**Лекция 2**

**2. Анализ и выравнивание динамических рядов**

* 1. **Анализ динамических рядов**
		1. **Определение динамического ряда и их виды**

Изучение изменения явлений во времени является одной из важных задач статистики. Решается эта задача при помощи составления и изучения рядов динамики.

**Ряд динамики** представляет собой ряд числовых значений определенного статистического показателя в последовательные моменты или периоды времени. Числовые значения того или иного статистического показателя, составляющие динамический ряд, принято называть уровнями ряда $(y\_{i}).$ Одной из **основных задач** исследования рядов динамики является **выявление определенной закономерности в изменении уровней ряда**, т. е. основной тенденции изменения уровней, именуемой трендом.

Основное требование динамического ряда – сопоставимость уровней. **Сопоставимость** **уровней** возможна по:

1. территории (требуется неизменность границ территорий);
2. кругу исследуемых объектов (подразумевается одинаковая полнота охвата исследуемого объекта);
3. времени регистрации (учет сезонности явления);
4. стоимостным показателям (неизменность цен в исследуемый период);
5. методологии расчета;
6. единицам измерения;
7. достоверности результатов (одинаковая репрезентативность выборки за разные периоды исследования).

 **Виды динамических рядов**:

**1.** **в зависимости от вида показателей**:

1) абсолютные;

2) относительные;

3) средние величины;

**2. в зависимости от отношений уровня динамического ряда к определенным моментам**:

1) моментные – ряды, уровни которых характеризуют величину явления по состоянию на определенные моменты времени;

 2) интервальные – ряды, уровни которых характеризуют величину изучаемого показателя, полученную в итоге за определенный пери времени;

**3. в зависимости от расстояния между датами:**

1) равные ряды;

2) неравные (равные интервалы между датами не соблюдаются);

**4. в зависимости от содержания показателей ряда:**

1) частные – характеризуют явление изолированно, односторонне;

2) агрегированные – производные из частных и характеризуют явление комплексно.

**2.1.2 Показатели динамического ряда**

При составлении уровней динамического ряда анализируются следующие **показатели** **ряда**:

1) **абсолютный прирост** ($A\_{i})$: $A\_{i}=y\_{i}-y\_{i-1}$, где *i* = 1,…, *n*, *n* – число уровней ряда (цепной прирост);

2) **базисный прирост** $∆\_{б}:$$∆\_{б}=y\_{i}-y\_{0}$, $y\_{0}$ – базисный уровень ряда;

3) **коэффициент роста** ($K\_{i}$) определяется как отношение последующего к предыдущему уровню ряда: $K\_{i} = y\_{i} / y\_{i-1}$ ;

4) **цепной темп роста** ($T\_{р}^{ц}$): $T\_{р}^{ц}=\left(y\_{i}:y\_{i-1}\right)\*100$;

5) **базисные темпы роста** ($T\_{р}^{б}$): $T\_{р}^{б}=\left(y\_{i}:y\_{0}\right)\*100;$

6) **темп прироста** ($T\_{i}$) – это отношение абсолютного прироста к уровню предыдущего периода (%): $T\_{i} =( A\_{i}/ y\_{i-1} )\*100$;

7) **значение 1% прироста** ($P\_{i}$): $P\_{i}={A\_{i}}/{T\_{i}} или P\_{i} = y\_{i-1} / 100$;

8) **средний уровень динамического ряда** (*ỹ*) определяется как среднее арифметическое приведенного ряда: $ \tilde{y}= \sum\_{}^{}y\_{i}/ n$;

9) **средний абсолютный прирост ряда** (*Г*) определяет скорость развития явления во времени: *Г* $= (y\_{n} – y\_{0} ) / (n - 1) .$ Этот показатель дает возможность определить, насколько в среднем за единицу времени должен измениться уровень ряда, чтобы отправляясь от начального уровня за данное число периодов, можно было бы достичь конечного уровня;

10) **средний коэффициент роста** ($\tilde{K}$), который показывает, во сколько раз в среднем за единицу времени изменился уровень динамического ряда: $\acute{K} =\sqrt[m]{\prod\_{}^{}K\_{i}}$, *m* – число коэффициентов роста.

**2.2. Выравнивание динамических рядов**

Для исключения влияния случайных компонентов динамические ряды подвергаются выравниванию. Выравнивание (сглаживание) динамического ряда может быть проведено несколькими способами:

1. **метод укрупнения интервалов** – переход от первоначальных значений динамического ряда к ряду с большими временными промежутками;

выбирается количество уровней в укрупненном интервале и находится среднее значение для этих данных;

1. **метод скользящей средней** – расчет средних уровней динамического ряда по укрупненным интервалам при смещении начала отчета на один временный период, применять можно трех-, пяти- и т.д. скользящие средние. Выбор интервала сглаживания зависит от четности и нечетности ряда;

3) **выравнивание по среднему абсолютному приросту**:

 $Y\_{ср,i}=y\_{0}+Г(i-1)$, где $Y\_{ср, i}$ – выравненное значение показателя, $y\_{0}$ – начальное (базисное) значение динамического ряда, *Г* – средний абсолютный прирост;

1. **выравнивание по среднему коэффициенту роста**:

 $Y\_{ср,i}=y\_{0}\acute{K}(i-1)$, где $\acute{K}$ – средний коэффициент роста;

5) **выравнивание по способу наименьших квадратов**. Проводится с учетом предполагаемой тенденции изменения показателя.

При линейной тенденции выравнивание идет с учетом уравнения прямой: $Y\_{ср,i}=a\_{0}+a\_{1}t$ , где $a\_{0}$ и $a\_{1}$ - параметры линейного уравнения; *t* – порядковый номер года в динамическом ряду.

При параболическом выравнивании уравнение имеет вид:$ Y\_{ср,i}=a\_{0}+a\_{1}t+a\_{2}t^{2}$.

Для нахождения неизвестных коэффициентов используем метод наименьших квадратов.

Выравнивание динамических рядов различными методами приводятся в таблицах 2.1 –5.3. Для наглядного представления полученных результатов строятся графики и диаграммы. Как видно из графиков, отражающих выровненные значения показателей по всем метода наиболее близки к фактическим значениям результаты выравнивания по способу наименьших квадратов.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Год | Производственные затраты | Сумма за три года | Выравненная урожайность | Среднее за три года |
| 1 | 100 | - | - |  |
| 2 | 105 | 307,0 | 102,3 | 102,3 |
| 3 | 102 | 318,0 | 106,0 |  |
| 4 | 111 | 328,0 | 109,3 |  |
| 5 | 115 | 346,0 | 115,3 | 115,3 |
| 6 | 120 | 365,0 | 121,7 |  |
| 7 | 130 | 390,0 | 130,0 |  |
| 8 | 140 | 435,0 | 145,0 | 145,0 |
| 9 | 165 | 481,0 | 160,3 |  |
| 10 | 176 | 529,0 | 176,3 |  |
| 11 | 188 | 577,0 | 192,3 | 192,3 |
| 12 | 213 | 651,0 | 217,0 |  |
| 13 | 250 | 722,0 | 240,7 |  |
| 14 | 259 | 779,0 | 259,7 | 259,7 |
| 15 | 270 | - | - |  |

Выравнивание динамического ряда производственных затрат по среднему абсолютному приросту Х1i =100+12,14\*(i-1)

|  |  |  |
| --- | --- | --- |
| Год | Производственные затраты | Выровненное значение |
| 1 | 100 | 100,0 |
| 2 | 105 | 112,1 |
| 3 | 102 | 124,3 |
| 4 | 111 | 136,4 |
| 5 | 115 | 148,6 |
| 6 | 120 | 160,7 |
| 7 | 130 | 172,9 |
| 8 | 140 | 185,0 |
| 9 | 165 | 197,1 |
| 10 | 176 | 209,3 |
| 11 | 188 | 221,4 |
| 12 | 213 | 233,6 |
| 13 | 250 | 245,7 |
| 14 | 259 | 257,9 |
| 15 | 270 | 270,0 |

Выравнивание динамического ряда производственных затрат по среднему коэффициенту роста Х1i =100,0\*1,074i-1

|  |  |  |
| --- | --- | --- |
| Год | Производственные затраты | Выровненное значение |
| 1 | 100 | 100,0 |
| 2 | 105 | 107,4 |
| 3 | 102 | 115,2 |
| 4 | 111 | 123,7 |
| 5 | 115 | 132,8 |
| 6 | 120 | 142,6 |
| 7 | 130 | 153,1 |
| 8 | 140 | 164,3 |
| 9 | 165 | 176,4 |
| 10 | 176 | 189,4 |
| 11 | 188 | 203,3 |
| 12 | 213 | 218,2 |
| 13 | 250 | 234,3 |
| 14 | 259 | 251,5 |
| 15 | 270 | 270,0 |

Выравнивание производственных затрат способом наименьших квадратов

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Год | Производственные затраты | Ранг года | Уi \*ti | t2 | Выровненная производственные затраты, |
| 1 | 100 | -7 | -700,0 | 49 | 72,2 |
| 2 | 105 | -6 | -630,0 | 36 | 85,2 |
| 3 | 102 | -5 | -510,0 | 25 | 98,1 |
| 4 | 111 | -4 | -444,0 | 16 | 111,1 |
| 5 | 115 | -3 | -345,0 | 9 | 124,1 |
| 6 | 120 | -2 | -240,0 | 4 | 137,0 |
| 7 | 130 | -1 | -130,0 | 1 | 150,0 |
| 8 | 140 | 0 | 0,0 | 0 | 162,9 |
| 9 | 165 | 1 | 165,0 | 1 | 175,9 |
| 10 | 176 | 2 | 352,0 | 4 | 188,8 |
| 11 | 188 | 3 | 564,0 | 9 | 201,8 |
| 12 | 213 | 4 | 852,0 | 16 | 214,8 |
| 13 | 250 | 5 | 1250,0 | 25 | 227,7 |
| 14 | 259 | 6 | 1554,0 | 36 | 240,7 |
| 15 | 270 | 7 | 1890,0 | 49 | 253,6 |
| Итого | 2444 | 0,0 | 3628,0 | 280,0 |  |

Статистическая группировка земельно-кадастровых показателей и построение статистических таблиц При обработке статистических материалов возникает необходимость выделения однородных групп, типов, а затем уже описание этих определенными количественными характеристиками. Расчленение совокупности на группы, однородные по какому-либо признаку, на группировкой. Для ее проведения необходимо определить те характеристики (группировочные признаки), по которым она будет пров и их значений, отделяющих одну группу от другой (интервалы группировки). В данной работе группировка будет проводиться по таким важным экономическим показателям, как урожайность, стоимость валовой продукции и чистого дохода на 1 га сельскохозяйственных угодий. Для этого в таблицах 7.1 – 7.5 производится расчет валовой продук чистого дохода и производственных затрат по всем культурам в 20 хозяйствах с использованием следующих формул: - валовая продукция: , где Вi – удельный вес культуры в данном хозяйстве; Уi – урожайность культуры; Цi – кадастровая цена за 1 ц продукции. - производственные затраты: , где Пi – производственные затраты по i-ой культуре. - чистый доход: . По способу составления группировка может быть первичной и вторичной. Величина интервала превичной группировки (h) определяе формуле: h = (Xmax – Xmin ) / K, где Xmax и Xmin – максимальное и минимальное значения признака; К – количество групп, на которые подразделяются хозяйства. Определяется по формуле Стерджеса: K = 1+1,32\*lg n. При вторичной группировке количество хозяйств в группе одинаково или несильно отличается, введены открытые интервалы. При комбинированной группировке производится выделение подгрупп в каждой группе по какому-либо другому показателю (возможна группировка и более чем по двум показателям). В таблицах 9.1-9.8 выделены группы хозяйств по урожайности культур, которые подразделяются на подгруппы по баллу бонитета по свойствам почв. Анализируя полученные результаты, видим, что при общей зави урожайности от количества применяемых удобрений по группам, наблюдается более высокая отдача на земле с более высоким баллом бонитета почв. Для выявления причин, определяющих величину чистого дохода, строится аналитическая таблица 10, в которую включаются 5 факто валовая продукция, производственные затраты, основные производственные фонды, затраты живого труда и затраты удобрений.

Любая информация должна быть представлена так, чтобы ей удобно было пользоваться. Существует несколько способов представления информации: текстовый, табличный, графический. Наиболее эффективным способом представления больших объемов земельно-кадастровых (статистических) данных являются таблицы. 6 Графический метод используется как в дополнение к таблицам, так и самостоятельно. С помощью графиков достигается наглядность взаимосвязи явлений, характеристики структуры и динамики. Используя текстовый, табличный и графический способы представления информации, возможно, провести полноценный анализ распределения земель

Таблица 1 Динамика распределения земель Пензенской области по категориям за период с 2008 по 2013 года, га Годы Категория земель 2008 2009 2010 2011 2012 2013 1. Земли сельскохозяйственного назначения 3074,9 3074,2 3074.1 3073,5 3073,3 3072,9 2. Земли населенных пунктов 227,3 227,5 227,7 228,0 228,2 228,2 3. Земли промышленности, энергетики транспорта, связи, радиовещания, телевидения, информатики, земли для обеспечения космической деятельности, земли обороны, безопасности и земли иного специального назначения 42,1 42,7 42,8 43,1 43,1 43,5 4. Земли особо охраняемых территорий и объектов 9,1 9,1 9,1 9,1 9,1 9,1 5. Земли лесного фонда 964,5 964,5 964,5 964,5 964,5 964,5 6. Земли водного фонда 14,8 14,8 14,8 14,8 14,8 14,8 7. Земли запаса 2,5 2,4 2.2 2,2 2,2 2,2 Итого земель 4335,2 4335,2 4335,2 4335,2 4335,2 4335,2

5. ОСНОВЫ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ ДАННЫХ 5.1. Расчёт показателей вариации Вариация является одной из важнейших категорий, применяемых в статистической науке. Явления, подверженные вариации лежат в области исследования статистической науки, в то время как явления неизменные, статичные, постоянные в статистике не рассматриваются. Вариация - это принятие единицами совокупности или их группами различных, отличающихся друг от друга, значений признака. Вариация является результатом воздействия на единицы совокупности множества факторов. Синонимами термина «вариация» являются понятия «изменение», «изменчивость», «вариативность». Необходимость в измерении вариации возникает из-за того, что в средней величине не проявляется степень колеблемости отдельных значений признаков (вариант) вокруг среднего уровня. В зависимости от однородности в совокупности, степень колеблемости может быть большой или маленькой. Вариацией называется изменчивость только тех явлений, на которые воздействуют внешние факторы и причины. Тогда как о явлениях, изменяющихся в силу своей внутренней природы нельзя говорить, что они подвержены вариации. Например, рост человека, меняющийся в течение жизни. Различие индивидуальных значений признака внутри изучаемой совокупности в статистике называется вариацией признака. Она возникает в результате того, что его индивидуальные значения складываются под совокупным влиянием разнообразных факторов, которые по-разному сочетаются в каждом отдельном случае. Средняя величина это абстрактная, обобщающая характеристика признака изучаемой совокупности, но она не показывает строения совокупности, которое весьма существенно для ее познания. Средняя величина не дает представления о том, как отдельные значения изучаемого признака 36 группируются вокруг средней, сосредоточены ли они вблиз

группируются вокруг средней, сосредоточены ли они вблизи или значительно отклоняются от нее. В некоторых случаях отдельные значения признака близко примыкают к средней арифметической и мало от нее отличаются. В таких случаях средняя хорошо представляет всю совокупность. В других, наоборот, отдельные значения совокупности далеко отстают от средней, и средняя плохо представляет всю совокупность. Колеблемость отдельных значений характеризуют показатели вариации. Термин «вариация» произошел от латинского variation «изменение, колеблемость, различие». Однако не всякие различия принято называть вариацией. Под вариацией в статистике понимают такие количественные изменения величины исследуемого признака в пределах однородной совокупности, которые обусловлены перекрещивающимся влиянием действия различных факторов. Различают вариацию признака: случайную и систематическую. Анализ систематической вариации позволяет оценить степень зависимости изменений в изучаемом признаке от определяющих ее факторов. Например, изучая силу и характер вариации в выделяемой совокупности, можно оценить, насколько однородной является данная совокупность в количественном, а иногда и качественном отношении, а, следовательно, насколько характерной является исчисленная средняя величина. Степень близости данных отдельных единиц хi к средней измеряется рядом абсолютных, средних и относительных показателей. Абсолютные и средние показатели вариации и способы их расчета Для характеристики совокупностей и исчисленных величин важно знать, какая вариация изучаемого признака скрывается за средним. Для характеристики колеблемости признака используется ряд показателей. Наиболее простой из них - размах вариации. Размах вариации это разность между наибольшим ( ) и наименьшим ( ) значениями вариантов. xmak xmin 37 Достоинством этого показателя является простота расчёта. Точнее характеризует вариацию признака показатель, основанный на учёте всех значений признака. К таким показателям относится среднее линейное отклонение, дисперсия и среднее квадратическое отклонение, представляющие собой среднюю арифметическую из отклонений индивидуальных значений признака от средней арифметической. Чтобы дать обобщающую характеристику распределению отклонений, исчисляют среднее линейное отклонение d, которое учитывает различие всех единиц изучаемой совокупности. Среднее линейное отклонение определяется как средняя арифметическая из отклонений индивидуальных значений от средней, без учета знака этих отклонений: . Порядок расчета среднего линейного отклонения следующий: 1) по значениям признака исчисляется средняя арифметическая: ; 2) определяются отклонения каждой варианты от средней: x x i ; 3) рассчитывается сумма абсолютных величин отклонений: ; 4) сумма абсолютных величин отклонений делится на число значений: . Если данные наблюдения представлены в виде дискретного ряда распределения с частотами, среднее линейное отклонение исчисляется по формуле средней арифметической взвешенной: R xmak xmin d x x n x x x x x x n n / / / / / / ... / / 1 2 x x n xi /x x / i / x x / n i 38 Порядок расчета среднего линейного отклонения взвешенного следующий: 1) вычисляется средняя арифметическая взвешенная: ; 2) определяются абсолютные отклонения вариант от средней: / /; 3) полученные отклонения умножаются на частоты ; 4) находится сумма взвешенных отклонений без учета знака: ; 5) сумма взвешенных отклонений делится на сумму частот: . Расчет дисперсии и среднего квадратического отклонения по индивидуальным данным и в рядах распределения Основными обобщающими показателями вариации в статистике являются дисперсии и среднее квадратическое отклонение. Дисперсия - это средняя арифметическая квадратов отклонений каждого значения признака от общей средней. Дисперсия обычно называется средним квадратом отклонений и обозначается . В зависимости от исходных данных дисперсия может вычисляться по средней арифметической простой или взвешенной: – дисперсия невзвешенная (простая); – дисперсия взвешенн

Среднее квадратическое отклонение представляет собой корень квадратный из дисперсии и обозначается S: – среднее квадратическое отклонение невзвешенное; – среднее квадратическое отклонение взвешенное. Среднее квадратическое отклонение - это обобщающая характеристика абсолютных размеров вариации признака в совокупности. Выражается оно в тех же единицах измерения, что и признак (в метрах, тоннах, процентах, гектарах и т.д.). Среднее квадратическое отклонение является мерилом надежности средней. Чем меньше среднее квадратическое отклонение, тем лучше средняя арифметическая отражает собой всю представляемую совокупность. Вычислению среднего квадратического отклонения предшествует расчет дисперсии. Порядок расчета дисперсии взвешенной: 1) определяют среднюю арифметическую взвешенную ; 2) определяют отклонения вариант от средней ; 3) возводят в квадрат отклонение каждой варианты от средней ; 4) умножают квадраты отклонений на веса (частоты) 5) суммируют полученные произведения ; 6) полученную сумму делят на сумму весов S x x n i ( ) 2 S x x n n i i i ( ) 2 xn n x x i x x i 2 xi x ni 2 xi x ni 2 40 . Расчет дисперсии по индивидуальным данным и в рядах

распределения. по формуле Техника вычисления дисперсии сложна, а при больших значениях вариант и частот может быть громоздкой. Расчеты можно упростить, используя свойства дисперсии. Свойства дисперсии Уменьшение или увеличение весов (частот) варьирующего признака в определенное число раз дисперсии не изменяет. Уменьшение или увеличение каждого значения признака на одну и ту же постоянную величину А дисперсии не изменяет. Уменьшение или увеличение каждого значения признака в какое-то число раз к соответственно уменьшает или увеличивает дисперсию в раз, а среднее квадратическое отклонение в к раз. Дисперсия признака относительно произвольной величины всегда больше дисперсии относительно средней арифметической на квадрат разности между средней и произвольной величиной: . Если А равна нулю, то приходим к следующему равенству: , т.е. дисперсия признака равна разности между средним квадратом значений признака и квадратом средней. Каждое свойство при расчете дисперсии может быть применено самостоятельно или в сочетании с другими. Порядок расчета дисперсии простой: 1) определяют среднюю арифметическую ; x x n n i i i 2 S x x 2 2 2 k 2 S SA x A 2 2 2 ( ) S x x 2 2 2 x x n 41 2) возводят в квадрат среднюю арифметическую ; 3) возводят в квадрат каждую варианту ряда ; 4) находят сумму квадратов вариант ; 5) делят сумму квадратов вариант на их число, т.е. определяют средний квадрат ; 6) определяют разность между средним квадратом признака и квадратом средней . Расчет дисперсии в интервальном ряду распределения. Порядок расчета дисперсии взвешенной (по формуле ): 1) определяют среднюю арифметическую ; 2) возводят в квадрат полученную среднюю ; 3) возводят в квадрат каждую варианту ряда ; 4) умножают квадраты вариант на частоты ; 5) суммируют полученные произведения ; 6) делят полученную сумму на сумму весов и получают средний квадрат признака ; 7) определяют разность между средним значением квадратов и квадратом средней арифметической, т.е. дисперсию

Показатели относительного рассеивания Для характеристики меры колеблемости изучаемого признака исчисляются показатели колеблемости в относительных величинах. Они позволяют сравнивать характер рассеивания в различных распределениях (различные единицы наблюдения одного и того же признака в двух совокупноx x n 2 2 xi 2 xi 2 x x n 2 i 2 x x 2 2 S x x 2 2 2 x xn n x 2 xi 2 x ni i 2 xi ni 2 x x n n i i i 2 2 S x x 2 2 2 42 стях, при различных значениях средних, при сравнении разноименных совокупностей). Расчет показателей меры относительного рассеивания осуществляют как отношение абсолютного показателя рассеивания к средней арифметической, умножаемое на 100%. 1. Коэффициент осцилляции отражает относительную колеблемость крайних значений признака вокруг средней: 2. Относительное линейное отклонение характеризует долю усредненного значения абсолютных отклонений от средней величины: 3. Коэффициент вариации: Учитывая, что среднеквадратическое отклонение дает обобщающую характеристику колеблемости всех вариантов совокупности, коэффициент вариации является наиболее распространенным показателем колеблемости, используемым для оценки типичности средних величин. При этом исходят из того, что если V больше 40 %, то это говорит о большой колеблемости признака в изучаемой совокупности.