ГОСУДАРСТВЕННЫЙ КОМИТЕТ ПО НАУКЕ И ТЕХНИКЕ АЗЕРБАЙДЖАНСКОЙ РЕСПУБЛИКИ

БАКИНСКИЙ НАУЧНО-УЧЕБНЫЙ ЦЕНТР

РЕФЕРАТ

АСПИРАНТА КАФЕДРЫ ДЕТСКОЙ ХИРУРГИИ

АМУ имени Н. НАРИМАНОВА

МУХТАРОВА ЭМИЛЯ ГАСАН оглы

НА ТЕМУ:

КОРРЕЛЯЦИОННЫЕ МОМЕНТЫ. КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ

БАКУ - 1999

ВВЕДЕНИЕ

***Теория вероятности* есть математическая наука, изучающая закономерности в случайных явлениях.**

**Что же понимается под случайными явлениями?**

**При научном исследовании физических и технических задач, часто приходится встречаться с явлениями особого типа, которые принято называть случайными. *Случайное явление* - это такое явление, которое при неоднократном воспроизведении одного и того же опыта протекает несколько по-иному.**

**Приведем пример случайного явления.**

**Одно и то же тело несколько раз взвешивается на аналитических весах: результаты повторных взвешиваний несколько отличаются друг от друга. Эти различия обусловливаются влиянием различных второстепенных факторов, сопровождающих операцию взвешивания, таких как случайные вибрации аппаратуры, ошибки отсчета показаний прибора и т.д.**

**Очевидно, что в природе нет ни одного физического явления, в котором не присутствовали бы в той или иной мере элементы случайности. Как бы точно и подробно ни были фиксированы условия опыта, невозможно достигнуть того, чтобы при повторении опыта результаты полностью и в точности совпадали.**

**Случайности неизбежно сопутствуют любому закономерному явлению. Тем не менее, в ряде практических задач этими случайными элементами можно пренебречь, рассматривая вместо реального явления его упрощенную схему, т.е. *модель*, и предполагая, что в данных условиях опыта явление протекает вполне определенным образом. При этом из бесчисленного множества факторов, влияющих на данное явление, выделяют самые главные, основные, решающие. Влиянием остальных, второстепенных факторов просто пренебрегают. Изучая закономерности в рамках некоторой теории, основные факторы, влияющие на то или иное явление, входят в понятия или определения, которыми оперирует рассматриваемая теория.**

**Как и всякая наука, развивающая общую теорию какого-либо круга явлений, теория вероятностей также содержит ряд основных понятий, на которых она базируется. Естественно, что не все основные понятия могут быть строго определены, так как определить понятие - это значит свести его к другим, более известным. Этот процесс должен быть конечным и заканчиваться на первичных понятиях, которые только поясняются.**

**Одним из первых понятий в теории вероятности вводится понятие события.**

**Под *событием* понимается всякий факт, который в результате опыта может произойти или не произойти.**

**Приведем примеры событий.**

**А - рождение мальчика или девочки;**

**В - избрание того или иного дебюта в шахматной игре;**

**С - принадлежность к тому или иному зодиакальному знаку.**

**Рассматривая вышеперечисленные события, мы видим, что каждое из них обладает какой-то степенью возможности: одна большей, другие - меньшей. Для того чтобы количественно сравнивать между собой события по степени их возможности, очевидно, нужно с каждым событием связать определенное число, которое тем больше, чем более возможно событие. Такое число называют вероятностью события. Таким образом, вероятность события есть численная характеристика степени объективной возможности события.**

**За единицу вероятности принимают вероятность достоверного события, равную 1, а диапазон изменения вероятностей любых событий - числа от 0 до 1.**

**Вероятность обычно обозначают буквой Р.**

**Рассмотрим на примере извечной проблемы шекспировского Гамлета "быть или не быть?" как можно определить вероятность события.**

**Вполне очевидно, что человек, предмет и всякое иное явление может находиться в одном из двух и не более состояний: наличия ("быть") и отсутствия ("не быть"). Т.е., возможных событий две, а произойти может только одно. Это означает, что вероятность, например бытия, равна 1/2.**

**Помимо понятия события и вероятности, одним из основных понятий теории вероятностей является понятие случайной величины.**

***Случайной величиной* называется величина, которая в результате опыта может принять то или иное значение, причем неизвестно заранее какое именно.**

**Случайные величины, принимающие только отдельные друг от друга значения, которые можно заранее перечислить, называются *прерывными или дискретными случайными величинами.***

**Например:**

1. **Число выживших и умерших больных.**
2. **Общее количество детей из поступивших за ночь в больницу больных.**

Случайные величины, возможные значения которых непрерывно заполняют некоторый промежуток, называют *непрерывными случайными величинами.*

**Например, ошибка взвешивания на аналитических весах.**

**Отметим, что современная теория вероятности преимущественно оперирует случайными величинами, а не событиями, на которые в основном опиралась "классическая" теория вероятностей.**

КОРРЕЛЯЦИОННЫЕ МОМЕНТЫ. КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ.

***Корреляционные моменты, коэффициент корреляции* - это числовые характеристики, тесно связанные во введенным выше понятием случайной величины, а точнее с системой случайных величин. Поэтому для введения и определения их значения и роли необходимо пояснить понятие системы случайных величин и некоторые свойства присущие им.**

**Два или более случайные величины, описывающих некоторое явление называют *системой или комплексом случайных величин.***

Систему нескольких случайных величин X, Y, Z, …, W принято обозначать через (X, Y, Z, …, W).

**Например, точка на плоскости описывается не одной координатой, а двумя, а в пространстве - даже тремя.**

Свойства системы нескольких случайных величин не исчерпываются свойствами отдельных случайных величин, входящих в систему, а включают также взаимные связи (зависимости) между случайными величинами. Поэтому при изучении системы случайных величин следует обращать внимание на характер и степень зависимости. Эта зависимость может быть более или менее ярко выраженной, более или менее тесной. А в других случаях случайные величины оказаться практически независимыми.

**Случайная величина Y называется *независимой* от случайной величины Х, если закон распределения случайной величины Y не зависит от того какое значение приняла величина Х.**

**Следует отметить, что зависимость и независимость случайных величин есть всегда явление взаимное: если Y не зависит от Х, то и величина Х не зависит от Y. Учитывая это, можно привести следующее определение независимости случайных величин.**

**Случайные величины Х и Y называются независимыми, если закон распределения каждой из них не зависит от того, какое значение приняла другая. В противном случае величины Х и Y называются *зависимыми.***

***Законом распределения* случайной величины называется всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями.**

**Понятие "зависимости" случайных величин, которым пользуются в теории вероятностей, несколько отличается от обычного понятия "зависимости" величин, которым пользуются в математике. Так, математик под "зависимостью" подразумевает только один тип зависимости - полную, жесткую, так называемую функциональную зависимость. Две величины Х и Y называются функционально зависимыми, если, зная значение одного из них, можно точно определить значение другой.**

**В теории вероятностей встречаются несколько с иным типом зависимости - *вероятностной зависимостью*. Если величина Y связана с величиной Х вероятностной зависимостью, то, зная значение Х, нельзя точно указать значение Y, а можно указать её закон распределения, зависящий от того, какое значение приняла величина Х.**

**Вероятностная зависимость может быть более или менее тесной; по мере увеличения тесноты вероятностной зависимости она все более приближается к функциональной. Т.о., функциональную зависимость можно рассматривать как крайний, предельный случай наиболее тесной вероятностной зависимости. Другой крайний случай - полная независимость случайных величин. Между этими двумя крайними случаями лежат все градации вероятностной зависимости - от самой сильной до самой слабой.**

Вероятностная зависимость между случайными величинами часто встречается на практике. Если случайные величины Х и Y находятся в вероятностной зависимости, то это не означает, что с изменением величины Х величина Y изменяется вполне определенным образом; это лишь означает, что с изменением величины Х величина Y

имеет тенденцию также изменяться (возрастать или убывать при возрастании Х). Эта тенденция соблюдается лишь в общих чертах, а в каждом отдельном случае возможны отступления от неё.

Примеры вероятностной зависимости.

Выберем наугад одного больного с перитонитом . случайная величина Т - время от начала заболевания, случайная величина О - уровень гомеостатических нарушений. Между этими величинами имеется явная зависимость, так как величина Т является одной из наиболее главных причин, определяющих величину О.

В то же время между случайной величиной Т и случайной величиной М, отражающей летальность при данной патологии, имеется более слабая вероятностная зависимость, так как случайная величина хоть и влияет на случайную величину О, однако не является главной определяющей.

Тем более, если рассматривать величину Т и величину В (возраст хирурга), то данные величины практически независимы.

До сих пор мы обсуждали свойства систем случайных величин, давая только словесное разъяснение. Однако существуют числовые характеристики, посредством которых исследуются свойства как отдельных случайных величин, так и системы случайных величин.

Одной из важнейших характеристик случайной величины нормального распределения является *математическое ожидание.*

Рассмотрим дискретную случайную величину Х, имеющую возможные значения Х1, Х2, ... , Хn с вероятностями р1, р2, ... , рn. нам требуется охарактеризовать каким-то числом положение значений случайной величины на оси абсцисс с учетом того, что эти значения имеют различные значения. Для этой цели обычно пользуются так называемым "средним взвешенным" из значений Хi, причем каждое значение Хi при осреднении должно учитываться с "весом", пропорциональным вероятности этого значения. Таким образом, если обозначить "среднее взвешенное" через М[X] или mx, получим

или, учитывая, что , то

 (1).

*Математическим ожиданием* случайной величины называется сумма произведений всех возможных значений случайной величины на вероятности этих значений.

Для большей наглядности рассмотрим одну механическую интерпретацию введенного понятия. Пусть на оси абсцисс расположены точки с абсциссами х1, х2, …, хn, в которых сосредоточены соответственно массы р1, р2, … , рn, причем . Тогда математическое ожидание есть не что иное, как абсцисса центра тяжести данной системы материальных точек.

Формула (1) для математического ожидания соответствует случаю дискретной случайной величины. Для непрерывной величины Х математическое ожидание, естественно, выражается не суммой, а интегралом:

(2),

где - плотность распределения величины Х.

Формула (2) получается из формулы (1), если в ней заменить отдельные значения Хi непрерывно изменяющимся параметром Х, соответствующие вероятности рi элементом вероятности *f(x)dx*, конечную сумму - интегралом.

В механической интерпретации математическое ожидание непрерывной случайной величины сохраняет тот же смысл - абсциссы центра тяжести в случае, когда масса распределения по оси абсцисс непрерывна с плотностью *f(x).*

Следует отметить, что математическое ожидание существует не для всех случайных величин, что, однако, по мнению некоторых ученых, не представляет для практики существенного интереса.

Помимо математического ожидания важное значение имеют также другие числовые случайной величины - *моменты*.

Понятие момента широко применяется в механике для описания распределения масс ( статистические моменты, моменты инерции и т.д.). Совершенно теми же приемами пользуются в теории вероятностей для описания основных свойств распределения случайной величины. Чаще всего применяются на практике моменты двух видов: начальные и центральные.

*Начальным моментом* s-го порядка прерывной случайной величины Х называется сумма вида

Очевидно это определение совпадает с определением начального момента порядка s в механике, если на оси абсцисс в точках х1, …, хn сосредоточена масса р1, …, рn.

Для непрерывной случайной величины Х начальным моментом s-го порядка называется интеграл

Очевидно, что

 ,

т.е. начальный момент s-го порядка случайной величины Х есть не что иное, как математическое ожидание s-ой степени этой случайной величины.

Перед тем как дать определение центрального момента введем понятие *"центрированной случайной величины".*

Пусть имеется случайная величина Х с математическим ожиданием mx. *Центрированной случайной величиной,* соответствующей величине Х, называется отклонение случайной величины Х от её математического ожидания

Нетрудно видеть, что математическое ожидание центрированной случайной величины равно нулю.

Центрирование случайной величины равносильно переносу начала координат в точку, абсцисса которой равна математическому ожиданию.

*Центральным моментом* порядка s случайной величины Х называется математическое ожидание s-ой степени соответствующей центрированной случайной величины:

 .

Для прерывной случайной величины s-й центральный момент выражается суммой

,

а для непрерывной - интегралом

 .

Важнейшее значение имеет второй центральный момент, который называют *дисперсией* и обозначают D[X]. Для дисперсии имеем

 .

*Дисперсия случайной величины* есть характеристика рассеивания, разбросанности значений случайной величины около её математического ожидания. Само слово "дисперсия" означает "рассеивание".

Механической интерпретацией дисперсии является не что иное, как момент инерции заданного распределения масс относительно центра тяжести.

На практике часто применяется также величина

 ,

называемая *средним квадратичным отклонением* (иначе - "стандартом") случайной величины Х.

Теперь перейдем к рассмотрению характеристик систем случайных величин.

*Начальным моментом порядка k,s системы (Х, Y)* называется математическое ожидание произведения Xk и Ys,

 xk,s=M[XkYs].

*Центральным моментом порядка k,s системы (Х, Y)* называется математическое ожидание произведения k-ой и s-ой степени соответствующих центрированных величин:

 ,

где , .

Для прерывных случайных величин

 ,

где рij - вероятность того , что система (Х, Y) примем значения (xi, yj), а сумма рассматривается по всем возможным значениям случайных величин X,Y.

Для непрерывных случайных величин

 ,

где f(x,y) - плотность распределения системы.

Помимо чисел k и s, характеризующих порядок момента по отношению к отдельным величинам, рассматривается ещё *суммарный порядок момента k+s*, равный сумме показателей степеней при Х и Y. Соответственно суммарному порядку моменты классифицируют на первый, второй и т.д. На практике обычно применяются только первые и вторые моменты.

*Первые начальные моменты* представляют собой математические ожидания величин Х и Y, входящих в систему

 σ1,0=mx σ0,1=my.

Совокупность математических ожиданий mx  , my представляет собой характеристику положения системы. Геометрически это координаты средней точки на плоскости, вокруг которой происходит рассеивание точки (Х, Y).

Важную роль на практике играют также *вторые центральные моменты* систем. Два из них представляют собой дисперсии величин Х и Y

 ,

характеризующие рассеивание случайной точки в направлении осей Ox и Oy.

Особую роль играет второй смещенный центральный момент:

 ,

называемый *корреляционным моментом* (иначе - "моментом связи")случайных величин Х и Y.

*Корреляционный момент* есть характеристика системы случайных величин, описывающая, помимо рассеивания величин Х и Y, еще и связь между ними. Для того, чтобы убедиться в этом отметим, что корреляционный момент независимых случайных величин равен нулю.

Заметим, что корреляционный момент характеризует не только зависимость величин, но и их рассеивание. Поэтому для характеристики связи между величинами (Х;Y) в чистом виде переходят от момента Kxy к характеристике

 , (3)

где σx, σy - средние квадратичные отклонения величин Х и Y. Эта характеристика называется *коэффициентом корреляции* величин Х и Y.

Из формулы (3) видно, что для независимых случайных величин коэффициент корреляции равен нулю, так как для таких величин kxy=0.

Случайные величины, для которых rxy=0, называют *некоррелированными (несвязанными).*

Отметим однако, что из некоррелированности случайных величин не следует их независимость.

Коэффициент корреляции характеризует не всякую зависимость, а только так называемую линейную зависимость. Линейная вероятностная зависимость случайных величин заключается в том, что при возрастании одной случайной величины другая имеет тенденцию возрастать (или же убывать) по линейному закону. Т.о., коэффициент корреляции характеризует степень тесноты линейной зависимости между случайными величинами.

Для определения коэффициента корреляции имеется несколько методов. Однако мы приведем пример с использованием коэффициента корреляции смешанных моментов Пирсона, где

 (4)

с применением таблицы данных (в нашем примере относительного содержания Т-лимфоцитов в % и уровня IgG в г/л):

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| x | Y | x' | y' | x'y' | x'2 | y'2 |
| 69,8 | 17,1 | 1,9 | 1,65 | 3,135 | 3,61 | 2,7225 |
| 69,5 | 16,9 | 1,6 | 1,45 | 2,32 | 2,56 | 2,1025 |
| 68,8 | 16,2 | 0,9 | 0,75 | 0,675 | 0,81 | 0,5625 |
| 67,5 | 14,8 | -0,4 | -0,65 | 0,26 | 0,16 | 0,4225 |
| 66,4 | 14,1 | -1,5 | -1,35 | 2,025 | 2,25 | 1,8225 |
| 65,5 | 13,6 | -2,4 | -1,85 | 5,76 | 5,76 | 3,4225 |
| ∑=679 | ∑=154,5 |  |  | ∑=12,855 | ∑=15,15 | ∑=11,055 |

Подставив полученные значения в формулу (4), получим

Т.е., коэффициент корреляции динамики Т-лимфоцитов и иммуноглобулина G у детей при перитонитах равен 0,9933, что говорит о высокой связи между данными показателями.

 СПИСОК ЛИТЕРАТУРЫ.

1. Гублер Е.В., Генкин А.А. Применение непараметрических критериев статистики в медико-биологических исследованиях. - Л.: Медицина, 1973.
2. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и её инженерные приложения. - М.: Наука, 1988.
3. Применение вычислительной техники и математической теории эксперимента в научных исследованиях (учебное пособие).// Под ред. М-Б.А. Бабаева. - Баку, "Елм". - 1999. - 85 стр.