорить о выполнении этих операций над группой чисел вообще, произведем операции на символах вместо чисел.

Последовательность Х1, Х2, ..., Хп представляет собой группу из n чисел, каждое число которой можно записать как Xi. Х1 + Х2 заменяет сумму первого и второго чисел. Порядок индексов обычно совершенно произволен. С тем же успехом можно было бы использовать Х2 + Х1 Х1 + Х2 + Х10 представляет собой сумму первого, второго и десятого номеров.

Часто мы хотим сложить все числа группы. Если в группе имеется 5 чисел, то n = 5, а сумма всех чисел равна Х1 + Х2 + … + Х5 Х1 + Х2 + … + Xn обозначает сумму всех n чисел в группе, когда точное значение n не сговорено.

Сокращение записи для Х1 + Х2 + … + Xn, которое часто употребляется, выглядит так: Х*i*



Х*i* обозначает Х1 + Х2 + … + Xn



Х*i =* Х1 + Х2 + X3 Х*i =* Х3 + Х4 + X5



Σ - это греческая прописная буква "сигма". Х*i* читается как "сумма Х*i* когда *i* пробегает значения от 1 до 5". Х*i* читается как "сумма Х*i* когда *i* пробегает значения от 1 до n ".



Общепризнанно, что краткое обозначение Σ является экономным. Статистики извлекают из этою большую пользу.

Сложение чисел, умноженных, например, на 6 или возведенных в квадрат (это значит умноженных на самих себя), осуществляется, как обычно. Допустим, мы хотим умножить каждое из n чисел на 2 и сложить результаты. Искомая сумма есть

2X1 + 2X2+...+2Xn.

Но вы наверняка заметите, что эта сумма - то же самое, что и

2(X1 + X2+...+Xn).

Используя Σ-обозначение, мы можем заменить (X1 + 2X2+...+2Xn) на Х*i* Результат можно записать так:



2X1 + 2X2+...+2Xn = 2Х*i = 2*Х*i*



Этот результат возник не вследствие какого-либо магического свойства числа 2: с числами 4, 60 или 131,4 результат будет тот же. В самом деле, если с представляет собой какое-либо постоянное число (то есть число, которое не зависит от *i*), то

*с*X1 + *с*X2+...+*с*Xn = *с*Х*i = с*Х*i*  (Правило 1)



Если постоянное число (константу) *с* прибавить к каждому из n чисел, то получим

X1 + *с,* X2+ *с, …,* Xn + *с*

Сумма этих значений

(X1 + *с*) + (X2+ *с*) + … + (Xn +*с*) = ( X*i* +*с*)



При сложении мы всегда можем перегруппировать числа в любом порядке до того, как складывать

( X*i* +*с*) = (X1 + X2+...+Xn ) + (*с* + *с + … + с*)



Первая сумма в круглых скобках справа дает Х*i*



Какова же вторая сумма в круглых скобках? Сколько *с* сложено? Ответ: n. Поэтому вторая сумма равна *nс*. Следовательно,

( X*i* +*с*) = Х*i + с* = Х*i + nс*  (Правило 2)



Другое важное выражение - сумма квадратов n чисел

(X*1* X*1*) + (X*2* X*2*) + ... + (X*n* X*n* ) = + + … + ,



которое символически изображается как Х



Аналогично

+ + … + = Х



хотя в элементарной статистике это выражение встречается редко.

Заметим, что Х*i* символически изображает единственное число: число, которое получается в результате сложения n чисел.



Х*i* может быть 10, 13 или 1300. сХ*i* это произведение двух чисел *с* и Х*i* . (Х*i*) (Х*i*) является произведением числа (некоторой суммы), умноженного на самого себя. Мы также запишем это следующим образом:



(Х*i*) (Х*i*) = (Х*i*)2



Если Х1 = 3, Х2 = 6, а Х3 = 1, то Х*i =* 10, а (Х*i*)2 = 100.



Обычным в статистическом анализе является выражение

(X*i* +*с*)2 = (X1 + с)2 + (X2+ с)2 + ... + (Xn +с)2



(X*i* +*с*)2 , равное (X*i* +*с*) (X*i* +*с*), иначе можно записать так:

X*i* + *с*



Действительно, тогда

(X*i* +*с*)2 = (Х + 2сХi  +с2)



Выражение в скобках можно записать n раз следующим образом:

Х + 2сХ1  +с2



Х + 2сХ2  +с2



… … …

… … …

Х + 2сХn  +с2



Чему равна сумма первого столбца данного выражения? Она равна Х+ Х + … + Х = Х. Какова сумма второго столбца? Она составляет



2сХ1 + 2сХ2 + … + 2сХn = 2с (Х1 + Х2 + … + Хn),

что более кратко можно записать как 2с Хi . Какова сумма третьего столбца? Она представляет собой *с2 + с2 + ... + с2  = nc2 .* Складывая суммы этих трех столбцов, имеем



(X*i* +*с*)2  = Х + 2сХi.+ *nc2.* (Правило 3)



Хотя такие действия правильны, в них нет необходимости. Вместо этого можно "распределить" знак суммирования перед каждым членом и получить непосредственно тот же результат:

(X*i* +*с*)2 = (Х + 2сХi  +с2) = Х + 2сХi + с2 =



= Х + 2сХi.+ *nc2.*

