

МЕТОД КОРРЕЛЯЦИИ

При проведении исследования в биологии или медицине, как правило, регистрируются множество учетных признаков. Представляет интерес вопрос об их взаимном изменении, т.е. обнаружение зависимостей между ними. Выявление наличия таких взаимосвязей является одной из важнейших задач любой науки, в том числе и медицины.

Различают две формы количественных связей между явлениями или процессами: функциональную и корреляционную. Под ФУНКЦИОНАЛЬНОЙ понимают такую связь, при которой любому значению одного из признаков соответствует строго определенное значение другого. В точных науках, таких, как физика, химия и другие, может быть установлена функциональная взаимосвязь. Например, зависимость площади круга от длины окружности в геометрии, или в физике длина пути, пройденной телом в свободном падении, от времени. Наиболее известным видом функциональной зависимости является линейная, которая выражается математической формулой: $y = ax + b$.

В биологии и медицине установить функциональную зависимость, как правило, не удается. Объекты этих исследований имеют большую изменчивость и зависят от огромного числа факторов, измерить которые просто невозможно. В этом случае определяется наличие КОРРЕЛЯЦИОННОЙ связи, при которой значению каждой средней величины одного признака соответствует несколько значений другого взаимосвязанного с ним признака. Например: связь между ростом и массой тела человека. У группы людей с одинаковым ростом наблюдается различная масса тела, однако она варьирует в определенных пределах вокруг средней величины. Поэтому такую зависимость нужно оценивать с использованием понятия случайной величины с привлечением подходов теории вероятности. Такую форму зависимостей называют «Корреляционной».

При поиске зависимости между признаками может быть обнаружена взаимосвязь, различная по направлению и силе:

- Прямая (при увеличении одного признака увеличивается второй);
- Обратная (при увеличении одного признака второй уменьшается).

Степень взаимосвязи признаков по силе (тесноте) принято обозначать как:

- Отсутствие;
- Слабая;
- Средняя;
- Сильная;
- Полная.

Способами выявления корреляционной взаимосвязи между признаками являются:

- Визуальные (таблицы и графики).
- Статистические (корреляция и регрессия).

Следует подчеркнуть, что обнаружение корреляции между двумя признаками еще не говорит о существовании причинной связи между ними, а лишь указывает на возможность таковой или на наличие фактора, определяющего изменение обеих переменных совместно.

Приёмы визуализации данных позволяют обнаружить корреляционную зависимость лишь при небольшом числе наблюдений и только приблизительно. Для обнаружения корреляционной взаимосвязи с помощью таблицы в ней располагают ранжированные вариационные ряды и затем определяют совместное изменение признаков. График более наглядно демонстрирует такую зависимость и позволяет оценить ее форму: линейная, параболическая, тригонометрическая и др.

Наиболее точным способом обнаружения взаимосвязи между признаками является вычисление коэффициента корреляции. В зависимости от природы обрабатываемых данных применяются параметрические или непараметрические методы вычисления этого коэффициента.

При вычислении коэффициента корреляции исследователь получает возможность судить о силе связи (степени сопряженности) и ее направлении, а также с требуемой долей вероятности делать вывод о проявлении этой связи в генеральной совокупности. Чем больше коэффициент корреляции, тем с большей степенью уверенности можно говорить о наличии корреляционной зависимости между признаками. Если каждому заданному значению одного признака соответствуют близкие друг к другу, тесно расположенные около средней величины значения другого признака, то связь является более тесной. Когда эти значения сильно варьируют, связь менее тесная. Таким образом, мера корреляции указывает, насколько тесно связаны между собой параметры.

Коэффициент корреляции может принимать значения от -1 до +1. Направление обнаруженной взаимосвязи определяют по знаку коэффициента корреляции. При его положительном значении обнаруженная связь является прямой, при отрицательном – обратной. Сила связи оценивается по модулю этого коэффициента. Условно выделяют следующие уровни корреляционной связи: отсутствие – 0; слабая – от 0 до 0,3; средняя – от 0,3 до 0,7; сильная – 0,7 и более; полная – 1. Однако обсуждать наличие корреляции имеет смысл только в тех случаях, когда она статистически значима ($p < 0,05$). Поэтому после вычисления коэффициента корреляции производится определение его ошибки репрезентативности и критерия достоверности.

Наиболее часто применяемыми в настоящее время методами обнаружения корреляции являются параметрический анализ по Пирсону и непараметрический анализ по Спирмену. Этими методами проверяется нулевая гипотеза (H_0) об отсутствии связи между параметрами. Если такая гипотеза отклоняется при

заданном уровне значимости (p), можно говорить о наличии взаимосвязи между параметрами.

Корреляционный анализ по Пирсону используется при решении задачи исследования линейной связи двух нормально распределенных параметров. Кроме проверки на нормальность распределения каждого параметра, до проведения корреляционного анализа рекомендуется строить график в координатах оцениваемых параметров, чтобы визуально определить характер зависимости.

КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ПИРСОНА (r_{xy}) или коэффициент линейной корреляции, был разработан в 90-х годах XIX века Карлом Пирсон, Фрэнсисом Эджуортом и Рафаэлем Уэлдоном в Англии. Он рассчитывается по формуле:

$$r_{xy} = \frac{COV_{XY}}{\sigma_X \sigma_Y} = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2 \sum(Y - \bar{Y})^2}}$$

- где: r_{xy} – коэффициент линейной корреляции Пирсона;
 COV_{XY} – ковариация признаков X и Y ;
 σ_X – среднее квадратическое отклонение признака X ;
 σ_Y – среднее квадратическое отклонение признака Y ;
 \bar{X} – средняя арифметическая признака X ;
 \bar{Y} – средняя арифметическая признака Y .

В медицинской литературе встречается упрощенная запись этой формулы:

$$r_{xy} = \frac{\sum d_x d_y}{\sqrt{\sum d_x^2 \sum d_y^2}},$$

- где: r_{xy} – коэффициент линейной корреляции Пирсона;
 d_x – отклонение каждой варианты признака x от средней этого признака: $d_x = x - M_x$,
 d_y – отклонение каждой варианты признака y от средней этого признака: $d_y = y - M_y$.

В программе Excel значение коэффициент линейной корреляции Пирсона может быть вычислено функцией =КОРРЕЛ(Диапазон ячеек 1-го ряда; Диапазон ячеек 2-го ряда).

Для прогнозирования уровня корреляции в генеральной совокупности определяют ошибку репрезентативности этого коэффициента m_r . Она вычисляется по формуле:

$$m_r = \sqrt{\frac{1-r_{xy}^2}{n-2}},$$

где: m_r – ошибка репрезентативности коэффициента корреляции;
 r_{xy} – коэффициент линейной корреляции Пирсона;
 n – число парных вариант.

Достоверность коэффициента линейной корреляции оценивается по коэффициенту Стьюдента (t_r), который вычисляется с использованием его ошибки:

$$t_r = \frac{r_{xy}}{m_r},$$

где: t_r – коэффициент достоверности Стьюдента;
 r_{xy} – коэффициент линейной корреляции Пирсона;
 m_r – ошибка репрезентативности коэффициента корреляции.

Если число парных вариант $n > 30$, то при $t_r > 2$ связь считается достоверной при уровне значимости $p < 0,05$. Если число парных вариант $n < 30$, то критическое значение $t_{r-Крит.}$ находят по таблице критических значений Стьюдента при степени свободы $df = n - 2$. В программе Excel это значение вычисляется функцией =СТЮДРАСПОБР(Уровень значимости p ; Степени свободы df).

С целью уменьшения объема вычислений может применяться функция =КОРРЕЛ(Диапазон1; Диапазон2) или надстройка «Анализ данных» и ее модуль «Корреляционный анализ».

Отсутствие линейной корреляции еще не означает, что параметры полностью независимы. Связь между ними может быть нелинейной, или признаки, используемые в вычислениях, могут не подчиняться нормальному закону распределения. Поэтому, помимо вычисления коэффициента линейной корреляции, прибегают к использованию непараметрических коэффициентов корреляции. К ним относятся:

- Коэффициент ранговой корреляции Спирмена;
- Коэффициент ранговой корреляции Кендалла;
- Коэффициент корреляции знаков Фехнера;
- Коэффициент множественной ранговой корреляции (конкордации).

КОРРЕЛЯЦИОННЫЙ АНАЛИЗ ПО СПИРМЕНУ применяется для обнаружения взаимосвязи двух параметров, если распределение хотя бы одного из них отлично от нормального.

Каждому показателю x и y присваивается ранг. На основе полученных рангов рассчитываются их разности d . Затем вычисляется коэффициент корреляции (ρ) по формуле:

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)},$$

где: ρ – коэффициент корреляции Спирмена;
 d – разность рангов;

n – число парных вариантов.

Ошибка репрезентативности коэффициента корреляции Спирмена определяется по формуле:

$$m_{\rho} = \sqrt{\frac{1-\rho^2}{n-2}},$$

а коэффициент достоверности Стьюдента:

$$t_{\rho} = \frac{\rho}{m_{\rho}},$$

где: t_{ρ} – коэффициент достоверности Стьюдента;

ρ – коэффициент корреляции Спирмена;

m_{ρ} – ошибка репрезентативности коэффициента корреляции Спирмена.

Оценка коэффициента корреляции Спирмена и его достоверности выполняется так же, как и коэффициента линейной корреляции Пирсона.

Пример решения задачи на выявление корреляционной зависимости.

Условие задачи: выполнены измерения уровня запыленности на рабочих местах работников с учетом температуры в помещении (таблица 1).

Таблица 1

Результаты измерений запыленности в помещении с учетом температуры окружающей среды

| Измерение на рабочем месте | Температура воздуха С° | Запыленность мг/м ³ |
|----------------------------|------------------------|--------------------------------|
| 1. Слесарь | 20 | 0,2 |
| 2. Электрик | 21 | 0,25 |
| 3. Сварщик | 21 | 0,24 |
| 4. ... | 19 | 0,08 |
| 5. ... | 19 | 0,08 |
| 6. ... | 19 | 0,07 |
| 7. ... | 22 | 0,3 |
| 8. ... | 22 | 0,28 |
| 9. ... | 25 | 0,33 |
| 10. ... | 24 | 0,31 |
| 11. ... | 21 | 0,26 |
| 12. ... | 21 | 0,27 |

Задание: определить силу и направление зависимости между температурой окружающей среды и уровнем запыленности помещения с помощью:

а) таблицы;

- б) графического изображения взаимосвязи между признаками;
- в) коэффициента корреляции Пирсона;
- г) коэффициента корреляции Спирмена.

Решение: запустите программу Excel, переименуйте Лист 1, обозначив названием «Коррел-я». На этом листе введите данные и решение задачи, как показано ниже, сохраните изменения.

а) Требуется выполнить ранжирование вариационных рядов и поместить их рядом друг с другом, как показано на рисунке 1, а затем проанализировать совместное возрастание или убывание значений.

| Корреляция по Пирсону | | | | | | | | |
|-----------------------|--------------------------|------------------------|-------|-------|-------------|---------|---------|--|
| Варианта | Темпера-тура воздуха (x) | Запыленность мг/м3 (y) | d_x | d_y | $d_x * d_y$ | d_x^2 | d_y^2 | |
| 1 | 19 | 0,07 | | | | | | |
| 2 | 19 | 0,08 | | | | | | |
| 3 | 19 | 0,08 | | | | | | |
| 4 | 20 | 0,2 | | | | | | |
| 5 | 21 | 0,24 | | | | | | |
| 6 | 21 | 0,25 | | | | | | |
| 7 | 21 | 0,26 | | | | | | |
| 8 | 21 | 0,27 | | | | | | |
| 9 | 22 | 0,3 | | | | | | |
| 10 | 22 | 0,28 | | | | | | |
| 11 | 24 | 0,31 | | | | | | |
| 12 | 25 | 0,33 | | | | | | |

Рис. 1

Вывод: в таблице наблюдается совместное прямое изменение двух изучаемых параметров.

б) Построение графика и оценка формы взаимосвязи между признаками (рис. 2).



Рис. 2

Вывод: график зависимости совместного изменения двух изучаемых параметров показывает наличие взаимосвязи, которая приближенно оценивается как линейная.

в) Вычисление и оценка коэффициента корреляции методом Пирсона. Вычисление отклонений вариант от средней арифметической (см. рис. 3)

| К18 | | f _x | | | | | | |
|-----|------------------------------|--------------------------------|-------------------------------|----------------------|----------------------|------------------------------------|----------------------------------|----------------------------------|
| | A | B | C | D | E | F | G | H |
| 1 | Корреляция по Пирсону | | | | | | | |
| 2 | Варианта | Температура воздуха (x) | Запыленность мг/м3 (y) | d_x | d_y | d_x*d_y | d_x² | d_y² |
| 3 | 1 | 19 | 0,07 | -2,2 | -0,153 | 0,330 | 4,7 | 0,0233 |
| 4 | 2 | 19 | 0,08 | -2,2 | -0,143 | 0,309 | 4,7 | 0,0203 |
| 5 | 3 | 19 | 0,08 | -2,2 | -0,143 | 0,309 | 4,7 | 0,0203 |
| 6 | 4 | 20 | 0,2 | -1,2 | -0,023 | 0,026 | 1,4 | 0,0005 |
| 7 | 5 | 21 | 0,24 | -0,2 | 0,018 | -0,003 | 0,0 | 0,0003 |
| 8 | 6 | 21 | 0,25 | -0,2 | 0,028 | -0,005 | 0,0 | 0,0008 |
| 9 | 7 | 21 | 0,26 | -0,2 | 0,038 | -0,006 | 0,0 | 0,0014 |
| 10 | 8 | 21 | 0,27 | -0,2 | 0,048 | -0,008 | 0,0 | 0,0023 |
| 11 | 9 | 22 | 0,3 | 0,8 | 0,078 | 0,065 | 0,7 | 0,0060 |
| 12 | 10 | 22 | 0,28 | 0,8 | 0,058 | 0,048 | 0,7 | 0,0033 |
| 13 | 11 | 24 | 0,31 | 2,8 | 0,088 | 0,248 | 8,0 | 0,0077 |
| 14 | 12 | 25 | 0,33 | 3,8 | 0,108 | 0,412 | 14,7 | 0,0116 |
| 15 | Средняя (M) = | 21,2 | 0,223 | Сумма (Σ) = | | 1,725 | 39,7 | 0,0976 |
| 16 | n= | 12 | | | | | | |

Рис. 3

Коэффициент корреляции вычисляется по формуле:

$$r_{xy} = \frac{\sum d_x d_y}{\sqrt{\sum d_x^2 \sum d_y^2}} = \frac{1,725}{\sqrt{39,7 \times 0,0976}} = \mathbf{0,88}.$$

В программе Excel может использоваться функция =КОРРЕЛ(Диапазон1;Диапазон2) или модуль «Корреляция», который вызывается командой «Данные» - «Анализ данных». Он производит создание таблицы, которая называется «Корреляционная матрица», что позволяет вычислить коэффициент корреляции для нескольких признаков одновременно. Результат вычислений, выполненный с помощью указанного модуля, показан на рис 4.

| | | |
|-------------------------|--------------------------------|-------------------------------|
| | <i>Температура воздуха (x)</i> | <i>Запыленность мг/м3 (y)</i> |
| Температура воздуха (x) | 1 | |
| Запыленность мг/м3 (y) | 0,876588407 | 1 |

Рис. 4

Оценка достоверности коэффициента корреляции с помощью критерия Стьюдента (с. рис 5):

$$m_r = \sqrt{\frac{1-r^2}{n-2}} = \sqrt{\frac{1-0,88^2}{12-2}} = \underline{0,152},$$

где: n – число парных вариант.

Критерий достоверности Стьюдента для коэффициента корреляции вычисляется по формуле:

$$t_r = \frac{r_{xy}}{m_r} = \frac{0,88}{0,152} = \underline{5,8}$$

| | A | B | C | D | E | F | G | H | I | J | K |
|----|-----------------------|-------------------------|------------------------|----------------|----------------|----------------------------------|-----------------------------|-----------------------------|---|---|---|
| 1 | Корреляция по Пирсону | | | | | | | | | | |
| 2 | Варианта | Температура воздуха (x) | Запыленность мг/м3 (y) | d _x | d _y | d _x *d _y | d _x ² | d _y ² | | | |
| 3 | 1 | 19 | 0,07 | -2,2 | -0,153 | 0,330 | 4,7 | 0,0233 | | | |
| 4 | 2 | 19 | 0,08 | -2,2 | -0,143 | 0,309 | 4,7 | 0,0203 | | | |
| 5 | 3 | 19 | 0,08 | -2,2 | -0,143 | 0,309 | 4,7 | 0,0203 | | | |
| 6 | 4 | 20 | 0,2 | -1,2 | -0,023 | 0,026 | 1,4 | 0,0005 | | | |
| 7 | 5 | 21 | 0,24 | -0,2 | 0,018 | -0,003 | 0,0 | 0,0003 | | | |
| 8 | 6 | 21 | 0,25 | -0,2 | 0,028 | -0,005 | 0,0 | 0,0008 | | | |
| 9 | 7 | 21 | 0,26 | -0,2 | 0,038 | -0,006 | 0,0 | 0,0014 | | | |
| 10 | 8 | 21 | 0,27 | -0,2 | 0,048 | -0,008 | 0,0 | 0,0023 | | | |
| 11 | 9 | 22 | 0,3 | 0,8 | 0,078 | 0,065 | 0,7 | 0,0060 | | | |
| 12 | 10 | 22 | 0,28 | 0,8 | 0,058 | 0,048 | 0,7 | 0,0033 | | | |
| 13 | 11 | 24 | 0,31 | 2,8 | 0,088 | 0,248 | 8,0 | 0,0077 | | | |
| 14 | 12 | 25 | 0,33 | 3,8 | 0,108 | 0,412 | 14,7 | 0,0116 | | | |
| 15 | Средняя (M) = | 21,2 | 0,223 | Сумма (Σ) = | | 1,725 | 39,7 | 0,0976 | | | |
| 16 | n= | 12 | | | | | | | | | |
| 17 | | | | | | | | | | | |
| 18 | | | r _{xy} = | 0,88 | | Функция =КОРРЕЛ(В3:В14;С3:С14) = | | 0,8765884 | | | |
| 19 | | | m _{xy} = | 0,152 | | | | | | | |
| 20 | | | t= | 5,8 | | | | | | | |
| 21 | | | | | | | | | | | |

Рис. 5

Вывод: зависимость изменения двух изучаемых параметров является сильной прямой и статистически достоверной при уровне значимости p<0,05.

г) вычисление и оценка коэффициента корреляции методом Спирмена.

В таблице вариационных рядов производится подсчет рангов как показано на рисунке 6. Каждому из 12 чисел присваивается порядковый номер по возрастанию в соответствии с его значением. В случае появления вариант одинаковых по величине находится среднее арифметическое значение ранга. В данном примере температура воздуха 19 встречается трижды и ему присваиваются ранги 1, 2 и 3, находим среднее арифметическое этих рангов

$(1+2+3)/3= 2$. Число 20 будет иметь ранг 4 (ни одно из чисел не будет иметь ранги 2 и 3).

Затем вычисляется разность рангов $d_r = x_i - y_i$, затем квадрат разности.

| | | | | | | | |
|----|-------------------------------|--------------------------------|--|------------------|-------------------|----------------------|----------------------------------|
| 21 | | | | | | | |
| 22 | Корреляция по Спирмену | | | | | | |
| 23 | Варианта | Температура воздуха (x) | Запыленность мг/м³ (y) | Ранг по x | Ранг по y | d_r | d_r² |
| 24 | 1 | 19 | 0,07 | 2 | 1 | 1 | 1 |
| 25 | 2 | 19 | 0,08 | 2 | 2,5 | -0,5 | 0,25 |
| 26 | 3 | 19 | 0,08 | 2 | 2,5 | -0,5 | 0,25 |
| 27 | 4 | 20 | 0,2 | 4 | 4 | 0 | 0 |
| 28 | 5 | 21 | 0,24 | 6,5 | 5 | 1,5 | 2,25 |
| 29 | 6 | 21 | 0,25 | 6,5 | 6 | 0,5 | 0,25 |
| 30 | 7 | 21 | 0,26 | 6,5 | 7 | -0,5 | 0,25 |
| 31 | 8 | 21 | 0,27 | 6,5 | 8 | -1,5 | 2,25 |
| 32 | 9 | 22 | 0,28 | 9,5 | 9 | 0,5 | 0,25 |
| 33 | 10 | 22 | 0,3 | 9,5 | 10 | -0,5 | 0,25 |
| 34 | 11 | 24 | 0,31 | 11 | 11 | 0 | 0 |
| 35 | 12 | 25 | 0,33 | 12 | 12 | 0 | 0 |
| 36 | | | | | Сумма (Σ)= | | 7 |
| 37 | | | | | | | |

Рис. 6

Вычисление коэффициента корреляции Спирмена (см. рис. 7):

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} = 1 - \frac{6 \times 7}{12 \times (12^2 - 1)} \quad \rho = \underline{\underline{0,97}}$$

Вычисление ошибки репрезентативности коэффициента корреляции:

$$m = \sqrt{\frac{1 - \rho^2}{n - 2}} = \sqrt{\frac{1 - 0,97^2}{12 - 2}} \quad m = 0,077$$

Вычисление коэффициента достоверности Стьюдента для коэффициента корреляции (с рис.7):

$$t = \frac{\rho}{m} = \frac{0,97}{0,077} \quad t = \underline{\underline{12,6 > 2}}$$

| | | | | | | | | |
|----|-------------------------------|--------------------------------|--|------------------|------------------|----------------------|----------------------------------|--|
| 21 | | | | | | | | |
| 22 | Корреляция по Спирмену | | | | | | | |
| 23 | Варианта | Температура воздуха (x) | Запыленность мг/м³ (y) | Ранг по x | Ранг по y | d_r | d_r² | |
| 24 | 1 | 19 | 0,07 | 2 | 1 | 1 | 1 | |
| 25 | 2 | 19 | 0,08 | 2 | 2,5 | -0,5 | 0,25 | |
| 26 | 3 | 19 | 0,08 | 2 | 2,5 | -0,5 | 0,25 | |
| 27 | 4 | 20 | 0,2 | 4 | 4 | 0 | 0 | |
| 28 | 5 | 21 | 0,24 | 6,5 | 5 | 1,5 | 2,25 | |
| 29 | 6 | 21 | 0,25 | 6,5 | 6 | 0,5 | 0,25 | |
| 30 | 7 | 21 | 0,26 | 6,5 | 7 | -0,5 | 0,25 | |
| 31 | 8 | 21 | 0,27 | 6,5 | 8 | -1,5 | 2,25 | |
| 32 | 9 | 22 | 0,28 | 9,5 | 9 | 0,5 | 0,25 | |
| 33 | 10 | 22 | 0,3 | 9,5 | 10 | -0,5 | 0,25 | |
| 34 | 11 | 24 | 0,31 | 11 | 11 | 0 | 0 | |
| 35 | 12 | 25 | 0,33 | 12 | 12 | 0 | 0 | |
| 36 | | | | | Сумма (Σ)= | | 7 | |
| 37 | | | | | | | | |
| 38 | | | $\rho =$ | 0,97 | | | | |
| 39 | | | | | | | | |
| 40 | | | $m =$ | 0,077 | | | | |
| 41 | | | | | | | | |
| 42 | | | $t =$ | 12,6 | | | | |
| 43 | | | | | | | | |

Рис.7

Вывод: корреляционная связь двух изучаемых параметров является сильной прямой и статистически достоверной при уровне значимости $p < 0,05$.

Практические задания.

✦ Задания выполняются по вариантам. Прежде чем приступить к выполнению задания, уточните свой вариант у преподавателя.

Задание 1. Задача на выявление корреляционной зависимости.

Ход выполнения задания

1. Запустите программу Excel
2. Переименуйте Лист 1, обозначив название «Корреляция».
3. Создайте таблицу и выполните ранжирование вариационных рядов. Сделайте вывод.
4. Постройте график и оцените форму взаимосвязи между признаками. Сделайте вывод.
5. Методом Пирсона вычисление и оцените коэффициент корреляции. Сделайте вывод.

6. Вычислите и оцените коэффициент корреляции методом Спирмена. Сделайте вывод.
7. Сохраните файл.

Вариант 1

При проведении комплексных медицинских осмотров у 20 пациентов измерено два признака: вес (Р) и артериальное давление (АД). Данные измерений приведены в таблице.

| Номер пациента | Вес Р, кг | Артериальное давление АД (мм.рт.ст.) |
|----------------|-----------|--------------------------------------|
| 1 | 40 | 110 |
| 2 | 75 | 165 |
| 3 | 55 | 125 |
| 4 | 70 | 160 |
| 5 | 60 | 135 |
| 6 | 50 | 120 |
| 7 | 45 | 130 |
| 8 | 55 | 125 |
| 9 | 60 | 130 |
| 10 | 45 | 125 |
| 11 | 50 | 120 |
| 12 | 65 | 150 |
| 13 | 60 | 130 |
| 14 | 45 | 125 |
| 15 | 65 | 150 |
| 16 | 50 | 120 |
| 17 | 65 | 155 |
| 18 | 55 | 120 |
| 19 | 45 | 120 |
| 20 | 55 | 130 |

Определите силу и направление зависимости между весом и артериальным давлением с помощью таблицы, графического изображения взаимосвязи между признаками, коэффициентов корреляции Пирсона и Спирмена, сделайте вывод.

Вариант 2

Уровни систолического и диастолического давления (в мм. рт. ст.) у 12 здоровых юношей в возрасте 18 лет.

| Систолическое | Диастолическое |
|---------------|----------------|
| 105 | 65 |
| 115 | 70 |
| 115 | 65 |
| 110 | 65 |
| 110 | 70 |
| 120 | 75 |
| 120 | 70 |
| 125 | 75 |
| 110 | 70 |
| 125 | 80 |
| 120 | 80 |
| 120 | 75 |

Определите силу и направление зависимости между СД и ДД с помощью таблицы, графического изображения взаимосвязи между признаками, коэффициентов корреляции Пирсона и Спирмена, сделайте вывод.

Вариант 3

Результаты измерения роста и массы тела у 12 студентов медицинской академии в возрасте 20 лет.

| Рост, см | Масса тела, кг |
|----------|----------------|
| 157 | 56 |
| 158 | 55 |
| 160 | 57 |
| 165 | 57 |
| 167 | 58 |
| 162 | 60 |

| | |
|-----|----|
| 171 | 63 |
| 174 | 65 |
| 168 | 67 |
| 176 | 72 |
| 170 | 79 |
| 180 | 82 |

Определите силу и направление зависимости между ростом и массой тела с помощью таблицы, графического изображения взаимосвязи между признаками, коэффициентов корреляции Пирсона и Спирмена, сделайте вывод.

Вариант 4

Выполнены измерения уровня заболеваемости органов дыхания на вредных производствах от содержания в воздухе двуокиси углерода

| Уровень заболеваемости | Содержание в воздухе двуокиси углерода |
|------------------------|--|
| 1160 | 0,78 |
| 1155 | 0,88 |
| 1158 | 1,1 |
| 1157 | 1 |
| 1160 | 0,9 |
| 1161 | 0,9 |
| 1157 | 0,88 |
| 1159 | 0,75 |
| 1256 | 1,2 |
| 1260 | 1,2 |
| 1040 | 0,6 |
| 1039 | 0,6 |

Определите силу и направление зависимости между уровнем заболеваемости и содержанием в воздухе двуокиси углерода с помощью таблицы, графического изображения взаимосвязи между признаками, коэффициентов корреляции Пирсона и Спирмена, сделайте вывод.

Вариант 5

Вероятность смерти¹ женщин от сосудистых поражений головного мозга в зависимости от возраста.

| Возраст, в годах | Вероятность смерти |
|-------------------------|---------------------------|
| 15 | 5,0 |
| 20 | 5,6 |
| 25 | 5,7 |
| 30 | 5,7 |
| 35 | 5,6 |
| 40 | 7,6 |
| 45 | 7,7 |
| 50 | 9,3 |
| 55 | 10,7 |
| 60 | 10,5 |
| 70 | 14,1 |

Определите силу и направление зависимости между возрастом и вероятностью смерти от сосудистых поражений головного мозга с помощью таблицы, графического изображения взаимосвязи между признаками, коэффициентов корреляции Пирсона и Спирмена, сделайте вывод.

Вариант 6

В городе Н. было проведено изучение зависимости заболеваемости инфарктом миокарда по месяцам года в зависимости от среднемесячной температуры воздуха.

| Месяцы года | Заболеваемость инфарктом миокарда по месяцам (на 10 000 жителей) | Среднемесячная температура воздуха |
|--------------------|---|---|
| январь | 1,6 | -7,1 |
| февраль | 1,23 | -7,7 |
| март | 1,14 | -5,8 |

¹ В расчете на 10000 женщин.

| | | |
|----------|------|-------|
| апрель | 1,13 | -4,1 |
| май | 1,12 | +13 |
| июнь | 1,02 | +14,9 |
| июль | 0,91 | +18,8 |
| август | 0,82 | +15,6 |
| сентябрь | 1,06 | +9,0 |
| октябрь | 1,22 | +6,0 |
| ноябрь | 1,33 | -1,0 |
| декабрь | 1,4 | -7,7 |

Определите силу и направление зависимости между среднемесячной температурой и заболеваемостью инфарктом с помощью таблицы, графического изображения взаимосвязи между признаками взаимосвязи между признаками, коэффициентов корреляции Пирсона и Спирмена, сделайте вывод

Вариант 7

Изучалась зависимость между систолическим давлением мужчин в начальной стадии шока и возрастом. Результаты наблюдений приведены в таблице:

| Возраст | Систолическое давление |
|---------|------------------------|
| 68 | 114 |
| 37 | 149 |
| 50 | 146 |
| 53 | 141 |
| 75 | 114 |
| 66 | 112 |
| 52 | 124 |
| 65 | 105 |
| 74 | 141 |
| 65 | 120 |
| 54 | 124 |

Определите силу и направление зависимости между возрастом и давлением в начальной стадии шока с помощью таблицы, графического изображения взаимосвязи между признаками, коэффициентов корреляции Пирсона и Спирмена, сделайте вывод.

Вариант 8

Выполнены измерения показателей физического развития школьников, характеризующих их рост стоя и объем грудной клетки.

| Измерение | Рост, см | Объем грудной клетки, см |
|----------------|----------|--------------------------|
| 1. Чернов А.С. | 151 | 70,8 |
| 2. Галкин М.В. | 178 | 78,2 |
| 3. Попов А.М. | 152 | 71,1 |
| 4. ... | 160 | 73,2 |
| 5. ... | 160 | 73,3 |
| 6. ... | 178 | 78,2 |
| 7. ... | 170 | 76,1 |
| 8. ... | 170 | 76,3 |
| 9. ... | 143 | 67,5 |
| 10. ... | 170 | 76,1 |
| 11. ... | 150 | 70,5 |
| 12. ... | 172 | 76,6 |

Определите силу и направление зависимости между ростом и объем грудной клетки с помощью таблицы, графического изображения взаимосвязи между признаками, коэффициентов корреляции Пирсона и Спирмена, сделайте вывод.

Вариант 9

Врачом футбольной команды выполнены измерения показателей деятельности системы кровообращения и тренированности спортсменов, измерены частота пульса и систолический объем сердечного выброса.

| Измерение | Пульс, уд/мин | Объем сердечного выброса, мл |
|------------------|----------------------|-------------------------------------|
| 1. Васильев А.С. | 78 | 58 |
| 2. Морозов Н.Р. | 72 | 38 |
| 3. Родионов А.К. | 78 | 63 |
| 4. ... | 80 | 65 |
| 5. ... | 72 | 35 |
| 6. ... | 60 | 46 |
| 7. ... | 72 | 59 |
| 8. ... | 72 | 59 |
| 9. ... | 72 | 50 |
| 10. ... | 66 | 38 |
| 11. ... | 72 | 40 |
| 12. ... | 84 | 68 |

Определите силу и направление зависимости между пульсом и систолическим объемом с помощью таблицы, графического изображения взаимосвязи между признаками, коэффициентов корреляции Пирсона и Спирмена, сделайте вывод.

Вариант 10

Изучена еженедельная заболеваемость острыми респираторными инфекциями на территории Н. в зимний период (с декабря по февраль). Данные приведены в таблице.

| Неделя | Кол-во ОРЗ | Температура воздуха, -t C⁰ |
|---------------|-------------------|--|
| 1 | 30 | 20 |
| 2 | 31 | 21 |
| 3 | 33 | 22 |
| 4 | 34 | 23 |
| 5 | 34 | 21 |
| 6 | 36 | 25 |
| 7 | 38 | 25 |
| 8 | 39 | 29 |
| 9 | 38 | 28 |

| | | |
|----|----|----|
| 10 | 36 | 23 |
| 11 | 28 | 20 |
| 12 | 34 | 22 |

Определите силу и направление зависимости между средней еженедельной температурой в зимний период и количеством острых респираторных заболеваний (ОРЗ) с помощью таблицы, графического изображения взаимосвязи между признаками, коэффициентов корреляции Пирсона и Спирмена, сделайте вывод.

Вариант 11

Были получены следующие данные о весе (в г) левой камеры сердца и длине ядер (в μ) в мышцах сердца:

| Номер опыта | Вес (гр) | Длина ядер (μ) |
|-------------|----------|----------------------|
| 1 | 304 | 23,0 |
| 2 | 632 | 28,4 |
| 3 | 256 | 15,9 |
| 4 | 262 | 20,7 |
| 5 | 273 | 19,4 |
| 6 | 289 | 19,8 |
| 7 | 291 | 11,7 |
| 8 | 292 | 21,0 |
| 9 | 207 | 16,6 |
| 10 | 328 | 13,6 |
| 11 | 372 | 19,6 |
| 12 | 397 | 22,9 |
| 13 | 460 | 19,4 |
| 14 | 221 | 18,0 |

Определите силу и направление зависимости между весом левой камеры и длиной ядер с помощью таблицы, графического изображения взаимосвязи между признаками, коэффициентов корреляции Пирсона и Спирмена, сделайте вывод.

Вариант 12

Результаты измерения массы тела у 10 студентов медицинского института и возраст студентов.

| Порядковый номер студента | Возраст в годах | Масса тела в кг |
|---------------------------|-----------------|-----------------|
| 1 | 22 | 56 |
| 2 | 22 | 57 |
| 3 | 21 | 59 |
| 4 | 24 | 60 |
| 5 | 23 | 63 |
| 6 | 23 | 65 |
| 7 | 24 | 67 |
| 8 | 23 | 72 |
| 9 | 24 | 79 |
| 10 | 24 | 82 |

Определите силу и направление зависимости между возрастом и массой тела с помощью таблицы, графического изображения взаимосвязи между признаками, коэффициентов корреляции Пирсона и Спирмена, сделайте вывод

КОНТРОЛЬНЫЕ ВОПРОСЫ:

- 1) Для чего практическому врачу необходимо знать метод корреляции?
- 2) Какие две формы количественных связей между явлениями или процессами различают?
- 3) Что понимают под функциональной связью?
- 4) Что понимают под корреляционной связью?
- 5) Охарактеризуйте корреляционную связь по направлению.
- 6) Что понимают под силой корреляции? Как она определяется?
- 7) Какие способы расчета коэффициента корреляции существуют?
- 8) Каким образом можно вычислить коэффициент корреляции в Excel?

Список литературы:

- 1) Зайцев В. М. Прикладная медицинская статистика: учеб. пособие для студентов мед. вузов / Зайцев В. М., Лифляндский В. Г., Маринкин В. И. . - СПб. : Фолиант , 2003 . - 430 с.
- 2) Ситуационные задачи по медицинской статистике с примерами решений в программе Microsoft Excel: учеб.-метод. пособие к практ. занятиям по дисциплине "Мед. информатика" для спец. : 060101 65 - Леч. дело, 060103 65 - Педиатрия, 060201 65 Стоматология, 060105 65 - Мед.-профил. дело / Голубев А. Н., Грибина Л. Н., Дьяченко Т. С. и др. ; ВолгГМУ Минздрава РФ . - Волгоград : Изд-во ВолгГМУ , 2014 . – 254с.
- 3) Сабанов В. И. Медицинская информатика и автоматизированные системы управления в здравоохранении: учеб.-метод. пособие к практ. занятиям / Сабанов В. И., Голубев А. Н., Комина Е. Р. ; Федерал. агентство по здравоохранению и соц. развитию; ВолГМУ . - Волгоград : Изд-во ВолГМУ , 2006 . - 144 с.
- 4) Голубев А.Н., Грибина Л.Н., Бирюкова Л.Ф. и т.д. Тестовые задания по медицинской информатике и статистике. Учебное пособие / Под ред. проф. В.И. Сабанова. – Волгоград: Изд-во ВолгГМУ, 2014. – 426с.