

Министерство цифрового развития, связи и
массовых коммуникаций Российской Федерации

Федеральное государственное бюджетное образовательное учреждение
высшего образования «Сибирский государственный университет
телекоммуникаций и информатики» (СибГУТИ)

Кафедра вычислительных систем

ОТЧЕТ

по лабораторной работе №1
по дисциплине «**Прикладные задачи Теории Вероятностей**»

Выполнил:

студент гр. ИС-142

«__» сентября 2023 г.

/Наумов А.А./

Проверил:

профессор кафедры В.С., д.т.н.

«__» сентября 2023 г.

/Родионов А.С./

Оценка « _____ »

Новосибирск 2023

ОГЛАВЛЕНИЕ

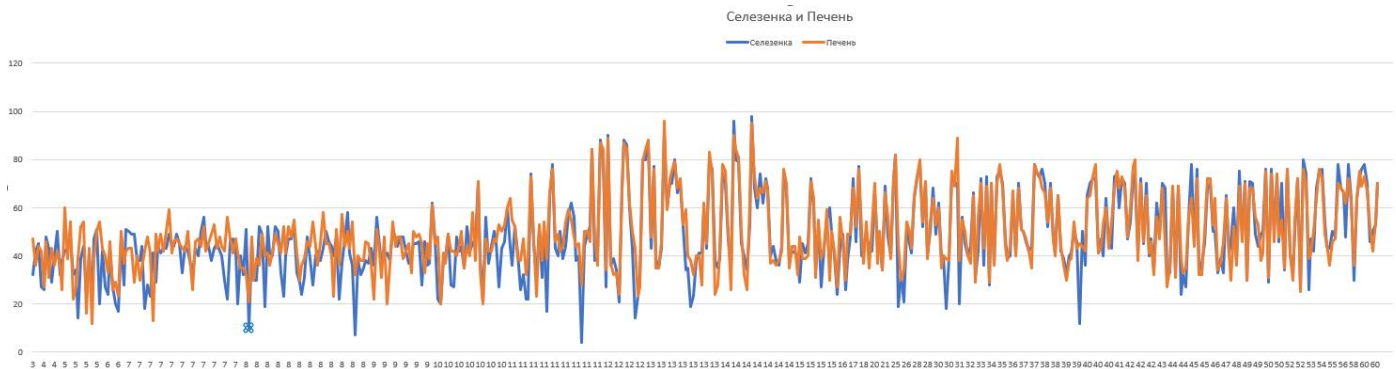
ВВЕДЕНИЕ	3
ОСНОВНАЯ ЧАСТЬ	3
ЗАКЛЮЧЕНИЕ	5

ВВЕДЕНИЕ

Цель лабораторной работы - анализ статистических данных и выдвижение с доказательством любой гипотезы в области математической статистики.

ОСНОВНАЯ ЧАСТЬ

Визуализация данных:



Исходя из графика можно сказать, что данные по “Селезенка” и “Печень” имеют взаимосвязь между собой. Также с увеличением возраста мы видим более частые всплески на графике, поэтому проверим корреляцию еще и между значениями: “Селезенка” и “Возраст”, “Печень” и “Возраст”.

Для проверки используем метод корреляции Пирсона:

$$r_{xy} = \frac{\sum (x_i - M_x) (y_i - M_y)}{\sqrt{\sum (x_i - M_x)^2 \cdot \sum (y_i - M_y)^2}}$$

Это статистическая мера, используемая для измерения степени линейной связи между двумя переменными.

Коэффициент корреляции Пирсона находится в диапазоне от -1 до 1. Значение -1 это полная отрицательная линейная корреляция, значение 1 - это положительная линейная корреляция, а значение 0 - отсутствие линейной корреляции вовсе.

Проверять будем с помощью функции “КОРРЕЛ” в Excel, ведь если посмотреть в справку по функции, можно найти формулу Пирсона, с помощью которой и считает программа (см. картинку ниже)

Функция КОРРЕЛ

- Уравнение для коэффициента корреляции имеет следующий вид:

$$\text{Correl}(X, Y) = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

Интерпретация значений корреляции:

Корреляция близка к 1 ($> +0.8$): Высокая положительная линейная корреляция. Это означает, что две переменные сильно и положительно взаимосвязаны: при увеличении одной переменной другая тоже увеличивается, и наоборот.

Корреляция близка к -1 (< -0.8): Высокая отрицательная линейная корреляция. Это означает, что две переменные сильно и отрицательно взаимосвязаны: при увеличении одной переменной другая уменьшается, и наоборот.

Корреляция близка к 0 (< 0.15): Отсутствие или очень слабая линейная корреляция. Это говорит о том, что две переменные не имеют значимой линейной связи. Изменения в одной переменной не влияют на изменения в другой.

Корреляция близка к 0.5 или -0.5: Средняя (умеренная) корреляция. Означает, что с увеличением (уменьшением) значений одной переменной, значения другой переменной склонны увеличиваться (уменьшаться), но это происходит не очень сильно.

Вычисления

Вычисления буду проводить с помощью Excel, функции “КОРРЕЛ”:

Корреляция между “Селезенка” и “Печень”: 0,865942

Корреляция между “Селезенка” и “Возраст”: 0,32137

Корреляция между “Печень” и “Возраст”: 0,266629

✓	f_x	=КОРРЕЛ(A2:A508;B2:B508)
D	E	
0,865942	Селезенка и Печень	
0,32137	Селезенка и возраст	
0,266629	Печень и возраст	

На основе полученных вычислений видим, что между “Селезенка” и “Печень” имеется сильная положительная линейная корреляция (близкая к 1), это говорит о том, что увеличение значений “Селезенка” ведет за собой увеличение значений и “Печень” и наоборот.

Между “Селезенка” и “Возраст”, есть слабая корреляция, при незначительном увеличении или уменьшении, сильных изменений не будет. Такая же ситуация с “Печень” и “Возраст”.

ЗАКЛЮЧЕНИЕ

На основе проведенного анализа, можно сделать вывод о том, что между “Селезенка” и “Печень” имеется сильная положительная линейная корреляция, а между “Селезенка” и “Возраст” и “Печень” и “Возраст” слабая линейная корреляция.

На мой непрофессиональный взгляд, сильная линейная корреляция между “Селезенка” и “Печень”, обусловлена наличием анатомического синдрома Гепатоспленомегалия, это синдром, характеризующийся одновременным увеличением селезенки и печени.

При анализе выявлено, что отрицательных линейных корреляций нет.