

Статистическое исследование продовольственных расходов в 2022 г. в Российской Федерации

*Константинов Александр
НИУ ВШЭ, Факультет экономических наук*

1. Введение

Потребительские расходы на продовольствие представляют собой значимую часть бюджета большинства домохозяйств, особенно характерно для небогатых слоев населения. Важно изучить, как изменяется доля этих расходов в доходах населения, чтобы выявить, насколько велика нагрузка расходов на основные потребности по отношению к доходам. Анализ доли продовольственных расходов позволяет оценить уровень финансовой устойчивости и благосостояния населения, а также способствует разработке мер социальной поддержки.

Цель данного исследования: рассмотреть продовольственные расходы населения России в динамике, выявить тенденции, а также оценить влияние состава домохозяйств на долю продовольственных расходов.

Будет рассмотрена динамика средних и медианных показателей доходов домохозяйств в 2022 г., будет проанализирована связь между количеством лиц, состоящих в домохозяйстве, и долей продовольственных расходов в доходах, а также влияние расположения (в сельской или городской местности) на характер данной связи.

Данное исследование основано на данных Росстата, полученных с официального портала <https://obdx.gks.ru/> за 2022 год, который является последним доступным периодом на момент проведения анализа. Эти данные содержат информацию о доходах и расходах домохозяйств по регионам России, что позволяет детально рассмотреть продовольственные расходы населения в контексте различных экономических и социальных факторов.

В силу высокого статуса Росстата доверие к данным очень высокое, более того, в файле имеется коэффициент взвешивания, который позволяет корректировать поиск среднего и статистик, получать менее смещенные оценки. Пропущенных значений нет, однако есть наблюдения, которые можем квалифицировать как выбросы (например, при расчете отношения расходов на продовольствие к доходам, у некоторых домохозяйств этот показатель превышает не только 100 процентов, но 1000 процентов. Очевидно, что это наблюдение является выбросом, более того, при учете выброса при пересчете средних значений по регионам, регионы демонстрируют 1000 и более процентов расходов в доходах.

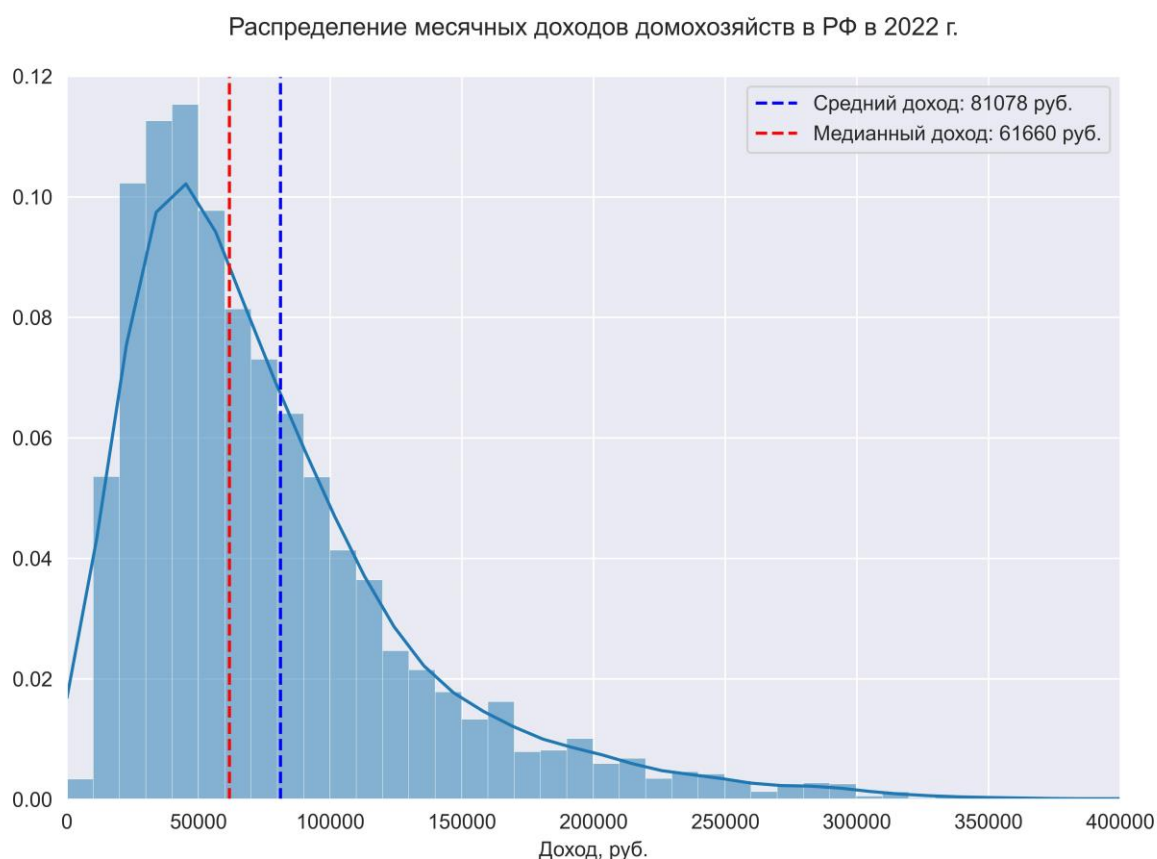
Можем обработать выбросы либо установив некоторый максимум, выше которого показатель не может оказаться, либо заменив его на более реалистичный с помощью метода ближайших соседей.

Общее количество наблюдений: 193256

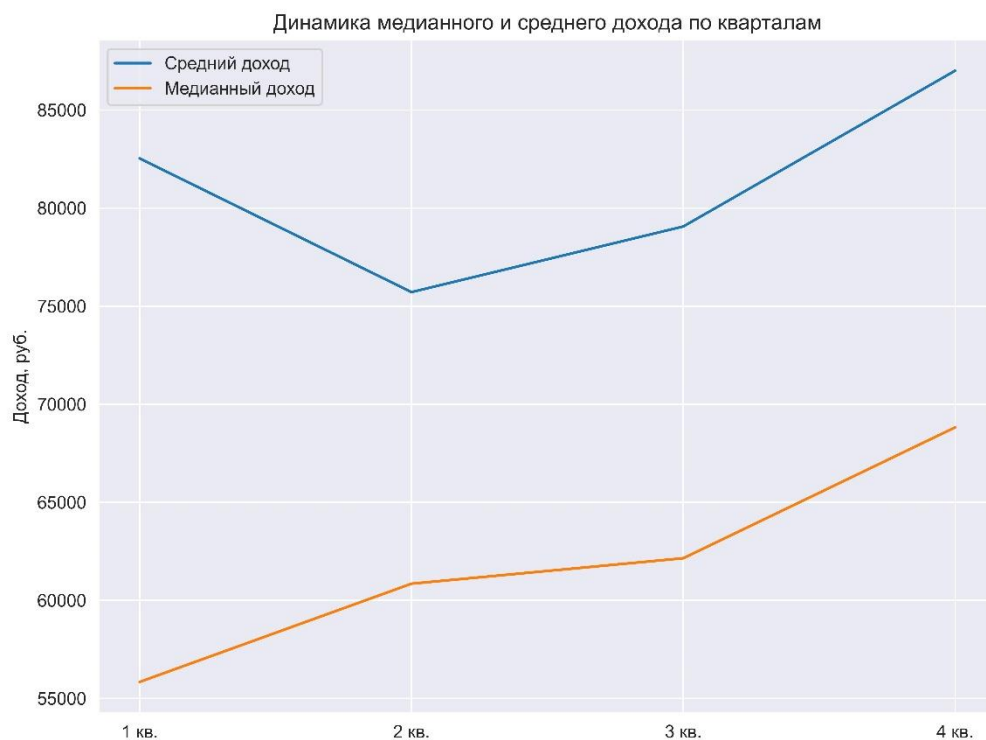
Количество наблюдений для каждого квартала: 48314.

2. Предварительный анализ доли распределения доходов домохозяйств в РФ в 2022 г.

Продовольственные расходы представляют собой важный аспект финансового благосостояния домохозяйств, отражая не только уровень жизни, но и потребительские предпочтения населения. Цель анализа — выявить динамику и особенности распределения продовольственных расходов в зависимости от различных факторов, а также оценить влияние экономических условий на эти расходы.



Видим асимметрию в распределении доходов и довольно весомый правый хвост распределения, который перетягивает среднее значение вправо.



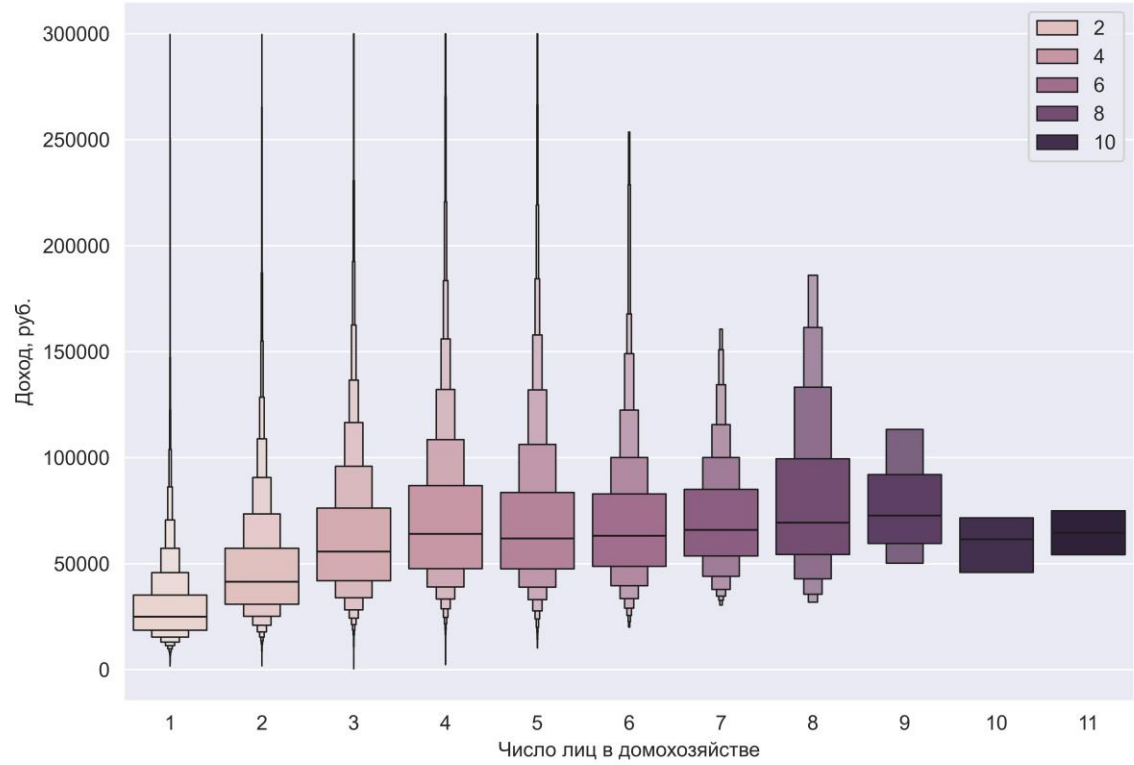
Средний доход падал во 2 квартале 2022 г., однако взял курс на повышение в 3 и 4 кварталах. Медианная зарплата росла на протяжении 2022 г. Это может быть последствием инфляции, а не роста реальных доходов населения.

Посмотрим на разницу в доходах между сельской местностью и городской местностью. Визуально можем оценить разность между медианами и стандартным отклонением в двух группах (выбросы не показаны, так как для некоторых домохозяйств демонстрируют экстремально большие значения).

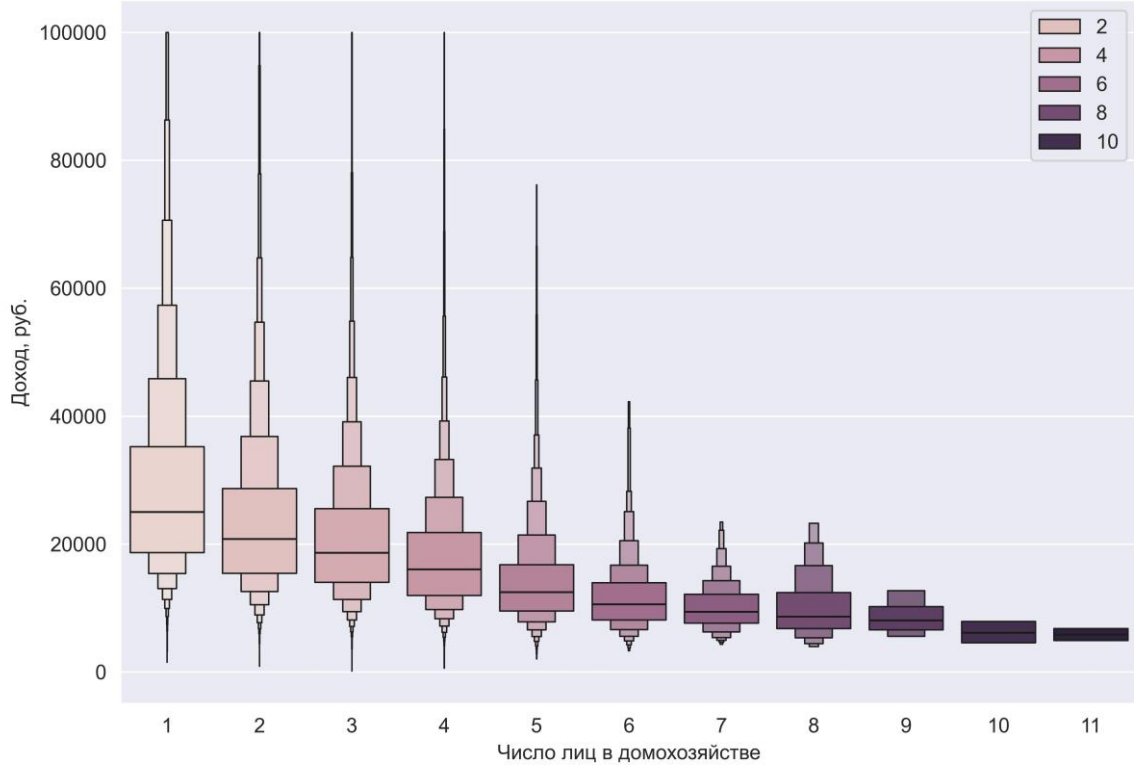


Укажем на другой важный принцип и его влияние на доход и среднедушевой доход домохозяйства – количество лиц в домохозяйстве.

Разница в доходах в зависимости от количества лиц в домохозяйстве



Разница в среднедушевом доходе в зависимости от количества лиц в домохозяйстве



Визуально можем оценить, что рост числа лиц в домохозяйстве от 1 до 4 сопровождается ростом медианного дохода в этих группах, в то время как медиана среднедушевого дохода падает с каждым новым членом семьи.

Выдвинем ряд гипотез для статистической проверки:

Гипотеза 1.

H_0 : Рост среднего дохода сопровождался ростом доли расходов на продовольственные товары.

А значит, наблюдая рост цен на продовольственные товары, мы можем сказать, что рост среднего дохода между 1 кв. и 4 кв. 2022 г. не способствовал большим тратам на товары роскоши, развлечения и т.п.

H_a : Рост среднего дохода не сопровождался ростом доли расходов на продовольственные товары.

Гипотеза 2.

H_0 : Доля продовольственных доходов выше для домохозяйств, в которых число лиц в домохозяйстве больше 2.

H_a : Доля продовольственных доходов не выше для домохозяйств, в которых число лиц в домохозяйстве больше 2.

Такое предположение может возникнуть в силу того, что домохозяйства с 3 и более лицами обычно состоят из работающих родителей и ребенка, а потому им приходится больше средств тратить на продовольствие.

Гипотеза 3.

Если гипотеза 2 будет подтверждена, можем выдвинуть следующую гипотезу.

H_0 : Тот факт, что домохозяйство располагается в сельской местности усиливает влияние количества членов семьи на долю продовольственных доходов.

H_a : Тот факт, что домохозяйство располагается в сельской местности не усиливает влияние количества членов семьи на долю продовольственных доходов.

3. Проверка гипотез

Гипотеза 1.

H_0 : Рост среднего дохода сопровождался ростом доли расходов на продовольственные товары.

H_a : Рост среднего дохода не сопровождался ростом доли расходов на продовольственные товары.

Переведем на математический язык.

H_0 : Отвергаются гипотезы $\mu_4^{\text{доход}} - \mu_1^{\text{доход}} \leq 0$ и $\mu_4^{\text{прод}} - \mu_1^{\text{прод}} \leq 0$.

H_a : Неверно хотя бы одно из условий в H_0 .

Где $\mu_i^{\text{доход}}$ – средний доход в i -ый квартал, $\mu_i^{\text{прод}}$ – среднее значение доли продовольственных расходов в доходах в i -ый квартал. Проверку делаем t-тестом, поскольку среднее значение асимптотически распределено равномерно, неизвестная дисперсия имеет распределение хи-квадрат.

Нулевая гипотеза принимается:

- разность средних по доходам действительно статистически значимо больше нуля с $p\text{value} = 9.7 \cdot 10^{-137}$ (можем отвергнуть на 1-процентном уровне значимости);
- разность средних генеральных совокупностей распределения долей продовольственных расходов больше 0 с $p\text{value} = 1.79 \cdot 10^{-6}$ (можем отвергнуть на 1-процентном уровне значимости).

Гипотеза 2.

H_0 : Доля продовольственных доходов выше для домохозяйств, в которых число лиц в домохозяйстве больше 2.

H_a : Доля продовольственных доходов не выше для домохозяйств, в которых число лиц в домохозяйстве больше 2.

$H_0: \mu^{\text{не более 2}} - \mu^{\text{более 2}} \leq 0$.

$H_a: \mu^{\text{не более 2}} - \mu^{\text{более 2}} > 0$.

Где $\mu^{\text{не более 2}}$ – среднее значение доли продовольственных расходов среди домохозяйств с числом лиц не более 2 и $\mu^{\text{более 2}}$ — с числом лиц более 2, соответственно.

Гипотеза отвергается ($p\text{value}$ при вычислениях не отличимо от 0), более того, мы принимаем гипотезу о том, что $\mu^{\text{не более 2}} - \mu^{\text{более 2}} < 0$ ($p\text{value}$ при вычислениях не отличимо от 1). Значит, наблюдается эффект, обратный

предполагаемому: доля продовольственных расходов возрастает для малых домохозяйств. Можно выдвинуть гипотезу о том, что это происходит за счет домохозяйств, состоящих из 1 человека: индивидуально, в силу низкого уровня доходов, вынужден тратить на продовольствие.

Гипотеза 3.

Если гипотеза 2 будет подтверждена, можем выдвинуть следующую гипотезу.

H_0 : Тот факт, что домохозяйство располагается в сельской местности усиливает влияние количества членов семьи на долю продовольственных доходов.

H_a : Тот факт, что домохозяйство располагается в сельской местности не усиливает влияние количества членов семьи на долю продовольственных доходов.

H_0 : $PROD_RATIO = \hat{\beta}_1 + \hat{\beta}_2 \cdot FAMSIZЕ + \hat{\beta}_3 \cdot FAMSIZЕ \cdot TOWN + \hat{\beta}_4 \cdot TOWN$, $\hat{\beta}_3 < 0$ при $\hat{\beta}_2 \geq 0$ или $\hat{\beta}_3 > 0$ при $\hat{\beta}_2 \leq 0$.

H_a : Неверна нулевая гипотеза.

Где $PROD_RATIO$ – предсказанное значение доли продовольственных расходов, $FAMSIZЕ$ – размер семьи, $TOWN$ – индикатор того, что домохозяйство находится в городской среде.

Результат:

	coef	std err	t	P> t	[0.025	0.975]
CONST	0.4130	0.002	253.441	0.000	0.410	0.416
FAM_SIZE	-0.0154	0.001	-26.503	0.000	-0.017	-0.014
TOWN	-0.0121	0.002	-6.091	0.000	-0.016	-0.008
FAM_SIZE*TOWN	-0.0061	0.001	-8.329	0.000	-0.008	-0.005

Видим, что размер все коэффициенты статистически значимы, как и регрессия в целом.

Базовый уровень трат на продовольствие для домохозяйства – 41,3 % от дохода (доверительный интервал [41.0 %, 41.6 %]), каждый новый член в среднем снижает долю трат на продовольствие на 1,54 % от дохода (доверительный интервал [–1.7 %, –1.4 %]), присутствие в городе снижает на 1,21 % (доверительный интервал [–1.6 %, –0.8 %]), а также проживание в городе усиливает отрицательный эффект большого размера семьи (доверительный интервал [–0.8 %, –0.5 %]).

Таким образом, гипотеза 3 не подтверждается, поскольку мы получаем: домохозяйство с большим количеством членов в среднем меньше тратит на продовольствие в долях от дохода, однако в сельской местности наблюдается обратный эффект.

4. Доверительные интервалы и корреляция



Коэффициент корреляции = 0.557. Тестом проверяется на статистическую значимость с $p\text{-value} = 0,00716$.

Построим доверительные интервалы для дисперсий месячных расходов на продовольствие и месячных доходов домохозяйств.

Доверительный интервал для месячных доходов домохозяйств: [1268810824.47, 1284912344.80], если извлечь квадратный корень из границ, получим [35620.37, 35845.67].

Доверительный интервалы для дисперсий месячных расходов на продовольствие: [73040037.95, 73966933.93], если извлечь квадратный корень из границ, получим [8546.35, 8600.40]. Разброс для стандартного отклонения не так велик по сравнению с суммами, которые в месяц тратятся на продовольствие. Плотность точек мы можем видеть на точечном графике.

5. Заключение

Проведенное исследование позволило подтвердить ряд важных закономерностей, связанных с долей продовольственных расходов в доходах домохозяйств в зависимости от уровня дохода, размера семьи и места проживания. В ходе проверки первой гипотезы мы установили, что рост среднего дохода в 2022 г. не сопровождался ростом доли расходов на продовольственные товары в доходах. Более того, разность средних по доходам и разность долей продовольственных расходов оказались статистически значимо положительными, что говорит о тенденции снижения доли доходов, потраченных на продовольствие, на начало и конец 2022 г.

Второе предположение, о том, что большие домохозяйства тратят большую долю дохода на продовольствие, также было опровергнуто. Причина этого эффекта может заключаться в том, что небольшие домохозяйства, особенно состоящие из одного человека, при низком уровне доходов вынуждены направлять большую часть средств на продовольствие.

Наконец, третья гипотеза о влиянии сельской среды на рост доли продовольственных расходов при увеличении численности семьи также была опровергнута. Мы обнаружили, что в сельской местности, напротив, больший размер семьи уменьшает долю продовольственных расходов, хотя и менее выражено, чем в городской среде. Основные эффекты в модели показали, что доля продовольственных расходов в доходах домохозяйств снижается с ростом числа членов семьи и наличием домохозяйства в городской среде.

Таким образом, полученные результаты позволяют сделать вывод о том, что доля доходов, потраченных на продовольствие, вступает в сложную зависимость с размером домохозяйства и расположением домохозяйства в городской или сельской местности. Зависимости можем выразить в виде причинных цепей.

- Размер домохозяйства \Rightarrow Доля доходов, потраченных на продовольствие
- Местоположение \Rightarrow Доля доходов, потраченных на продовольствие
- Местоположение \Rightarrow Размер домохозяйства \Rightarrow Доля доходов, потраченных на продовольствие