

*Автор:*

*Митрофанова Мария Николаевна*

студентка

*Научный руководитель:*

*Щербакова Ирина Викторовна*

старший преподаватель

ФГБОУ ВО «Саратовский государственный медицинский

университет им. В.И. Разумовского» Минздрава России

г. Саратов, Саратовская область

## **КОРРЕЛЯЦИЯ В МЕДИЦИНЕ**

*Аннотация:* грамотная статистическая обработка данных – неотъемлемая составная часть современной медицинской науки и практики. В статье подробно рассматривается процесс выявления корреляции между величинами, имеющими медико-биологическое значение.

*Ключевые слова:* коэффициент корреляции, корреляция, статистическая обработка данных, медицинская статистика.

Приступая к изучению методики установления корреляционной зависимости между величинами, характеризующими протекание того или иного процесса в медицине, представляется необходимым определить понятие корреляции, рассмотреть ее место в системе функциональных зависимостей. В рамках этого подхода отметим, что функциональной называется такая связь между двумя признаками, при которой каждому значению одного из них соответствует значение другого. Данный тип связи очень характерен для физико-математических величин – таких, как скорость движения и время прибытия объекта; масса тела и ускорение, приобретаемое им в процессе движения; радиус круга и длина окружности; уравнение прямой и угол ее наклона к оси координат и т. д.

О корреляционной связи говорят тогда, когда каждому значению одного признака соответствует несколько значений другого признака. Термин «корреляция» происходит от латинского слова *correlatio* («соотношение, взаимосвязь»)

и подразумевает *статистическую* взаимосвязь между двумя или более *случайными величинами*. При этом изменения значений одной или нескольких из этих величин сопутствуют систематическому изменению значений другой или других величин [1].

Корреляционная связь характерна для медико-биологических процессов. Методика установления корреляции используется при выявлении взаимосвязи между температурой тела и частотой пульса, между возрастом и массой тела, между массой тела и массой мозга, между массами самки и ее детеныша, между стажем работы и уровнем заболеваемости рабочих и др.

По характеру взаимосвязи корреляция может быть положительная либо отрицательная, линейная либо нелинейная. Также выделяются прямая корреляционная связь, при которой с увеличением одного признака возрастает и другой, и обратная, когда с увеличением одного признака другой уменьшается [2].

Установление наличия и вида корреляции возможно посредством построения корреляционного поля, под которым понимается область расположения в координатной плоскости точек, координатами которых являются значения случайных величин. Подтверждение корреляции между исследуемыми параметрами осуществляется с помощью расчетов по определённой методике.

Установление связи между случайными величинами, а также силы и направления характеризует коэффициент корреляции  $r$ .

Вычислить коэффициент корреляции можно по формуле:

$$r = \frac{\sum_{i=1} (X_i - \bar{X}) * (Y_i - \bar{Y})}{n * \sigma_x * \sigma_y} \quad (1)$$

где:  $X_i$  и  $Y_i$  – сопряженные значения величин  $X$  и  $Y$ ;

$\bar{X}$  и  $\bar{Y}$  – усреднённые по всей выборке значения  $X$  и  $Y$ ;

$\sigma_x$  и  $\sigma_y$  – средние квадратические отклонения  $X$  и  $Y$ ;  $n$  – объём выборки.

Значение коэффициента корреляции, вычисленное по приведённой формуле, может оказаться как положительным, так и отрицательным, поэтому для

установления силы корреляции оценивают модуль данной величины на основе критериев, приведенных в табл. 1.

Таблица 1

### Методика установления силы корреляции

Значение $ r $	Вид корреляции
$0 <  r  < 0,3$	слабая
$0,3 \leq  r  \leq 0,7$	средняя (значительная)
$0,7 <  r  < 1$	сильная

Разберем последовательность расчета коэффициента корреляции на следующем примере.

Изучали зависимость между массой тела и возрастом детей по эмпирическим данным, приведенным в табл. 2.

Таблица 2

### Исходные данные для расчета

Возраст, лет	1	2	3	4	5	6	7	8	9	10
Масса, кг	9	11	12,5	14	16	18	20	22	24	25

Предварительные данные о корреляции между массой тела  $X$  и возрастом  $Y$  можно получить на основе корреляционного поля (рисунок 1).

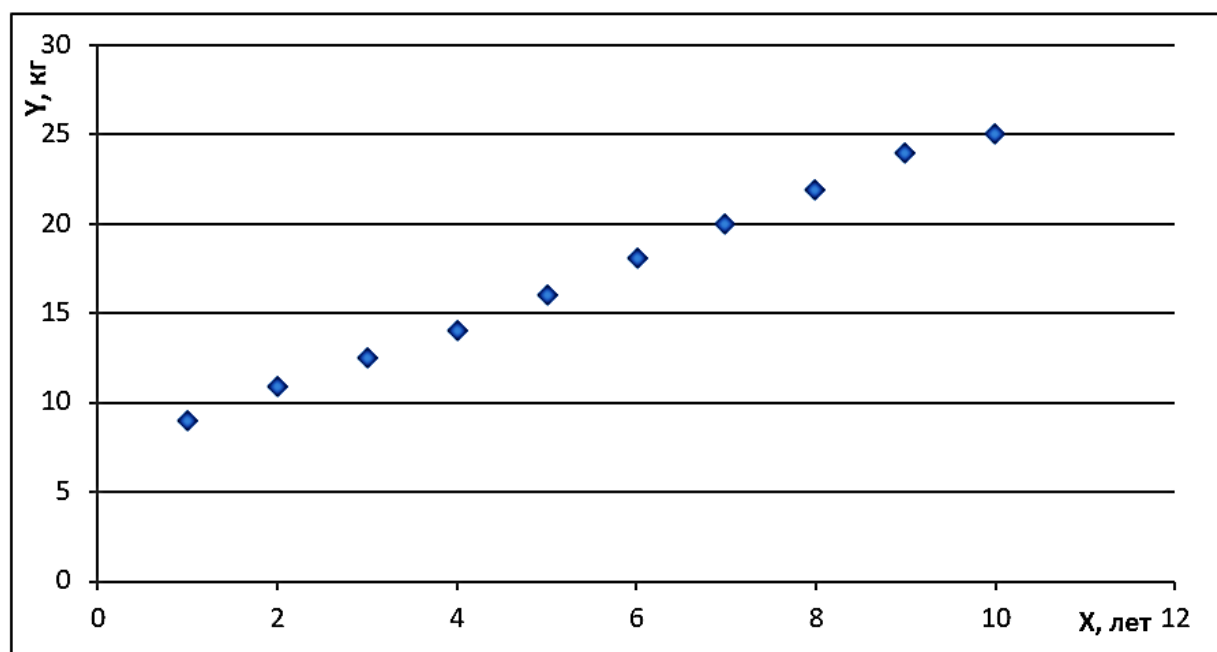


Рис. 1

### Корреляционное поле

Корреляционное поле свидетельствует о наличии корреляции между исследуемыми величинами X и Y. Так как с ростом X наблюдается к росту тенденция величины Y, корреляция является положительной. Характер зависимости приблизительно линейный. Следовательно, в качестве параметрического показателя может быть использован коэффициент корреляции [3].

Для расчета коэффициента корреляции удобно воспользоваться таблицей по нижеприведенной форме (табл. 3).

Таблица 3

Расчетная таблица

$X_i$	$Y_i$	$\bar{X}_i - \bar{X}$	$\bar{Y}_i - \bar{Y}$	$(\bar{X}_i - \bar{X})(\bar{Y}_i - \bar{Y})$	$(\bar{X}_i - \bar{X})^2$	$(\bar{Y}_i - \bar{Y})^2$
1	9	-4,5	-8,15	36,675	20,25	66,4225
2	11	-3,5	-6,15	21,525	12,25	37,8825
3	12,5	-2,5	-4,65	11,625	6,25	21,6225
4	14	-1,5	-3,15	6,975	2,25	9,9225
5	16	-0,5	-1,15	0,575	0,25	1,3225
6	18	0,5	0,85	0,425	0,25	0,7225
7	20	1,5	2,85	4,275	2,25	8,1225
8	22	2,5	4,85	12,125	6,25	23,5225
9	24	3,5	6,85	23,975	12,25	46,9225
10	25	4,5	7,85	35,325	20,25	61,6225
$\bar{X}=5,5$	$\bar{Y}=17,5$			$\Sigma=15,3$	$\Sigma=82,5$	$\Sigma=278,025$

Среднее квадратичное отклонение  $\sigma$ , необходимое для расчета коэффициента корреляции по формуле (1), можно найти следующим образом:

$$\sigma(x) = \sqrt{D(x)} \quad (2)$$

т.е. как квадратный корень из дисперсии

$$\bar{D}(x) = \frac{1}{n-1} * \sum_{i=1}^n (X_i - \bar{X})^2 \quad (3)$$

Вычисляем:

$$D_x=8,25; \sigma_x= 2,9$$

$$D_Y=27,8; \sigma_Y = 5,3$$

Рассчитаем коэффициент  $r$ , поставив в формулу (1) все необходимые величины:

$$r = 15,3 / (10 \times 2,9 \times 5,3) = 0,995$$

Полученное значение  $r$  подтверждает наличие сильной положительной корреляции между возрастом и массой детей. Последовательность такого расчёта представляется несложной даже для студентов младших курсов медицинских вузов [4; 5]. Однако статистическая обработка данных требует внимательности. Также следует подчеркнуть необходимость корректного представления исходных данных.

Зачастую в практических целях приходится выявлять зависимость между каким-либо признаком, а также влиянием на этот признак различных факторов. Это можно сделать с помощью статистического метода. Необходимо изучать корреляционную взаимосвязь, определять ее особенности, размеры и направление, а также оценивать ее достоверность [6; 7]. В медицине проводятся обязательные исследования для определения эффективности лекарственного препарата в зависимости от его дозы, между сроками госпитализации и частотой осложнений, между разными уровнями физических факторов окружающей среды и состоянием здоровья. Особенно реальную пользу применение аппарата корреляционного анализа может принести на стадии ранних исследований в областях, где характеры причин определённых явлений ещё недостаточно понятны. Установление корреляции, в свою очередь, является основой для проведения регрессионного анализа, позволяющего представить зависимость между изучаемыми величинами в виде уравнения, т.е. в аналитической форме, выявлять направления изменения признаков, тренды и прогнозировать значения медико-биологических параметров.

### ***Список литературы***

1. Гмурман В.Е. Теория вероятностей и математическая статистика: Учеб. пособие для вузов. – М.: Высш.шк., 2004.
2. Биометрия: Учеб.-метод. пособие / Сост. Г.А. Козлов, А.Е. Луньков, Б.А. Дворкин, С.В. Трубецкова. – Саратов: Изд-во Саратов. мед. ун-та, 2012.
3. Снитко А.В. Методика расчета коэффициента корреляции и составления уравнения регрессии // Научное и образовательное пространство: перспективы развития: материалы Междунар. науч.-практ. конф. / Редкол.: О.Н. Широков [и др.]. – Чебоксары: ЦНС «Интерактив плюс», 2015.
4. Магомедов А.М. Использование статистических методов в медицинских исследованиях / А.М. Магомедов, И.В. Щербакова // Бюллетень медицинских Интернет-конференций. – 2015. – Vol. 4. – Iss. 11. – P. 1270–1271.
5. Курышова В.А. Статистический анализ данных: просто или сложно? (точка зрения студента) / В.А. Курышова, И.В. Щербакова // Материалы Всерос. науч.-практ. интернет-конф. «YSRP-2014»: секция «Медицинская и биологическая физика», рубрика «Биофизика» [Электронный ресурс]. – Режим доступа: <http://medconfer.com/node/4457>
6. Уразгельдиева Л.М. К проблеме представления опытных данных медицинских исследований / Л.М. Уразгельдиева, В.Д. Мизерная // YSRP-2016: Материалы Всерос. науч.-практ. интернет-конф. студ. и молодых учёных науч.-образоват. мед. кластера «Нижевожский» [Электронный ресурс]. – Режим доступа: <http://medconfer.com/node/11083> (дата обращения: 12.12.2016).
7. Антонова В.М. Пример обработки и визуализации статистических данных на основе программы Excel / В.М. Антонова, И.В. Щербакова // Молодежный научный форум: Технические и математические науки: Сб. ст. по материалам XXXVIII студ. междунар. науч.-практ. конф. – М.: МЦНО, 2016. – №9 (38).