

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО
ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ЭКОНОМИКИ И
ФИНАНСОВ

ФАКУЛЬТЕТ СТАТИСТИКИ, УЧЁТА И ЭКОНОМИЧЕСКОГО АНАЛИЗА

КАФЕДРА СТАТИСТИКИ И ЭКОНОМЕТРИКИ

Курсовая работа
по общей теории статистики
на тему
«Зависимость внебрачной рождаемости от доли незамужних женщин
в России»

Выполнила студентка 359 группы Колесниченко Ксения Сергеевна
Научный руководитель д.э.н., профессор, Васильева Эвелина Карловна

Санкт-Петербург
2007 год

Введение	3
1 Показатели тесноты связи.	5
2 Корреляция	8
3 Расчёт показателей вариации	12
Средняя арифметическая.	14
Стандартная ошибка.	15
Медиана.	15
Мода.	16
Стандартное отклонение.	16
Дисперсия выборки.	17
Эксцесс.	17
Асимметричность.	19
Размах вариации	20
Минимум (Наименьший).	20
Максимум (Наибольший).	21
Сумма.	21
Счёт.	21
Уровень надёжности.	21
3.1 Динамический ряд	23
4 Заключение.....	25
5 Список использованной литературы	26

Введение

Я выбрала тему «Зависимость рождаемости от брачного состояния женщины» потому что считаю, что для нашей страны, как и для многих других развитых стран мира, сейчас очень актуальны вопросы о методах повышения уровня рождаемости. Существует много факторов рождаемости: политические, социально-экономические, правовые нормы, принятые в государстве, историко-культурные, демографические, этнические, образовательный уровень населения и другие.

В своей работе я хочу рассмотреть один из демографических факторов – брачное состояние женщины.

В наше время мам, рожающих детей, условно можно разделить на три группы:

- женщины, состоящие в зарегистрированном браке;
- женщины, состоящие в незарегистрированном браке;
- матери-одиночки.

Мне очень интересно рассмотреть такое явление, как внебрачная рождаемость. Считаю, что уровень внебрачной рождаемости напрямую зависит от брачного состояния женщины.

В своей работе я попробую доказать наличие прямой зависимости между брачным состоянием женщины и уровнем внебрачной рождаемости. Также я хочу показать, что доля внебрачной рождаемости растёт с каждым годом.

В моей работе источниками данных о рождаемости были взяты из Демографического ежегодника России (2006 года), данные о соотношении численности замужних и незамужних женщин – с официального сайта Всероссийской переписи населения 2002 года.

К сожалению, всех имеющихся данных не достаточно для того, чтобы провести полный анализ данных.

Так, например, данные о рождаемости за период с 1960 по 1980 годы приведены с 5-летним интервалом, нет данных о фактическом брачном

состоянии рожениц за 1960–1969 и 1981–1987 годы. Также отсутствуют данные о рождаемости с разделением по субъектам Федерации. Отсутствуют данные и о фактическом брачном состоянии женщин, рожавших в 2004-2005 годах с подразделением на субъекты Федерации. Отсутствуют данные и о разделении женщин, рожавших в период с 1960 по 2003 годы с разделением на городское и сельское население. Также отсутствуют данные о доле незамужних женщин с разделением на городское и сельское население по субъектам Федерации.

Кроме того, расчёт корреляции между долей незамужних женщин и долей внебрачной рождаемости придётся сделать некорректно. Так, данные о рождаемости приведены за 2004 и 2005 годы, а данные о доле незамужних женщин – по данным Всероссийской переписи 2002 года. Это вынужденная мера, т.к. наиболее точные данные о доле незамужних женщин можно получить только по данным переписей. Так как серьёзных сдвигов в соотношении замужних и незамужних женщин не было, я использовала эти данные.

1 Показатели тесноты связи.

Прежде всего рассчитаем показатели тесноты связи уровней внебрачной рождаемости в городских и сельских местностях.

Имеется таблица с данными о рождаемости у женщин, не состоящих в браке за 2004 и 2005 годы (см. Приложение 1).

Для погашения случайных колебаний в качестве расчётных данных будут использоваться усредненные данные. Данные осредняются по средней арифметической простой (см. Приложение 2)

$$\bar{x} = \frac{\sum x_i}{n} \quad (1.1)$$

Так же по формуле средней арифметической простой осредняются и численность и доля детей родившихся у замужних и незамужних женщин, численность и доля детей, родившихся по типу местности – городская или сельская.

В связи с тем, что отсутствуют данные о рождаемости в республике Ингушетия за 2004 год, в разделе с осреднёнными данными будут использоваться данные за 2005 год.

Определим ассоциацию, контингенцию, коэффициент взаимной сопряженности Пирсона, коэффициент Чупрова для выявления данных о зависимости между рождаемостью, брачным состоянием женщины, и типом местности, в которой проживает женщина.

Коэффициент ассоциации.

Условно обозначим

$$K_a = \frac{ad - bc}{ad + bc}; \quad (1.2)$$

$$K_a \approx \frac{775 \text{ тыс. чел.} \cdot 146 \text{ тыс. чел.} - 285 \text{ тыс. чел.} \cdot 306 \text{ тыс. чел.}}{775 \text{ тыс. чел.} \cdot 146 \text{ тыс. чел.} + 285 \text{ тыс. чел.} \cdot 306 \text{ тыс. чел.}} = 0,1303$$

Где а – замужние женщины, родившие детей в 2004-2005 годах, проживающие в городской местности (775113 человек)

b – замужние женщины, родившие детей в 2004-2005 годах, проживающие в сельской местности (285203 человек)

c – незамужние женщины, родившие детей в 2004-2005 годах, проживающие в городской местности (305745 человек)

d – незамужние женщины, родившие детей в 2004-2005 годах, проживающие в сельской местности (146210 человек)

Коэффициент контингенции

$$K_k = \frac{ad - bc}{\sqrt{(a+b)(b+d)(a+c)(c+d)}}; \quad (1.3)$$

$$K_k \approx \frac{775 \text{ тыс. чел.} \cdot 146 \text{ тыс. чел.} - 285 \text{ тыс. чел.} \cdot 306 \text{ тыс. чел.}}{\sqrt{1060 \text{ тыс. чел.} \cdot 431 \text{ тыс. чел.} \cdot 1081 \text{ тыс. чел.} \cdot 452 \text{ тыс. чел.}}} = 0,0553$$

Коэффициент взаимной сопряженности Пирсона

$$K_{\pi} = \sqrt{\frac{\varphi^2}{1 + \varphi^2}} \quad (1.4)$$

$$1 + \varphi^2 = \frac{\frac{775113^2}{1080857} + \frac{285203^2}{431413}}{1060316} + \frac{\frac{305745^2}{1080857} + \frac{146210^2}{431413}}{451955} = 0,702 + 0,301 = 1,003$$

$$\varphi^2 = 1,003 - 1 = 0,003$$

$$K_{\pi} = \sqrt{\frac{0,003}{1,003}} = \sqrt{0,00299} = 0,0547$$

Рассчитаем коэффициент Чупрова.

$$K_{\chi} = \sqrt{\frac{\varphi^2}{\sqrt{(K_1 - 1)(K_2 - 1)}}} = \sqrt{\frac{0,003}{\sqrt{(2 - 1)(2 - 1)}}} = \sqrt{\frac{0,003}{1}} = 0,0548 \quad (2.5)$$

Все показатели $<< 1$, что говорит об отсутствии связей.

Таким образом, мы видим, что в 2004-2005гг среднем по Российской Федерации брачное состояние матери не зависит от того, в каком типе местности она проживает.

Посчитаем соотношение внебрачных рождений у городского и сельского населения по субъектам Федерации, убрав из анализа города

Москва и Санкт-Петербург (т.к. в данных субъектах Федерации нет сельских местностей) и Усть-Ордынский Бурятский автономный округ (т.к. в данном субъекте Федерации нет городской местности), см. Приложение 3.

Используя инструмент «Описательная статистика» функции «Анализ данных» программы MS Excel, получим основные характеристики для данного дискретного ряда:

$\bar{x}=0,81$, что говорит о том, что в среднем уровень соотношения внебрачной рождаемости в городах на 19% ниже, чем в сельских местностях.

$As=1,78$, а это говорит о сильной правосторонней асимметрии.

Рассчитав самостоятельно σ_{as} , получаем значение 0,028, мы видим, что $\frac{As}{\sigma_{as}} >> 2$, а это значит, что распределение нельзя считать нормальным.

Коэффициент вариации для данного ряда составил 0,017.

Построим линейчатую диаграмму (См. приложение 4, диаграмма 1).

Из полученной диаграммы видно, что в России нет определённых типов субъектов, из которых можно было бы выделить несколько групп с одинаковым уровнем соотношений внебрачной рождаемости в городских и сельских местностях.

Мы видим, что в основной массе уровень внебрачной рождаемости в городах до 30 процентов ниже, чем в сельских местностях. Это обуславливается другими демографическими факторами рождаемости такими, как соотношение полов в составе населения.

Тем не менее, видно резко выделяющиеся регионы из основной массы регионов – Кабардино-Балкарскую республику, Мурманскую область, Чувашскую республику, Карачаево-Черкесскую республику, республики Татарстан, Ингушетия, в которых внебрачная рождаемость в городах значительно выше внебрачной рождаемости в сельской местности. Как мы видим, это регионы, в которых силен этнический фактор рождаемости, в этих регионах исторически осуждалась внебрачная рождаемость. Мы видим, что в

городской местности процесс изменения отношения общественности к внебрачной рождаемости идёт быстрее, чем в сельской местности.

2 Корреляция

По данным Всероссийской переписи населения 2002 года имеются данные о количестве замужних женщин различных возрастных групп (см. Приложение 5). Данные подразделяются по субъектам федерации. РФ.

<i>Незамужние</i>	<i>Женщины</i>	<i>Женщины</i>	<i>Женщины</i>
<i>женщины</i>	<i>= 16 – 49 лет</i>	<i>– 16 – 49 лет</i>	<i>– 16 – 49 лет,</i>
<i>16 – 49 лет</i>	<i>состоящие</i>	<i>не состоящие</i>	<i>не указавшие</i>
	<i>в официальном</i>	<i>в официальном</i>	<i>брачное</i>
	<i>браке</i>	<i>браке</i>	<i>состояние</i>

Доля незамужних женщин в возрасте 16-49 лет рассчитывается как соотношение незамужних женщин 16-49 лет в данном субъекте Федерации к общей численности женщин 16-49 лет в данном субъекте Федерации. Данные представлены в Приложении 6

Так как в 2004 году произошло объединение Пермской области и Коми-Пермяцкого автономного округа в единый регион – Пермский край, данные о доле незамужних женщин рассчитываются по формуле средней арифметической простой

$$\begin{aligned}
 & \text{Доля незамужних женщин } 16-49 \text{ лет}_{\text{Пермский край}} = \\
 & = \frac{\text{Доля незамужних женщин } 16-49 \text{ лет}_{\text{Коми-Пермяцкий автономный округ}} + \text{Доля незамужних женщин } 16-49 \text{ лет}_{\text{Пермская область}}}{2} \\
 & = \frac{0,53 + 0,55}{2} = 0,54
 \end{aligned}$$

Рассчитаем парную корреляцию, используя функцию «КОРРЕЛ» программы MS Excel. В качестве признака-фактора используем долю незамужних женщин в возрасте 16-49 лет в субъектах Федерации. В качестве признака-результата используем данные об уровне внебрачной рождаемости в субъекте Федерации.

В таблице 2.1 представлены значения коэффициента парной корреляции.

Таблица 2.1

Значения коэффициента парной корреляции по федеральным округам России

Федеральный округ Российской Федерации	Значение коэффициента парной корреляции
Центральный Федеральный округ	0,819
Северо-Западный Федеральный округ	0,244
- без учёта Ненецкого автономного округа	0,668
Южный Федеральный округ	-0,471
- без учёта республики Ингушетия	0,328
Приволжский федеральный округ	0,908
Уральский федеральный округ	0,996
Сибирский федеральный округ	0,738
Дальневосточный федеральный округ	0,361
- без учёта Чукотского автономного округа	0,521
Вся Россия	0,567
-без учёта исключённых регионов	0,767

Прежде всего, мы видим, что коэффициент парной корреляции в целом по России имеет значение 0,567, а это говорит о высокой связи между признаком-фактором и признаком-результатом.

Убрав из анализа 4 субъекта Федерации, мы получили коэффициент парной корреляции 0,767, что говорит об очень высокой зависимости между долей незамужних женщин и долей внебрачных рождений.

Прежде всего, хочется отметить, что применение статистико-математической модели является необходимым инструментом для анализа тенденций и факторов рождаемости в нашей стране.

Высокая зависимость между долей незамужних женщин и долей внебрачных рождений говорит о том, что в наше время рождения вне официально зарегистрированного брака и рождения детей матерями-одиночками, становятся нормой жизни, что, в свою очередь говорит об эмансипации женщин, о том, что в наше время женщина становится экономически независимой, и, наверное, самое главное, нет психологического осуждения рождения детей вне брака со стороны общества.

К сожалению, полноценно рассмотреть 3 типа матерей (замужних, состоящих в незарегистрированном браке и матерей-одиночек) невозможно

без проведения специальных выборочных обследований. Несомненно, что подход к прогнозированию рождаемости на основе более дифференцированных данных даст не только лучшие результаты, но и на государственном уровне будет проще выбрать именно те необходимые меры, которые нужно предпринять для повышения рождаемости в конкретных группах потенциальных матерей, направленные на реализацию желаемого числа детей.

Если рассмотреть данные о парной корреляции на федеральном уровне (см. Таблицу 2.1), можно выделить четыре федеральных округа с высокими показателями парной корреляции (более 0,738) – Центральный, Приволжский, Уральский и Сибирский. В таких федеральных округах, как Приволжский и Уральский, значения коэффициента парной корреляции очень высоки. Это говорит, прежде всего, о высокой однородности составов субъектов Федерации. Кроме того, стоит отметить, что в данных федеральных округах очень высока доля славянских народов.

Меня очень удивили значения коэффициента парной корреляции в Северо-Западном и Дальневосточном федеральных округах – 0,244 и 0,361 соответственно. Убрав из анализа Ненецкий Чукотский автономные округа, значения этого коэффициента значительно увеличилось – 0,668 и 0,521 соответственно. Это говорит о том, что в этих регионах очень сильны этнические факторы рождаемости.

Очень интересно значение коэффициента парной корреляции в Южном Федеральном округе – он оказался отрицательным. Если не проводить дополнительного анализа, то получится, что в Южном федеральном округе при росте доли незамужних женщин, уменьшается доля внебрачной рождаемости. Убрав из анализа один регион – республику Ингушетию, коэффициент стал положительным, хотя его значение не очень велико – 0,328. Это говорит о высокой неоднородности состава данного федерального округа. И это действительно так – в данном регионе находятся и субъекты Федерации с преимущественно славянским населением, и субъекты

Федерации, исконно населённые различными народами, в этом федеральном округе много и субъектов Федерации с преобладанием христианских народов, и субъектов Федерации с преобладанием мусульманских народов. В Ингушетии же, в которой недавно шли военные действия, произошёл дисбаланс между долей мужского и женского населения репродуктивного возраста. Традиционно в Ингушетии осуждается внебрачная рождаемость, из-за чего и получилась обратная зависимость между долей незамужних женщин и долей внебрачной рождаемости.

Из анализа корреляционных связей на федеральном уровне мы видим, что и федеральные округа сильно различаются по своей однородности. Таким образом, для повышения эффективности программ, направленных на повышение уровня рождаемости, их необходимо составлять для каждого федерального округа в отдельности. А для некоторых федеральных округов необходимо разрабатывать программы на уровне субъектов Федерации.

Здесь так же требуются специальные выборочные обследования, направленные на выявление факторов, препятствующих рождению желаемого числа детей.

3 Расчёт показателей вариации

Проведём анализ вариации для городского и сельского населения. Для женщин из городов уберём Усть-Ордынский Бурятский автономный округ, так в нём отсутствует городское население, а для анализа вариации сельского населения уберём Москву и Санкт-Петербург, так в них отсутствует сельское население. Используя инструмент «Описательная статистика» функции «Анализ данных» программы MS Excel, мы получаем основные характеристики данных вариационных рядов. Так же самостоятельно рассчитаем эти же показатели, предварительно сгруппировав их.

В таблице 3.3 приведены показатели вариации, рассчитанные компьютером, а в скобках – рассчитанные самостоятельно.

Все расчёты представлены в Приложениях 7 и 8.

Для определения количества групп воспользуемся формулой Стерджесса.

$$k = 1 + 3,332 \lg N; \quad (3.1)$$

где k – число групп,

N – число единиц совокупности.

Так, для городской местности получим число групп:

$$k = 1 + 3,332 \cdot \lg 87 = 1 + 3,332 \cdot 1,940 = 1 + 6,46408 = 7,46408 \approx 8 (\text{групп})$$

Для сельской местности получим число групп:

$$k = 1 + 3,332 \cdot \lg 86 = 1 + 3,332 \cdot 1,934 = 1 + 6,444 = 7,444 \approx 7 (\text{групп})$$

Таким образом, для городской местности мы построим 8 групп, для сельской местности – 7 групп.

Определим длину интервала h_i

$$h_i = \frac{x_{\max} - x_{\min}}{k}; \quad (3.2)$$

$$h_{i_{\text{город}}} = \frac{62,30 - 9,10}{8} = 6,65$$

$$h_{i_{\text{сельск}}} = \frac{68,65 - 8,00}{7} = 8,66$$

Для городской местности мы получили такую таблицу

Таблица 3.1

Распределение регионов городской местности

Доли внебрачной рождаемости, %	Число регионов
9,10-15,75	1
15,75-22,40	15
22,40-29,05	33
29,05-35,70	24
35,70-42,35	9
42,35-49,00	3
49,00-55,65	0
55,65-62,30	2

Из таблицы видно, что не все интервалы заполнены (это связано с очень высокими долями внебрачной рождаемости в Корякском автономном округе и в республике Тыва – 59,75% и 62,30% соответственно). Поэтому сократим количество или увеличим количество интервалов. При увеличении количества групп до 9 и 10 также остаются пустые группы. При уменьшении количества групп до 6 и 7 так же остаются незаполненные группы. Поэтому мы будем рассчитывать данные из 8 групп, одна из групп будет пустой.

Для сельской местности мы получили такую таблицу

Таблица 3.2

Распределение регионов сельской местности

Доля внебрачной рождаемости, %	Число регионов
8,00-16,66	2
16,66-25,32	8
25,32-33,98	27
33,98-42,64	20
42,64-51,30	16
51,30-59,96	9
59,96-68,62	4

Таблица 3.3

Показатели вариации у городского и сельского населения		
	Городское население	Сельское население
Среднее	29,42 (29,09)	37,62 (38,01)
Стандартная ошибка	0,91	1,28
Медиана	27,7 (27,53)	35,85 (36,56)
Мода	34,65 (23,08)	28,5 (31,64)
Стандартное отклонение	8,50 (2,24)	11,83 (11,99)
Дисперсия выборки	72,17 (67,98)	139,97 (143,8)
Эксцесс	3,08 (976,15)	0,20 (-0,45)
Асимметричность	1,19 (57,05)	0,40 (0,30)
Интервал	53,2	60,65
Минимум	9,1	8
Максимум	62,3	68,65
Сумма	2559,2	3235,6
Счет	87	86
Наибольший(1)	62,3	68,65
Наименьший(1)	9,1	8
Уровень надежности(95,0%)	1,81	2,536569537

Средняя арифметическая.

Компьютер словом «среднее» обозначает среднюю арифметическую величину.

При самостоятельном расчёте среднюю величину определим по формуле средней арифметической взвешенной.

$$\bar{x} = \frac{\sum x_i m_i}{\sum m_i}; \quad (3.3)$$

$$\bar{x}_{\text{гордск}} = \frac{2531}{87} = 29,09$$

$$\bar{x}_{\text{сельск}} = \frac{3269}{86} = 38,01$$

Средняя арифметическая показывает, что в 2004-2005 годах среднем, в городской местности 29,42% (29,09%) женщин родили детей вне официального брака. В сельской местности в 2004-2005 годах среднем, в 37,62% (38,01%) женщин родили детей вне официального брака.

Стандартная ошибка.

Здесь стандартная ошибка показывает коэффициент вариации. Эта величина абстрактно-теоретическая и интерпретации не подлежит.

Медиана.

При самостоятельном расчёте найдём медиану по следующей интерполяционной формуле.

$$Me = x_{k-1} + h_k \cdot \frac{\frac{1}{2} \sum_i m_i - F_{k-1}}{m_k}; \quad (3.4)$$

где x_{k-1} – нижняя граница медианного интервала;

h_k – длина медианного интервала;

F_{k-1} – накопленная частота интервала, предшествующего медианному;

m_k – частота медианного интервала.

Найдём медианный интервал. Для этого используем накопленные частоты. Соответственно номер медианы равен $\frac{1}{2} \sum_i m_i$

$$\text{Для городской местности } \frac{1}{2} \sum m_i = \frac{1}{2} \cdot 87 = 43,5$$

$$\text{Для сельской местности } \frac{1}{2} \sum m_i = \frac{1}{2} \cdot 86 = 43$$

$$Me_{\text{городск}} = 22,4 + 6,65 \cdot \frac{43,5 - 18}{33} = 27,53$$

$$Me_{\text{сельск}} = 33,98 + 8,66 \cdot \frac{43 - 37}{20} = 36,56$$

Медиана показывает, что в 2004-2005 годах половине регионов в городской местности вне брака было рождено более 27,7% (27,5%) детей, а в половине регионов – менее 27,7% детей (27,5%). Для сельской местности данный показатель интерпретируется аналогично: в 2004-2005 годах в половине регионов вне брака было рождено более 35,85% (36,56%) детей, а в половине регионов – менее 35,85% (36,56%) детей

Мода.

Мода показывает наиболее часто встречающееся значение признака. В связи с тем, что у нас данные округлены и ряд не интервальный, данные, выданные компьютером, интерпретировать не следует.

Сделаем самостоятельный расчёт моды. Так как у нас имеется интервальный ряд с равными интервалами, определим моду по интерполяционной формуле.

$$Mo = x_{k-1} + h_k \cdot \frac{m_k - m_{k-1}}{(m_k - m_{k-1}) + (m_k - m_{k+1})}; \quad (3.5)$$

где x_{k-1} – нижняя граница модального интервала;

h_k – длина модального интервала;

m_{k-1}, m_k, m_{k+1} частота интервала, соответственно предшествующего модальному, модального и следующего за модальным.

$$Mo_{городск} = 22,4 + 6,65 \cdot \frac{38-17}{38-17+38-28} = 23,08 \quad Mo_{сельск} = 25,32 + 8,66 \cdot \frac{27-8}{(27-8)+(27-20)} = 31,64$$

Стандартное отклонение.

Компьютер словами «стандартное отклонение» здесь показывает среднее квадратическое отклонение.

Самостоятельно рассчитаем СКО.

$$\sigma = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{\sum_i m_i}}; \quad (3.6)$$

$$\sigma_{\text{городск}} = \sqrt{\frac{5914,6}{87}} = 2,244998$$

$$\sigma_{\text{сельск}} = \sqrt{\frac{12366,4}{86}} = 11,99146983$$

Среднее квадратическое отклонение показывает среднее отклонение отдельных вариантов от их средней величины. В городской местности в среднем показатели внебрачной рождаемости отклоняются от средних значений внебрачной рождаемости в городах на 8,50% (2,24%). В сельской местности в среднем показатели внебрачной рождаемости отклоняются от средних значений внебрачной рождаемости в сельской местности на 11,83% (11,99%).

Дисперсия выборки.

Дисперсия – очень важный статистический показатель, который широко применяется, но само значение дисперсии интерпретации не подлежит.

Самостоятельно рассчитаем дисперсию.

$$\sigma^2 = \frac{\sum_i (x_i - \bar{x})^2 \cdot m_i}{\sum_i m_i}; \quad (3.7)$$

$$\sigma_{\text{городск.}}^2 = \frac{5914,6238}{87} = 67,98418$$

$$\sigma_{\text{сельск}}^2 = \frac{12366,4}{86} = 143,7953488$$

Эксцесс.

Эксцесс показывает, существует ли в рассматриваемой совокупности т.н. «ядро», т.е. насколько много единиц совокупности по своему признаку сосредоточено рядом со средней величиной.

Для городского населения эксцесс 3,08, а это значит, что в изучаемой массе явлений существует слабо варьирующее по данному признаку «ядро», окруженное рассеянным «галом».

Для сельского населения эксцесс 0,2, что говорит о нормальности распределения эксцесса.

Подтвердим эти выводы расчетами. Распределение можно считать нормальным, если показатель эксцесса не превышает своего двукратного среднего квадратического отклонения σ_{ex}

$$\sigma_{ex} = \sqrt{\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}} \quad (3.8)$$

где n- число единиц совокупности

В городской местности число единиц совокупности (субъектов федерации) 87, в сельской- 86.

$$\sigma_{ex_город.} = \sqrt{\frac{24 \cdot 87 \cdot 86^2}{84 \cdot 85 \cdot 90 \cdot 92}} = 0,51 \quad \sigma_{ex_сельск.} = \sqrt{\frac{24 \cdot 86 \cdot 85^2}{83 \cdot 84 \cdot 89 \cdot 91}} = 1,67$$

Теперь рассчитаем соотношение эксцесса к его двукратному среднему квадратическому отклонению σ_{ex} .

$$\frac{Ex_{гор.}}{\sigma_{ex_город.}} = \frac{3,08}{0,51} = 6,04 \quad \frac{Ex_{сельск.}}{\sigma_{ex_сельск.}} = \frac{0,2}{1,67} = 0,12$$

Теперь рассчитаем эксцесс самостоятельно

$$Ex = \frac{\sum_j (x'_j - \bar{x})^4}{\sum_j f} - 3 \quad (3.9)$$

$$Ex_{городск} = \frac{2145020}{2,24^4} - 3 = 979,15 = 976,15 \quad Ex_{сельск} = \frac{4551225}{11,99^4} - 3 = 2,55 - 3 = -0,45$$

Мы видим, что для городского населения соотношение эксцесса к его среднему квадратическому отклонению больше 2, а это значит, что наши выводы подтверждаются – в изучаемой массе явлений существует слабо варьирующее по данному признаку «ядро», окруженное рассеянным «галом».

Так же мы видим, что для сельского населения соотношение эксцесса к его среднему квадратическому отклонению близко к 0, а это значит, что наши выводы о нормальности распределения эксцесса подтверждаются.

Асимметричность.

Здесь асимметричность показывает коэффициент асимметрии. И для городского, и для сельского населения асимметрия положительна, а это значит, что и в городской, и в сельской местностях преобладают субъекты Федерации, в которых показатели внебрачной рождаемости выше средних.

Подтвердим эти выводы расчетами. Распределение можно считать нормальным, если показатель асимметрии не превышает своего двукратного среднего квадратического отклонения σ_{as} .

$$\sigma_{as} = \sqrt{\frac{6(n-1)n}{(n-2)(n+1)(n+3)(n+3)}} \quad (3.9)$$

$$\sigma_{as_город.} = \sqrt{\frac{6 \cdot 86 \cdot 87}{85 \cdot 88 \cdot 90^2}} = 0,027 \quad \sigma_{as_сельск.} = \sqrt{\frac{6 \cdot 85 \cdot 86}{84 \cdot 87 \cdot 89^2}} = 0,028$$

Теперь рассчитаем соотношение асимметрии к её двукратному среднему квадратическому отклонению σ_{as} .

$$\frac{As_{гор}}{\sigma_{as_город.}} = \frac{1,19}{0,027} = 44,04 \quad \frac{As_{сельск}}{\sigma_{as_сельск.}} = \frac{0,4}{0,028} = 14,28$$

Теперь рассчитаем асимметрию самостоятельно

$$As = \frac{\sum_j (x'_j - \bar{x})^3}{\sum_j f \sigma^3} \quad (3.10)$$

$$As_{городск} = \frac{55790}{2,24^3} = 57,05 \quad As_{сельск} = \frac{44254,71}{11,99^3} = 0,30$$

$$\frac{As_{городск}}{\sigma_{as_городск}} = \frac{57,05}{0,027} = 2112,96 \quad \frac{As_{сельск}}{\sigma_{as_сельск}} = \frac{0,30}{0,028} = 10,71$$

И в городской, и в сельской местности соотношение эксцесса к его среднему квадратическому отклонению больше 2 (т. при расчетах компьютера, и при самостоятельных расчётах), а это значит, что наши выводы подтверждаются – и в городской, и в сельской местностях преобладают субъекты федерации, в которых показатели внебрачной рождаемости выше средних.

Размах вариации

Здесь интервал показывает размах вариации. Размах вариации – оценка рассеяния признака. Здесь размах вариации показывает, насколько могут различаться максимальные и минимальные значения признака. Так, в городской местности показатели внебрачной рождаемости максимально различаются на 53,2%, а в сельской местности на 60,65%.

Так как интервальный ряд мы построили на основе дискретного, и в закрытых интервальных рядах размах вариации рассчитывается как разница между верхней границей последнего интервала и нижней границей первого интервала, то данный показатель полностью совпадает по своему значению с рассчитанным компьютером.

Данный показатель является довольно грубой оценкой.

Минимум (Наименьший).

Здесь минимум (наименьший) показывает наименьшее значение признака в предложенной совокупности. Так, мы видим, что в городской местности минимальная доля внебрачных рождений – 9,1% (в республике Ингушетия), а сельской местности минимальная доля внебрачных рождений – 8% (в республике Ингушетия)

Данный показатель полностью совпадает с нижней границей первого интервала

Максимум (Наибольший).

Здесь максимум (наибольший) показывает наибольшее значение признака в предложенной совокупности. Так, мы видим, что в городской местности максимальная доля внебрачных рождений – 62,3% (в республике Тыва), а сельской местности максимальная доля внебрачных рождений – 68,65% (в республике Тыва).

Данный показатель полностью совпадает с верхней границей последнего интервала.

Сумма.

Сумма показывает сумму значений признаков. В данном случае суммирование интерпретации не подлежит, так как признаки уже являются соотношениями.

Счёт.

Здесь счёт показывает сумму единиц совокупности. В городской местности 87 единиц совокупности (субъектов федерации), в сельской местности 86 единиц совокупности (субъектов федерации).

Уровень надёжности.

Здесь уровень надёжности показывает уровень значимости. И для городской, и для сельской местности выбран уровень значимости 95%. Это значит, что в 5 случаях из 100 может быть отвергнута правильная гипотеза.

Также рассчитаем другие показатели вариации.

Коэффициент вариации

$$V_1 = \frac{\bar{d}}{Me} \cdot 100\% \quad (3.11)$$

$$V_2 = \frac{\sigma}{x} \cdot 100\% \quad (3.12)$$

$$V_{1\text{городск}} = \frac{6,3875}{27,53} \cdot 100\% = 23,20\%$$

$$V_{1\text{сельск}} = \frac{9,6079}{36,56} \cdot 100\% = 26,28\%$$

$$V_{2_{\text{городск}}} = \frac{2,244998}{29,09} \cdot 100\% = 7,71\%$$

$$V_{2_{\text{сельск}}} = \frac{11,99146983}{38,01162791} \cdot 100\% = 31,54684$$

Среднее линейное отклонение.

$$\bar{d} = \frac{\sum_i |x_i - \bar{x}| m_i}{\sum m_i} \quad (3.13)$$

$$\bar{d}_{\text{городск}} = \frac{555,7150}{87} = 6,3875$$

$$\bar{d}_{\text{сельск}} = \frac{826,2794}{86} = 9,6079$$

3.1 Динамический ряд

Имеются данные о рождаемости у женщин с 1960 по 2005 год. Данные с 1960 по 1980 год даны с 5-летним интервалом, с 1980 по 2005 годы – ежегодно (см. Приложение 9).

В таблице представлены данные о рождаемости с разделением на рождения в зарегистрированном браке и внебрачные рождения.

Восстановим отсутствующие коэффициенты доли рождений в незарегистрированных браках в числе рождений в фактических браках за 1971-1974 годы и за 1976-1979 годы. Здесь признаком-фактором (здесь- x) будет год, а признаком результатом (y) - коэффициент. Мы видим, что в 1970 и 1975 годах коэффициент одинаковый - 0,05. Таким образом, линейное уравнение типа $y=ax+b$ в общем виде будет выглядеть как $y=0,05x$. Таким образом, в 1971-1974 годах доля рождений в незарегистрированных браках в числе рождений в фактических браках будет 0,05.

Для восстановления данных за 1976-1979 решим систему линейных уравнений.

$$\begin{cases} 1975 = 0,05a + b \\ 1980 = 0,04a + b \end{cases}$$

откуда

$$\begin{cases} b = 1975 - 0,05a \\ 1980 = 0,04a + (1975 - 0,05a) \end{cases}$$

откуда

$$\begin{cases} a = -500 \\ b = 2000 \end{cases}$$

Таким образом, уравнение принимает вид $y = -500x + 2000$

Подставляя в y значения годов (1976; 1977; 1978; 1979), получаем соответствующие значения x (0,048; 0,046; 0,044; 0,042)

Для восстановления данных за 1981-1987 решим систему линейных уравнений

$$\begin{cases} 1980 = 0,04a + b \\ 1988 = 0,06a + b \end{cases}$$

откуда

$$\begin{cases} a = \frac{1980 - b}{0,04} \\ 1988 = 0,06\left(\frac{1980 - b}{0,04}\right) + b \end{cases}$$

откуда

$$\begin{cases} b = 1964 \\ a = 400 \end{cases}$$

Таким образом, уравнение принимает вид $y = 400x + 1964$

Подставляя в y значения годов (1981; 1982; 1983; 1984; 1985; 1986; 1987), получаем соответствующие значения x (0,0425; 0,045; 0,0475; 0,05; 0,0525; 0,055; 0,0575).

Восстановленный ряд предложен в Приложении 10.

4 Заключение

В заключении хочу сказать, что свою основную гипотезу о том, что внебрачная рождаемость в современной России полностью определяется долей незамужних женщин.

5 Список использованной литературы

1. Демографический ежегодник России. 2006: Стат. сб. / Д 31 Росстат. - М., 2006, стр. 165-166
2. Елисеева И.И., Юзбашев М.М. Общая теория статистики.- М.: Финансы и статистика.- 2004
3. Демография и статистика населения/ Под. ред. И.И.Елисеевой.-М.: Финансы и статистика.-2006.
4. Теория статистики/Под. ред. Г.Л.Громыко.-М: Инфра-М.-2006