# Тема лекционного дистанционного занятия по дисциплине «Статистика» для студентов факультета ПСО (Право и организация социального обеспечения) 3 курс, преподаватель Гасанова С.Р.

**Лекция № 6.** Показатели вариации

**Показатели вариации.**При изучении варьирующего признака у единиц совокупности нельзя ограничиваться лишь расчетом средней величины из отдельных вариантов, так как одна и та же средняя может относиться далеко не к одинаковым по составу совокупностям.

Вариацией признака называется различие индивидуальных значений признака внутри изучаемой совокупности.

Термин «вариация» произошел от латинского variatio – 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. Колеблемость отдельных значений характеризуют показатели вариации. Чем больше вариация, тем дальше в среднем отдельные значения лежат друг от друга.

Различают вариацию признака в абсолютных и относительных величинах.

К абсолютным показателям относятся: размах вариации, среднее линейное отклонение, среднее квадратическое отклонение, дисперсия. Все абсолютные показатели имеют ту же размерность, что и изучаемые величины.

К относительным показателям относятся коэффициенты осцилляции, линейного отклонения и вариации.

*Показатели абсолютные.*Рассчитаем абсолютные показатели, характеризующие вариацию признака.

Размах вариации, представляет собой разность между максимальным и минимальным значением признака.

|  |  |
| --- | --- |
| R=Xmax–Xmin. | (6.1) |

Показатель размаха вариации не всегда применим, так как он учитывает только крайние значения признака, которые могут сильно отличаться от всех других единиц.

Более точно можно определить вариацию в ряду при помощи показателей, учитывающих отклонения всех вариантов от средней арифметической.

Таких показателей в статистике два: среднее линейное и среднее квадратическое отклонение.

Среднее линейное отклонение (L)представляет собой среднее арифметическое из абсолютных значений отклонений отдельных вариантов от средней.

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-dup6R8.png– для несгруппированных данных; | (6.2) |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-NFFWRB.png– для сгруппированных данных. | (6.3) |

Практическое использование среднего линейного отклонения заключается в следующем, с помощью этого показателя анализируется состав работающих, ритмичность производства, равномерность поставок материалов.

Недостаток этого показателя заключается в том, что он усложняет расчеты вероятного типа, затрудняет применение методов математической статистики.

Среднее квадратическое отклонение (https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-Qqf2wY.png) является наиболее распространенным и общепринятым показателем вариации. Оно несколько больше среднего линейного отклонения. Для умеренно асимметричных распределений установлено следующее соотношение между ними

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-L_ucku.png=1,25L | (6.4) |

Для его исчисления каждое отклонение от средней возводится в квадрат, все квадраты суммируются (с учетом весом), после чего сумма квадратов делится на число членов ряда и из частного извлекается корень квадратный.

Все эти действия выражает следующая формула

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-7BjxTC.png– для несгруппированных данных, | (6.5) |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-D07Vpx.png– для сгруппированных данных. | (6.6) |

т.е. среднее квадратическое отклонение представляет собой корень квадратный из средней арифметической квадратов отклонений от средней.

Среднее квадратическое отклонение является мерилом надежности средней. Чем меньше σ, тем лучше среднее арифметическое отражает собой всю представляемую совокупность.

Средняя арифметическая из квадратов отклонений вариантов значений признака от средней величины носит название дисперсии (https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-7PfWQF.png), которая рассчитывается по формулам

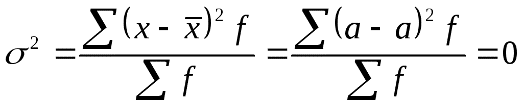
|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-pmzcYx.png– для несгруппированных, | (6.7) |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-30C70U.png– для сгруппированных. | (6.8) |

Отличительной особенностью данного показатели является то, что при возведении в квадрат (https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-srKhSf.png) удельный вес малых отклонений уменьшается, а больших увеличивается в общей сумме отклонений.

Дисперсия обладает рядом свойств, некоторые из них позволяют упростить её вычисление:

1. Дисперсия постоянной величины равна 0.

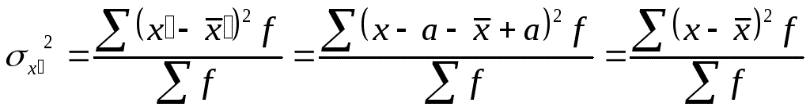
Если https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-o2U7hu.png, то иhttps://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-q_Gc7H.png.

Тогда .

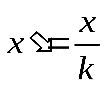
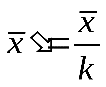
2. Если все варианты значений признака (x) уменьшить на одно и то же число, то дисперсия не уменьшится.

Пусть https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-TY6gVL.png, но тогда в соответствии со свойствами средней арифметической иhttps://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-qHxXz7.png.

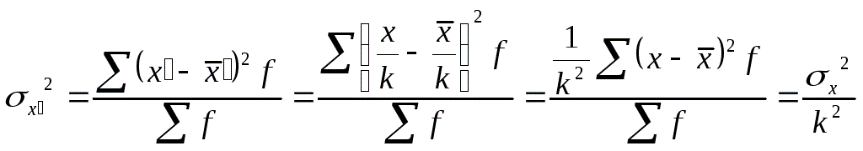
Дисперсия в новом ряду будет равна

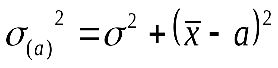
, т.е. дисперсия в рядуhttps://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-TZKvFT.pngравна дисперсии первоначального рядаhttps://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-cNaVc7.png.

3. Если все варианты значений признака уменьшить в одно и то же число раз (kраз), то дисперсия уменьшится вk2раз.

Пусть , тогда и.

Дисперсия же нового ряда https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-64L8tP.pngбудет равна



4. Дисперсия, рассчитанная по отношению к средней арифметической, является минимальной. Средний квадрат отклонений, рассчитанный относительно произвольного числа https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-63Ehqg.png, больше дисперсии, рассчитанной по отношению к средней арифметической, на квадрат разности между средней арифметической и числомhttps://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-9_x68l.png, т.е.. Дисперсия от средней имеет свойство минимальности, т.е. она всегда меньше дисперсий, исчисленных от любых других величин. В этом случае, когдаhttps://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-EoURUV.pngприравниваем к 0 и , следовательно, не вычисляем отклонения, формула принимает такой вид:

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-N_jx0J.png | (6.9) |

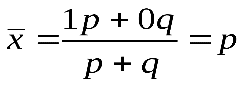
Выше был рассмотрен расчет показателей вариации для количественных признаков, но в экономических расчетах может ставиться задача оценки вариации качественных признаков*.*Например, при изучении качества изготовленной продукции, продукцию можно разделить на качественную и бракованную.

В таком случае речь идет об альтернативных признаках.

Альтернативными признаками называются такие, которыми одни единицы совокупности обладают, а другие нет. Например, наличие производственного стажа у абитуриентов, ученая степень у преподавателей ВУЗов и т.д. Наличие признака у единиц совокупности условно обозначаем через 1, а отсутствие – 0. Тогда, если долю единиц, обладающих признаком (в общей численности единиц совокупности), обозначить через р, а долю единиц, не обладающих признаком, через q, дисперсию альтернативного признака можно рассчитать по общему правилу. При этомp+q= 1 и, значит,q= 1–p.

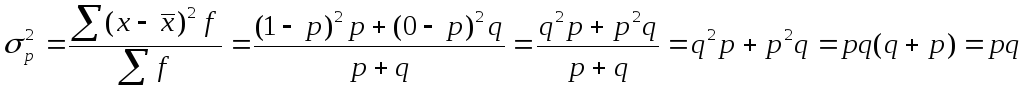
Сначала рассчитываем среднее значение альтернативного признака:

Рассчитаем среднее значение альтернативного признака

,

т.е. среднее значение альтернативного признака равно доле единиц, обладающих данным признаком.

Дисперсия же альтернативного признака будет равна:



Таким образом, дисперсия альтернативного признака равняется произведению доли единиц, обладающих данным признаком, на долю единиц, не обладающих данным признаком.

А среднее квадратическое отклонение будет равно https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-y2TgVb.png=https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-gYVYU1.png.

*Показатели относительные.*Для целей сравнения колеблемости различных признаков в одной и той же совокупности или же при сравнении колеблемости одного и того же признака в нескольких совокупностях представляют интерес показатели вариации, выраженные в относительных величинах. Базой для сравнения служит средняя арифметическая. Эти показатели вычисляются как отношение размаха вариации, среднего линейного отклонения или среднего квадратического отклонения к средней арифметической или медиане.

Чаще всего они выражаются в процентах и определяют не только сравнительную оценку вариации, но и дают характеристику однородности совокупности. Совокупность считается однородной, если коэффициент вариации не превышает 33%. Различают следующие относительные показатели вариации:

1. Коэффициент осцилляции отражает относительную колеблемость крайних значений признака вокруг средней.

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-0gIsPA.png. | (6.10) |

2. Относительное линейное отклонение характеризует долю усредненного значения абсолютных отношений от средней величины.

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-W8Brxh.png. | (6.11) |

3. Коэффициент вариации оценивает типичность средних величин.

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-ATpUOq.png. | (6.12) |

Чем меньше https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-A9VEBB.png, тем однороднее совокупность по изучаемому признаку и типичнее средняя. Еслиhttps://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-2SkexA.png≤33%, то распределение близко к нормальному, а совокупность считается однородной. Из приведенного примера вторая совокупность однородна.

**Виды дисперсий и правило сложения дисперсий.**Наряду с изучением вариации признака по всей совокупности в целом часто бывает необходимо проследить количественные изменения признака по группам, на которые разделяется совокупность, а также и между группами. Такое изучение вариации достигается посредством вычисления и анализа различных видов дисперсии.

При этом можно определить три показателя колеблемости признака в совокупности:

1. Общую вариацию совокупности, которая является результатом действия всех причин. Эта вариация может быть измерена общей дисперсией (https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-MIwkkd.png), характеризующей отклонения индивидуальных значений признака совокупности от общей средней

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-uF2S47.png. | (6.13) |

2. Вариацию групповых средних, выражающих отклонения групповых средних от общей средней и отражающих влияние того фактора, по которому произведена группировка. Эта вариация может быть измерена так называемой межгрупповой дисперсией (δ2)

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-BlcDOD.png, | (6.14) |

где https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-1S_nqe.png- групповые средние, аhttps://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-i_A8zf.png-общая средняя для всей совокупности, иhttps://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-YK_Ogx.png- численность отдельных групп.

3. Остаточную (или внутригрупповую) вариацию, которая выражается в отклонении отдельных значений признака в каждой группе от их групповой средней и, следовательно, отражает влияние всех прочих факторов кроме положенного в основу группировки. Поскольку вариацию в каждой группе отражает групповая дисперсия

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-tB7Bt7.png, | (6.15) |

то для всей совокупности остаточную вариацию будет отражать средняя из групповых дисперсий. Эту дисперсию называют средней из внутригрупповых дисперсий (https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-AG9W4R.png) и рассчитывается она по формуле

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-e35FxF.png. | (6.16) |

Общая вариация признака в совокупности должна определяться как сумма вариации групповых средних (за счет одного выделенного фактора) и остаточной вариации (за счет остальных факторов). Это равенство находит свое выражение в сложении дисперсий

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-uu8x8O.png. | (6.17) |

Это равенство, имеющее строго математическое доказательство, известно, как правило сложения дисперсий.

Правило сложения дисперсий позволяет находить общую дисперсию по её компонентам, когда индивидуальные значения признака неизвестны, а в распоряжении имеются только групповые показатели.

**Коэффициент детерминации.**Правило сложения дисперсии позволяет выявить зависимость результатов от определенных факторов при помощи коэффициента детерминации.

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-6Z9oDW.png, | (6.18) |

Этот коэффициент показывает долю (удельный вес) общей вариации изучаемого признака, обусловленную вариацией группировочного признака.

Корень квадратный из коэффициента детерминации носит название корреляционного отношения (https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-CW4oeY.png):

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-UDkdah.png | (6.19) |

Оно характеризует влияние признака, положенного в основание группировки, на вариацию результативного признака. Корреляционное отношение изменяется в пределах от 0 до 1. Если https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-obXny1.png, то группировочный признак не оказывает влияния на результативный. Еслиhttps://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-KyJ9JY.png, то результативный признак изменяется только в зависимости от признака, положенного в основание группировки, а влияние прочих факторных признаков равно нулю.

**Показатели асимметрии и эксцесса.**В области экономических явлений строго симметричные ряды встречаются крайне редко, чаще приходится иметь дело с асимметричными рядами.

В статистике для характеристики асимметрии пользуются несколькими показателями. Если учесть, что в симметричном ряду средняя арифметическая совпадает по значению с модой и медианой, то наиболее простым показателем асимметрии (https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-IG4Cke.png) будет разность между средней арифметической и модой, т.е.https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-OQF8ob.png=https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-AkXlci.png.

Если (https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-VP_xOa.png)>0, то на графике такой ряд будет иметь вытянутость вправо (правосторонняя асимметрия).

Если (https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-xVFFSD.png)<0, то на графике такой ряд будет иметь вытянутость влево (левосторонняя асимметрия).

Для сравнения асимметрии в нескольких рядах используют относительный показатель, полученный путем деления величины (https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-FAZDB8.png) на среднее квадратическое отклонение, т.е.

|  |  |
| --- | --- |
| Аs=https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-jLhhNk.png. | (6.20) |

Еще один показатель рассчитывается в вариационных рядах для характеристики крутости распределения. Это показатель эксцесса (https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-WP9t8S.png). При одной и той же средней арифметической эмпирический ряд может быть островершинным или низковершинным по сравнению с кривой нормального распределения.

Величину эксцесса рассчитывают по формуле

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-IDXNjo.png. | (6.21) |

Центральный момент четвертого порядка рассчитывается по формуле

|  |  |
| --- | --- |
| https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-Gtyppt.png. | (6.22) |

Если https://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-OStsmd.png>0, то эксцесс считают положительным (распределение островершинно), еслиhttps://studfile.net/html/2706/813/html_X8G5ptDHp2.Pl0p/img-pkteie.png<0, то эксцесс считается отрицательным (распределение низковершинно).