

УДК 311.313, 614.7

Заворотний Анатолий Анатольевич,
кандидат физ.-мат. наук, доцент, ЛГТУ, Липецк

Севостьянова Анастасия Романовна,
студент, ЛГТУ, Липецк

СТАТИСТИЧЕСКИЙ АНАЛИЗ ВЛИЯНИЯ ЗАБОЛЕВАНИЙ ОРГАНОВ ПИЩЕВАРЕНИЯ ОТ КОЛИЧЕСТВА ВЫБРОСОВ В СТОЧНЫЕ ВОДЫ В ЦФО ЗА 2022 ГОД

Аннотация: В работе представлен статистический анализ влияния общего количества сбрасываемых отходов в сточные воды на заболевания органов пищеварения в Центральном федеральном округе по данным за 2022 год. Выявлена четко прослеживаемая связь между указанными параметрами и даны общие рекомендации специалистам в области заболеваний пищеварения и экологии.

Ключевые слова: Статистический анализ, заболевания органов пищеварения, выбросы в сточные воды, регрессия, кривая Лоренца.

ВВЕДЕНИЕ. Проблема выбросов в сточные воды с каждым годом становится всё актуальнее в связи с тем, что увеличивается количество заводов и производств, что приводит к возрастанию степени загрязнения питьевой воды, несмотря на фильтры и разные методы очистки. Это подтверждается тем, что количество проб воды, не соответствующих санитарно-гигиеническим нормам, повышается с каждым годом [1]. Поэтому важно понимать, насколько сильно влияет загрязнение воды на нормальную жизнь, чтобы принять соответствующие меры по сокращению рисков заражения различными пищевыми болезнями или провоцированию их симптомов.

К указанной проблеме в последнее время обращались многие исследователи, поскольку она является весьма острой во всех регионах России. Негативное влияние загрязнения воды на развитие пищевых болезней изучалось в довольно большом ряде работ, например [2-6]. В этих исследованиях в основном затрагиваются только определенные города и области или сделан общий анализ влияния выбросов разного рода на здравоохранение.

АНАЛИЗ СТАТИСТИЧЕСКИХ ДАННЫХ. В данной статье целью было изучить по данным открытых официальных источников в Центральном федеральном округе указанную проблему и обратить внимание соответствующих специалистов данного региона на реальность существующих проблем в данном направлении.

В качестве входных контролируемых факторов были выбраны: 1) общее количество выбросов загрязняющих веществ, нефтепродукты (бензин, дизель), тяжёлые металлы (ртуть, цинк и иные); 2) общее количество зарегистрированных заболеваний органов пищеварения (воспаление пищевода, гастрит, гепатит и иные).

К неконтролируемым (то, на что в работе внимание не обращается, но что также влияет на пищевые заболевания) факторам можно отнести, например, следующие: состояние окружающей среды регионов в целом, их экономическое состояние, количество больниц, отделений для лечения болезней органов пищеварения, генетическая предрасположенность больных и т.д.

В качестве исходных данных были взяты показатели заболеваемости на 100 тыс. чел [7] и выбросов в тыс. тонн по регионам Центрального федерального округа России [8]. Эти данные сведены в таблице 1. Объём совокупности, т.е. количество областей ЦФО равно 18.



Таблица 1

Исходные данные

Регион	Выбросы, тыс. тонн [8]	Заболеваемость, на 100 000 чел [7]
	X	Y
Белгородская обл.	201,98	47495
Брянская обл.	62,83	12190
Владимирская обл.	97,6	30505
Воронежская обл.	220,71	150464
Ивановская обл.	86,49	11664
Калужская обл.	87,55	55916
Костромская обл.	1733,63	28437
Курская обл.	91,7	30127
Липецкая обл.	88,7	55979
г. Москва	7619,44	1811904
Московская обл.	1642,09	390437
Орловская обл.	52,75	15675
Рязанская обл.	118,94	24850
Смоленская обл.	59,41	14317
Тамбовская обл.	67,2	18502
Тверская обл.	2104,03	49235
Тульская обл.	166,64	27741
Ярославская обл.	153,29	14317

Количественные признаки X и Y являются непрерывными величинами (т. е. в пределах вариации могут принимать любые значения, отличающиеся друг от друга на сколь угодно малую величину) [9], поэтому для дальнейшего удобства анализа в таблице 2 представлена группировка исходных данных в виде интервальных вариационных рядов.

Таблица 2

Ранжированные исходные данные

Заболевания (100 тыс. чел)				Выбросы (тыс. тонн)			
x1	y1	PI (%)	QI (%)	x2	y2	P2 (%)	Q2 (%)
11664-360041	16	88,89	4,18	53-1517	14	77,78	4,20
360041-708418	1	94,44	16,21	1517-2981	3	94,44	16,23
708418-1056795	0	94,44	36,08	2981-4446	0	94,44	36,10
1056795-1405172	0	94,44	63,79	4446-5910	0	94,44	63,81
1405172-1811904	1	100,00	100,00	5910-7619	1	100,00	100,00

Распределение количества заболеваний органов пищеварения и выбросов в сточные воды по регионам ЦФО наглядно отображены в гистограммах на рисунке 1.



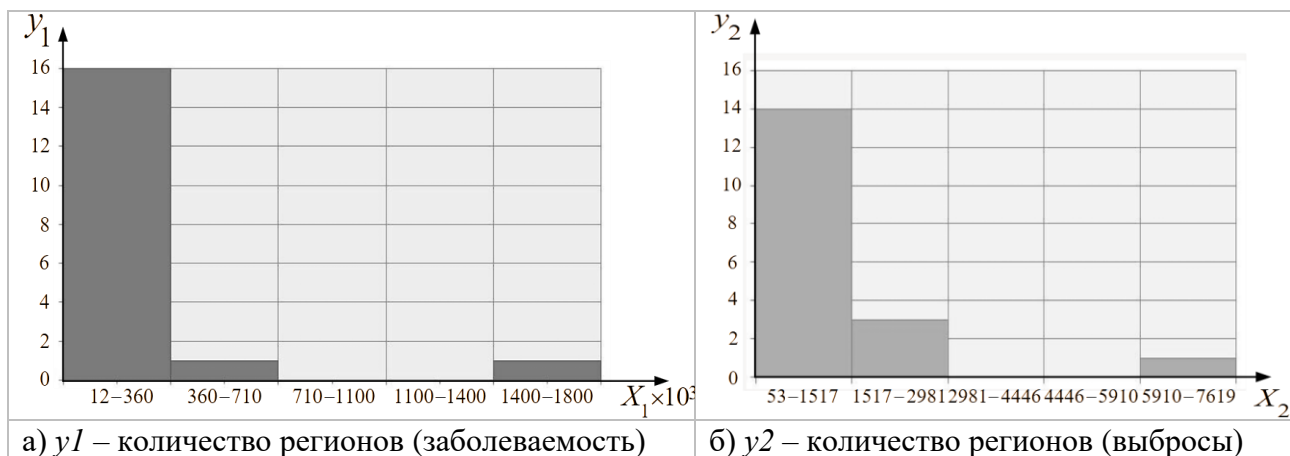


Рис. 1 – Гистограммы распределения контролируемых факторов по регионам

Гистограмма на рисунке 1,а показывает, что наиболее часто встречаются показатели заболеваний в интервале от 11664 до 360041. На рисунке 1,б видно, что в подавляющем большинстве регионов загрязнение сточных вод относительно невысоко на 100 тыс. чел., т.е. общее количество зарегистрированных в медицинских учреждениях, случаев соответствующих заболеваний относительно невысокое и лежит в интервале до 36%. Только в двух регионах ЦФО (г. Москва и Воронежская область) количество заболеваний высоко, что может косвенно объясняться плохой экологией регионов, а также высокой плотностью населения в них (до 1517 тыс. тонн).

Таким образом, в большинстве регионов ЦФО видно, что и количество заболеваний и количество выбросов сравнительно невелико и не находится в каких-либо критических показателях.

Для более наглядного качественного исследования неравномерности распределения заболеваний и выбросов по регионам на рисунке 2 представлены кривые Лоренца по каждому фактору. Для построения этих кривых были рассчитаны накопленные параметры P_i и x_i , $i=1,2$, представленные в таблице 2, при этом P_i – накопленный коэффициент для исследуемых факторов (выбросы, заболевания), а Q_i – накопленный коэффициент по областям в каждом случае.

Видно, что кривые для обоих факторов являются выгнутыми и расположены достаточно далеко над диагональю, которая представляет собой линию равномерного распределения показателей по регионам.

Таким образом, видна высокая неравномерность заболеваний ЖКТ и выбросов в сточные воды.

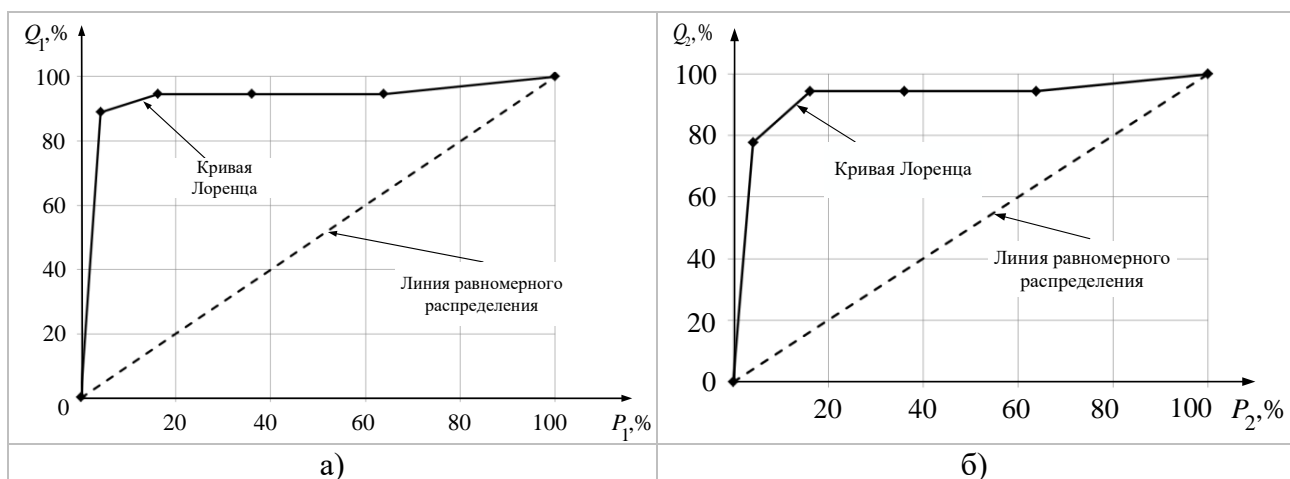


Рис. 2. Кривые Лоренца исследуемых факторов



Для оценки и численного подтверждения неравномерности распределения, был рассчитан особый параметр – коэффициент Джини (формула 1), который определяется как отношение площади S_2 фигуры между кривой Лоренца и линией равномерного распределения к площади S_1 под линией равномерного распределения:

$$G_s = \frac{S_0 - S_1}{S_1} = \frac{S_2}{S_1} \quad (1)$$

Значение коэффициентов Джини G является весьма высоким, в обоих случаях близким к единице (0,859 и 0,861 соответственно), что говорит о сильном отклонении показателей от равномерного. Значит, заболеваемость по областям распределена очень неравномерно, что может объясняться наличием или отсутствием в разных областях разных социальных и экономических факторов: больниц и отделений с разным профилем, наличие или отсутствие мест в больнице и персонала. Количество выбросов по областям распределено также неравномерно, что может объясняться наличием или отсутствием в разных областях разных загрязняющих факторов: заводов с разным профилем производства, наличие или отсутствие разных полезных ископаемых, распространённость водоемов и т.д.

Проведён расчёт числовых характеристик положения и рассеяния, которые охарактеризуют конкретный количественный разброс значений изучаемых генеральных совокупностей. Среднее арифметическое по заболеваемости составляет 444222,84 на 100 тыс. чел., а по выбросам – 1870,79 тыс. тонн. Стандартное отклонение 598916,89 и 2517,40, которое показывает сильный разброс значений рассматриваемых данных.

Для установления непосредственно взаимосвязи между представленными двумя генеральными совокупностями проведён корреляционный анализ.

Для того чтобы визуально оценить, есть ли линейная зависимость между генеральными совокупностями и есть ли эта зависимость в принципе, было построено корреляционное поле, которое отмечено точками синего цвета на рисунке 3.

По разбросу точек на рисунке 3 уже можно выдвинуть предположение о существующей линейной взаимосвязи.

Для аналитического выявления взаимосвязи вычислены корреляционные коэффициенты.

Расчет коэффициента корреляции по формуле

$$r = \frac{\sum_{i=1}^n (x_i - x_{cp})(y_i - y_{cp})}{\sqrt{\sum_{i=1}^n (x_i - x_{cp})^2 \cdot \sum_{i=1}^n (y_i - y_{cp})^2}} \quad (2)$$

показал значение в 0,77, что указывает на то, что около 77% изменений в заболеваемости органов пищеварения можно объяснить изменениями в сбросе сточных вод. Положительный знак указывает на то, что с увеличением сброса сточных вод число заболеваний пищеварительной системы также имеет тенденцию к увеличению. Коэффициент корреляции считается сильным, поскольку он находится в диапазоне от 0,5 до 0,9. Это говорит о том, что между двумя переменными существует значительная линейная зависимость. Коэффициент детерминации r^2 показывает долю влияния признака на связь, которая составляет 60 %.

Корреляционное отношение η оказалось равным 0,3, что согласно шкале Чеддока показывает наличие умеренной прямой связи между общим количеством выбросов в ЦФО и заболеваемостью по регионам.

Для того чтобы найти конкретную взаимосвязь проведён регрессионный анализ. Регрессия – это нахождение уравнения связи между исследуемыми параметрами [9]. В результате расчета получили следующий вид линейной регрессии:

$$y = b + ax = -23830,53 + 219,63x \quad (3)$$



Доверительные интервалы для коэффициентов уравнения (3) равны: $-250738 < a < 203076,99$; $103,23 < b < 336,04$. При этом коэффициент a , который является точкой пересечения линии регрессии с осью Y указывает на то, что в значении коэффициента существует некоторая неопределенность из-за широты интервала.

Коэффициент b является коэффициентом наклона, который представляет собой угол наклона линии регрессии. Включение в доверительный интервал для b нуля означает, что существует статистически значимая линейная зависимость между X и Y .

По полученному уравнению (3) найдены *утеор* для каждого значения x и построен график зависимости (рисунок 3), который позволяет визуально оценить разброс значений теоретических от эмпирических.

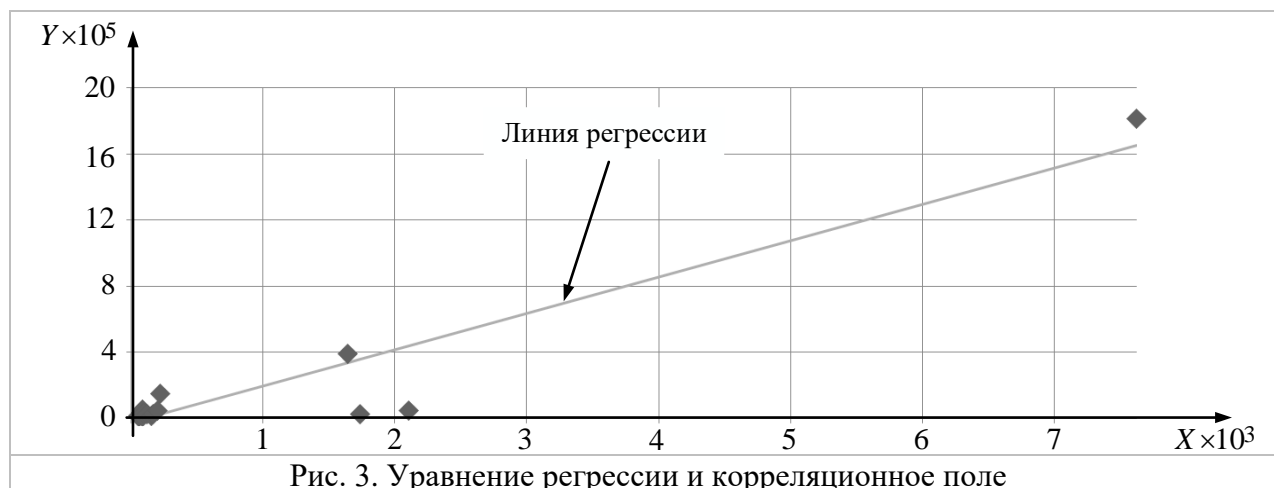


Рис. 3. Уравнение регрессии и корреляционное поле

Далее была проведена оценка изменчивости зависимой переменной Y (т.е. болезни органов пищеварения) от значения сбросов в сточные воды X . Для этого для уравнения (3) был рассчитан коэффициент детерминации $r_{reg}^2 = 0,8954$, который позволил численно объяснить влияние загрязнения сточных вод на болезни. Полученный результат позволяет заключить, что 89,54% вариации значения признака Y может быть объяснено вариацией значения фактора X , т.е. практически 90% заболеваний ЖКТ может быть объяснено выбросами в сточные воды (это, если не учитываются другие факторы, которые могут оказать влияние на болезни пищеварительной системы).

Воспользовавшись критерием Стьюдента, получен результат о том, что уравнение регрессии соответствует истинному значению в генеральной совокупности.

Информативно ли уравнение регрессии (3) или нет, позволяет определить критерий Фишера, который оказался равен 0,9, поэтому на уровне значимости в 5% можно утверждать, что регрессионная прямая достоверно отражает поведение рассматриваемых факторов.

ВЫВОД. В связи с полученными данными можно сделать вывод о том, что заболевания органов пищеварения и выбросы в сточные воды тесно связаны. Это означает, что достаточно большое влияние на здоровье человека оказывает экологическая ситуация, а также, в частности, методы и тщательность очищения воды на производствах, и количество разного рода отходов, которые имеют отрицательное влияние на морских, речных и озерных живых существ. В связи с этим необходимо более строго регулировать и следить за фильтрацией воды на производствах и фабриках и придумывать новые методы очищения грязной воды, которые смогут улучшить состояние здоровья граждан и минимизировать экологический ущерб.

Проведенный математико-статистический анализ позволяет предложить ряд рекомендаций для врачей:

1) быть более бдительными при диагностике и лечении пациентов с желудочно-кишечными симптомами, особенно в районах с высоким уровнем сброса сточных вод (эти районы явно определены в таблице 1);



2) собрать подробный экологический анамнез, включая опрос пациентов об их близости к местам сброса сточных вод и потенциальном воздействии загрязненных источников воды;

3) проводить обследование пациентов с желудочно-кишечными заболеваниями на наличие сопутствующих заболеваний, которые могут быть связаны с воздействием окружающей среды, таких как проблемы с дыханием или кожные заболевания.

При обследовании пациентов с желудочно-кишечными заболеваниями нужно учитывать потенциальную роль факторов окружающей среды, включая сброс сточных вод, в развитии и прогрессировании заболевания.

Проводить целенаправленное тестирование на инфекционные агенты и токсины, которые могут быть связаны со сбросом сточных вод, такие как паразитарные инфекции или бактериальные токсины.

Список литературы:

1. Сайт 76.ru [Электронный ресурс] – Ярославль 2013. – Ярославская вода – самая грязная в ЦФО – Электрон. дан. – Режим доступа: <https://76.ru/text/gorod/2014/12/24/61642151/> свободный. (Дата обращения 21.06.2024г.)

2. **Иванов С.В.** Влияние качества воды на здоровье населения / С.В. Иванов, Э.Л. Федорова, Э.Э. Темиров // Международный журнал прикладных и фундаментальных исследований. – 2017. – № 3. – с. 186-189.

3. **Степанов Н.А.** Характеристика влияния качественного состава питьевой воды на здоровье человека / Н.А. Степанов, Е.И. Заводова // Медицина труда и экология человека. – 2015. – № 3. – с. 207-212.

4. **Сазонова О.В.** Анализ риска здоровью населения, обусловленного загрязнением питьевой воды (опыт Самарской области) / О.В. Сазонова, А.К. Сергеев, Л.В. Чупахина и др. // Анализ риска здоровью. – 2021. – № 2. – с. 41-51. – DOI: 10.21668/health.risk/2021.2.04

5. **Клейн С.В.** Приоритетные факторы риска питьевой воды систем централизованного питьевого водоснабжения, формирующие негативные тенденции в состоянии здоровья населения / С.В. Клейн, С.А. Вековщина // Анализ риска здоровью. – 2020. – № 3. – с. 49-60. – DOI: 10.21668/health.risk/2020.3.06

6. **Соловьев Н.А.** Источники химического загрязнения и их влияние на гидросферу и здоровье человека / Н.А. Соловьев, И.Г. Широкова // Царскосельские чтения. – 2011. – № 15. – с. 402-409.

7. Заболеваемость населения по субъектам Российской Федерации [Электронный ресурс] – Электрон. дан. – Режим доступа: <https://statprivat.ru/zdo?r=5> свободный. (Дата обращения 21.06.2024г.)

8. Официальный сайт Федеральной службы государственной статистики / Окружающая среда [Электронный ресурс] – Электрон. дан. – Режим доступа: <https://rosstat.gov.ru/folder/11194> свободный. (Дата обращения 21.06.2024г.)

9. Теория статистики: учебник / под ред. Г. Л. Громыко. – М.: ИНФРА-М, 2005. – 476 с.

