**Методические рекомендации к выполнению типовых заданий расчетной части контрольных работ**

Контрольные работы по статистике включают три ***типовых*** расчетных задания.

**Задание 1.** Исследование структуры совокупности.

**Задание 2.** Выявление наличия корреляционной связи между признаками, установление направления связи, измерение тесноты и силы связи, а также оценка статистической значимости показателя силы связи.

**Задание 3.** Применение метода выборочных наблюдений.

Методика комплексного применения статистических методов при выполнении расчетных заданий излагается в данном разделе с использованием демонстрационного примера.

**Образец**

**выполнения и оформления Заданий 1-3**

**контрольных работ**

При проведении статистического наблюдения за деятельностью коммерческих банков одного из регионов РФ за исследуемый период получены выборочные данные об объеме кредитных вложений и сумме прибыли по 30-ти банкам (выборка 20%-ная, механическая).

В проводимом статистическом исследовании эти банки выступают как ***единицы выборочной совокупности***. ***Генеральную совокупность*** образуют ***все коммерческие банки региона***. Анализируемыми признаками изучаемых единиц совокупности являются ***Объем кредитных вложений***и***Сумма прибыли***банка.

Выборочные данные представлены в табл.1.

Таблица 1

Исходные данные

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Номер  банка  п/п | Объем кредитных вложений,  млн руб. | Сумма прибыли,  млн руб. | Номер  банка  п/п | Объем кредитных вложений,  млн руб. | Сумма прибыли, млн руб. |
| 1 | 150,0 | 45,1 | 16 | 167,1 | 58,0 |
| 2 | 40,0 | 6,2 | 17 | 130,0 | 47,0 |
| 3 | 180,0 | 67,0 | 18 | 171,0 | 64,7 |
| 4 | 88,3 | 27,3 | 19 | 148,3 | 46,2 |
| 5 | 170,0 | 62,5 | 20 | 150,0 | 53,7 |
| 6 | 169,0 | 60,0 | 21 | 180,0 | 67,0 |
| 7 | 70,0 | 16,9 | 22 | 198,1 | 68,0 |
| 8 | 112,0 | 20.9 | 23 | 200,0 | 70,0 |
| 9 | 170,0 | 65,0 | 24 | 211,0 | 80,1 |
| 10 | 93,3 | 16,0 | 25 | 190,0 | 67,7 |
| 11 | 136,4 | 69,0 | 26 | 205,0 | 72,0 |
| 12 | 120,0 | 35,0 | 27 | 225,0 | 84,0 |
| 13 | 135,4 | 53,4 | 28 | 230,0 | 87,0 |
| 14 | 173,0 | 66,2 | 29 | 240,0 | 90,2 |
| 15 | 160,0 | 56,0 | 30 | 230,0 | 85,0 |

**Задание 1**

По исходным данным (табл.1) необходимо выполнить следующее:

1. Построить статистический ряд распределения банков по ***Объему кредитных вложений***, образовав ***четыре*** группы с равными интервалами.
2. Графическим методом и путем расчётов определить значения *моды* и *медианы* полученного ряда распределения.
3. Рассчитать характеристики ряда распределения: *среднюю арифметическую*, *среднее квадратическое отклонение*, *коэффициент вариации*.
4. Вычислить *среднюю арифметическую* по исходным данным (табл. 1.1), сравнить её с аналогичным показателем, рассчитанным в п. 3 для интервального ряда распределения. Объяснить причину их расхождения.

**Сделать выводы** по результатам выполнения Задания 1.

**Выполнение Задания 1**

Целью выполнения данного Задания является изучение состава и структуры выборочной совокупности банков путем построения и анализа статистического ряда распределения банков по признаку *Объем кредитных вложений*.

**1.Построение интервального ряда распределения банков по объему кредитных вложений**

Для построения интервального вариационного ряда, характеризующего распределение банков по объему кредитных вложений, необходимо вычислить ***величину и границы интервалов ряда***.

При построении ряда с равными интервалами величина интервала ***h*** определяется по формуле

**,** (1)

где ****– наибольшее и наименьшее значения признака в исследуемой совокупности,  ***k*-** число групп интервального ряда.

Число групп ***k*** задается в условии задания или рассчитывается по формуле Г.Стерджесса

***k=1+3,322 lg n*,** (2)

где ***n* -** число единиц совокупности.

Определение величины интервала по формуле (1) при заданных *k* = 4, *xma****x*** = 240 млн руб., *xmin* = 40 млн руб.:

****

При *h* = 50 млн руб. границы интервалов ряда распределения имеют следующий вид (табл. 2):

Таблица 2

|  |  |  |
| --- | --- | --- |
| Номер группы | Нижняя граница,  млн руб. | Верхняя граница,  млн руб. |
| 1 | 40 | 90 |
| 2 | 90 | 140 |
| 3 | 140 | 190 |
| 4 | 190 | 240 |

Для построения интервального ряда необходимо подсчитать число банков, входящих в каждую группу (***частоты групп***). При этом возникает вопрос, в какую группу включать единицы совокупности, у которых значения признака выступают одновременно и верхней, и нижней границами смежных интервалов (для демонстрационного примера – это 90, 140, 190 млн руб.). Отнесение таких единиц к одной из двух смежных групп рекомендуется осуществлять ***по принципу полуоткрытого интервала*** **[ ).** Т.к. при этом верхние границы интервалов не принадлежат данным интервалам, то соответствующие им единицы совокупности включаются не в данную группу, а в следующую. В последний интервал включаются и ***нижняя***, и ***верхняя границы***.

Процесс группировки единиц совокупности по признаку *Объем кредитных вложений* представлен во вспомогательной (разработочной) таблице 3 (графа 4 этой таблицы необходима для построения аналитической группировки в Задании 2).

###### Таблица 3

Разработочная таблица для построения интервального ряда распределения и аналитической группировки

|  |  |  |  |
| --- | --- | --- | --- |
| Группы банков по объему кредитных вложений, млн руб. | Номер банка | Объем кредитных вложений, млн руб. | Сумма прибыли,  млн руб. |
| **1** | **2** | **3** | **4** |
| 40 – 90 | 2 | 40,0 | 6,2 |
|  | 7 | 70,0 | 16,9 |
|  | 4 | 88,3 | 27,3 |
| Всего | 3 | 198,3 | 50,4 |
| 90 – 140 | 10 | 93,3 | 16,0 |
|  | 8 | 112,0 | 20,9 |
|  | 12 | 120,0 | 35,0 |
|  | 17 | 130,0 | 47,0 |
|  | 13 | 135,4 | 53,4 |
|  | 11 | 136,4 | 69,0 |
| Всего | 6 | 727,1 | 241,3 |
| 140 – 190 | 19 | 148,3 | 46,2 |
|  | 1 | 150,0 | 45,1 |
|  | 20 | 150,0 | 53,7 |
|  | 15 | 160,0 | 56,0 |
|  | 16 | 167,1 | 58,0 |
|  | 6 | 169,0 | 60,0 |
|  | 5 | 170,0 | 62,5 |
|  | 9 | 170,0 | 65,0 |
|  | 18 | 171,0 | 64,7 |
|  | 14 | 173,0 | 66,2 |
|  | 3 | 180,0 | 67,0 |
|  | 21 | 180,0 | 67,0 |
| Всего | 12 | 1988,4 | 711,4 |
| 190 – 240 | 25 | 190,0 | 67,7 |
|  | 22 | 198,1 | 68,0 |
|  | 23 | 200,0 | 70,0 |
|  | 26 | 205,0 | 72,0 |
|  | 24 | 211,0 | 80,1 |
|  | 27 | 225,0 | 84,0 |
|  | 28 | 230,0 | 87,0 |
|  | 30 | 230,0 | 85,0 |
|  | 29 | 240,0 | 90,2 |
| Всего | 9 | 1929,1 | 704,0 |
| ИТОГО | 30 | 4842,9 | 1707,1 |

На основе групповых итоговых строк «Всего» табл. 3 формируется итоговая табл. 4, представляющая ***интервальный ряд распределения банков по объему кредитных вложений***.

Таблица 4

Распределение банков по объему кредитных вложений

|  |  |  |
| --- | --- | --- |
| Номер группы | Группы банков по объему кредитных вложений, млн руб.,  *х* | Число банков,  *f* |
| 1 | 40 – 90 | 3 |
| 2 | 90 – 140 | 6 |
| 3 | 140 – 190 | 12 |
| 4 | 190 – 240 | 9 |
|  | Итого | 30 |

Помимо частот групп в абсолютном выражении в анализе интервальных рядов используются ещё три характеристики ряда, приведенные в графах 4 - 6 табл. 1.4. Это *частоты групп в относительном выражении*, *накопленные (кумулятивные) частоты Sj*,получаемые путем последовательного суммирования частот всех предшествующих (j-1) интервалов, и ***накопленные частости***, рассчитываемые по формуле .

Таблица 5

Структура банков по объему кредитных вложений

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| № группы | Группы банков по объему кредитных вложений, млн руб. | Число банков, *fj* | | Накопленная  частота,  *Sj* | Накопленная  частоcть, % |
| в абсолютном выражении | в % к итогу |
| **1** | **2** | **3** | **4** | **5** | **6** |
| 1 | 40 – 90 | 3 | 10,0 | 3 | 10,0 |
| 2 | 90 – 140 | 6 | 20,0 | 9 | 30,0 |
| 3 | 140 – 190 | 12 | 40,0 | 21 | 70,0 |
| 4 | 190 – 240 | 9 | 30,0 | 30 | 100,0 |
|  | Итого | 30 | 100,0 |  |  |

**Вывод.** Анализ интервального ряда распределения изучаемой совокупности банков показывает, что распределение банков по объему кредитных вложений не является равномерным: преобладают банки с кредитными вложениями от 140 млн руб. до 190 млн руб. (это 12 банков, доля которых составляет 40%); 30% банков имеют кредитные вложения менее 140 млн руб., а 70% – менее 190 млн руб.

**1.2. Нахождение моды и медианы полученного интервального ряда распределения графическим методом и путем расчетов**

Мода и медиана являются ***структурными средними величинами***, характеризующими (наряду со средней арифметической) центр распределения единиц совокупности по изучаемому признаку.

***Мода Мо***для дискретного ряда – это значение признака, наиболее часто встречающееся у единиц исследуемой совокупности[[1]](#footnote-1). В интервальном вариационном ряду модой приближенно считается ***центральное значение модального интервала*** (имеющего наибольшую частоту). Более точно моду можно определить графическим методом по гистограмме ряда (рис.1).

Рис. 1 Определение моды графическим методом

Конкретное значение моды для интервального ряда рассчитывается по формуле:

 (3)

где ***хМo***– нижняя граница модального интервала,

***h*** –величина модального интервала,

***fMo*** – частота модального интервала,

***fMo-1*** – частота интервала, предшествующего модальному,

***fMo+1***– частота интервала, следующего за модальным.

Согласно табл.1.3 модальным интервалом построенного ряда является интервал 140 – 190 млн. руб., так как его частота максимальна (f3 = 12).

Расчет моды по формуле (3):



**Вывод.** Для рассматриваемой совокупности банков наиболее распространенный объем кредитных вложений характеризуется средней величиной 173,33 млн руб.

***Медиана Ме*** – это значение признака, приходящееся на середину ранжированного ряда. По обе стороны от медианы находится одинаковое количество единиц совокупности.

Медиану можно определить графическим методом по кумулятивной кривой (рис. 2). Кумулята строится по накопленным частотам (табл. 5, графа 5).

Рис. 2. Определение медианы графическим методом

Конкретное значение медианы для интервального ряда рассчитывается по формуле:

, (4)

где ***хМе***– нижняя граница медианного интервала,

***h*** – величина медианного интервала,

– сумма всех частот,

***fМе*** – частота медианного интервала,

***SMе-1*** – кумулятивная (накопленная) частота интервала, предшествующего медианному.

Для расчета медианы необходимо, прежде всего, определить медианный интервал, для чего используются накопленные частоты (или частости) из табл. 5 (графа 5). Так как медиана делит численность ряда пополам, она будет располагаться в том интервале, где накопленная частота ***впервые равна*** полусумме всех частот  или превышает ее (т.е. все предшествующие накопленные частоты меньше этой величины).

В демонстрационном примере медианным интервалом является интервал 140 – 190 млн. руб., так как именно в этом интервале накопленная частота Sj = 21 впервые превышает величину, равную половине численности единиц совокупности (=).

Расчет значения медианы по формуле (4):



**Вывод.** В рассматриваемой совокупности банков половина банков имеют в среднем объем кредитных вложений не более 165 млн руб., а другая половина – не менее 165 млн руб.

**3. Расчет характеристик ряда распределения**

Для расчета характеристик ряда распределения , ***σ*,** ***σ*2**, ***Vσ*** на основе табл. 5 строится вспомогательная табл. 6 (**** – середина j-го интервала).

Таблица 6

Расчетная таблица для нахождения характеристик ряда распределения

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Группы банков по объему кредитных вложений, млн руб. | Середина интервала, | Число банков,  ***fj*** |  |  |  |  |
| **1** | **2** | **3** | **4** | **5** | **6** | **7** |
| 40 – 90 | 65 | 3 | 195 | -95 | 9025 | 27075 |
| 90 – 140 | 115 | 6 | 690 | -45 | 2025 | 12150 |
| 140 – 190 | 165 | 12 | 1980 | 5 | 25 | 300 |
| 190 – 240 | 215 | 9 | 1935 | 55 | 3025 | 27225 |
| Итого |  | 30 | 4800 |  |  | 66750 |

Расчет средней арифметической взвешенной:

 (5)

Расчет дисперсии:

 (6)

Расчет среднего квадратического отклонения:

****

Расчет коэффициента вариации:

 (7)

**Вывод.** Анализ полученных значений показателей  и ***σ*** говорит о том, что средний объем кредитных вложений банков составляет 160 млн руб., отклонение от среднего объема в ту или иную сторону составляет в среднем 47,17 млн руб. (или 29,5%), наиболее характерные значения объема кредитных вложений находятся в пределах от 112,83 млн руб. до 207,17 млн руб. (диапазон ).

Значение **Vσ** = 29,5% не превышает 33%, следовательно, вариация кредитных вложений в исследуемой совокупности банков незначительна и совокупность по данному признаку качественно однородна. Расхождение между значениями , **Мо** и **Ме** незначительно (=160млн руб., **Мо**=173,33млн руб., **Ме**=165млн руб.), что подтверждает вывод об однородности совокупности банков. Таким образом, найденное среднее значение объема кредитных вложений банков (160 млн руб.) является типичной, надежной характеристикой исследуемой совокупности банков.

**4.Вычисление средней арифметической по исходным данным**

Для расчета применяется формула средней арифметической простой:

, (8)

Причина расхождения средних величин, рассчитанных по формулам (8) и (5), заключается в том, что по формуле (8) средняя определяется по фактическим значениям исследуемого признака для всех 30-ти банков, а по формуле (5) средняя вычисляется для интервального ряда, когда в качестве значений признака берутся середины интервалов **** и, следовательно, значение средней будет менее точным (за исключением случая равномерного распределения значений признака внутри каждой группы).

**Задание 2**

По исходным данным табл. 1 с использованием результатов выполнения Задания 1 необходимо выполнить следующее:

1. Установить наличие и характер корреляционной связи между признаками О*бъем* *кредитных вложений* и С*умма* *прибыли*, используя метод аналитической группировки.
2. Оценить тесноту и силу корреляционной связи, используя коэффициент детерминации и эмпирическое корреляционное отношение.

3. Оценить статистическую значимость показателя силы связи.

**Сделать выводы** по результатам выполнения Задания 2.

**Выполнение Задания 2**

Целью выполнения данного Задания является выявление наличия корреляционной связи между факторным и результативным признаками, установление направления связи, оценка тесноты и силы связи.

Факторный и результативный признаки либо задаются в условии задания, либо определяются путем проведения предварительного теоретического анализа. Лишь после того, как выяснена экономическая сущность явления и определены факторный и результативный признаки, приступают к проведению корреляционного анализа данных.

По условию Задания 2 факторным является признак *Объем* *кредитных вложений* **(X)**, результативным – признак С*умма* *прибыли* **(Y)***.*

**1. Установление наличия и характера связи между признаками О*бъем кредитных вложений* и *Сумма прибыли* методом аналитической группировки**

Применение метода аналитической группировки

При использовании метода аналитической группировки строится интервальный ряд распределения единиц совокупности по факторному признаку **Х** и для каждой j-ой группы ряда определяется среднегрупповое значение ****** результативного признака **Y**. Если с ростом значений фактора **Х** от группы к группе ***средние*** значения ****** ***систематически*** возрастают (или убывают), между признаками **X** и **Y** имеет место корреляционная связь.

Используя разработочную таблицу 3, строим аналитическую группировку, характеризующую зависимость между факторным признаком **Х** – О*бъем кредитных вложений*и результативным признаком **Y** –С*умма прибыли*. Макет аналитической таблицы имеет следующий вид (табл. 7):

Таблица 7

Зависимость суммы прибыли банков от объема кредитных вложений

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Номер группы | Группы банков по объему кредитных вложений,  млн руб. | Число банков | Сумма прибыли,  млн руб. | |
| всего | в среднем на один банк |
| 1 |  |  |  |  |
| 2 |  |  |  |  |
| 3 |  |  |  |  |
| 4 |  |  |  |  |
| Итого |  |  |  |  |

Групповые средние значения ****** получаем из таблицы 3 (графа 4), основываясь на итоговых строках «Всего». Построенную аналитическую группировку представляет табл. 8.

Таблица 8

Зависимость суммы прибыли банков от объема кредитных вложений

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Номер группы | Группы банков по объему кредитных вложений,  млн руб.,  *х* | Число банков,  ***fj*** | Сумма прибыли,  млн руб. | |
| всего | в среднем на один банк, |
| **1** | **2** | **3** | **4** | **5=4:3** |
| 1 | 40 – 90 | 3 | 50,4 | 16,800 |
| 2 | 90 – 140 | 6 | 241,3 | 40,217 |
| 3 | 140 – 190 | 12 | 711,4 | 59,283 |
| 4 | 190 – 240 | 9 | 704,0 | 78,222 |
|  | Итого | 30 | 1707,1 | 56,90 |

**Вывод**. Анализ данных табл. 8 показывает, что с увеличением объема кредитных вложений от группы к группе систематически возрастает и средняя прибыль по каждой группе банков, что свидетельствует о наличии прямой корреляционной связи между исследуемыми признаками.

**2. Измерение тесноты и силы корреляционной связи с использованием коэффициента детерминации и эмпирического корреляционного отношения**

Для измерения тесноты и силы связи между факторным и результативным признаками рассчитывают специальные показатели – эмпирический коэффициент детерминации  и эмпирическое корреляционное отношение .

***Эмпирический коэффициент детерминации***  оценивает силу связи, определяя, насколько вариация результативного признака **Y** объясняется вариацией фактора **Х** (остальная часть вариации **Y** объясняется вариацией прочих факторов). Показатель  рассчитывается как доля межгрупповой дисперсии в общей дисперсии по формуле

, (9)

где  – общая дисперсия признака **Y**,

 – межгрупповая (факторная) дисперсия признака **Y**.

Значения показателя  изменяются в пределах . При отсутствии корреляционной связи между признаками **Х** и **Y** имеет место равенство  =**0**, а при наличии функциональной связи между ними - равенство =**1**.

***Общая дисперсия***  характеризует вариацию результативного признака, сложившуюся под влиянием всех действующих на **Y** факторов (систематических и случайных). Этот показатель вычисляется по формуле

, (10)

где ***yi*** – индивидуальные значения результативного признака;

– общая средняя значений результативного признака;

***n*** – число единиц совокупности.

Общая средняя  вычисляется как средняя арифметическая простая по всем единицам совокупности:

 (11)

или как средняя взвешенная по частоте групп интервального ряда:

 (12)

Для вычисления  удобно использовать формулу (11), т.к. в табл. 8 (графы 3 и 4 итоговой строки) имеются значения числителя и знаменателя формулы.

Расчет  по формуле (11):



Для расчета общей дисперсии  применяется вспомогательная таблица 12.

Таблица 12

Вспомогательная таблица для расчета общей дисперсии

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Номер  банка  п/п | Прибыль, млн руб. |  |  |  |
| **1** | **2** | **3** | **4** | **5** |
| 1 | 45,1 | -11,803 | 139,3187 | 2034,01 |
| 2 | 6,2 | -50,703 | 2570,8280 | 38,44 |
| 3 | 67,0 | 10,097 | 101,9427 | 4489,00 |
| 4 | 27,3 | -29,603 | 876,3573 | 745,29 |
| 5 | 62,5 | 5,597 | 31,3227 | 3906,25 |
| 6 | 60,0 | 3,097 | 9,5893 | 3600,00 |
| 7 | 16,9 | -40,003 | 1600,2667 | 285,61 |
| 8 | 20,9 | -36,003 | 1296,2400 | 436,81 |
| 9 | 65,0 | 8,097 | 65,5560 | 4225,00 |
| 10 | 16,0 | -40,903 | 1673,0827 | 256,00 |
| 11 | 69,0 | 12,097 | 146,3293 | 4761,00 |
| 12 | 35,0 | -21,903 | 479,7560 | 1225,00 |
| 13 | 53,4 | -3,503 | 12,2733 | 2851,56 |
| 14 | 66,2 | 9,297 | 86,4280 | 4382,44 |
| 15 | 56,0 | -0,903 | 0,8160 | 3136,00 |
| 16 | 58,0 | 1,097 | 1,2027 | 3364,00 |
| 17 | 47,0 | -9,903 | 98,0760 | 2209,00 |
| 18 | 64,7 | 7,797 | 60,7880 | 4186,09 |
| 19 | 46,2 | -10,703 | 114,5613 | 2134,44 |
| 20 | 53,7 | -3,203 | 10,2613 | 2883,69 |
| 21 | 67,0 | 10,097 | 101,9427 | 4489,00 |
| 22 | 68,0 | 11,097 | 123,1360 | 4624,00 |
| 23 | 70,0 | 13,097 | 171,5227 | 4900,00 |
| 24 | 80,1 | 23,197 | 538,0853 | 6416,01 |
| 25 | 67,7 | 10,797 | 116,5680 | 4583,29 |
| 26 | 72,0 | 15,097 | 227,9093 | 5184,00 |
| 27 | 84,0 | 27,097 | 734,2293 | 7056,00 |
| 28 | 87,0 | 30,097 | 905,8093 | 7569,00 |
| 29 | 90,2 | 33,297 | 1108,6680 | 8136,04 |
| 30 | 85,0 | 28,097 | 789,4227 | 7225,00 |
| Итого | 1707,1 | 1650,197 | 14192,2897 | 111331,97 |

Расчет общей дисперсии по формуле (10):



Общая дисперсия может быть также рассчитана по формуле

,

где  – средняя из квадратов значений результативного признака,

 – квадрат средней величины значений результативного признака.

Для демонстрационного примера





Тогда



***Межгрупповая дисперсия***  измеряет ***систематическую вариацию*** результативного признака, обусловленную влиянием признака-фактора **Х** (по которому произведена группировка). Воздействие фактора **Х** на результативный признак **Y** проявляется в отклонении групповых средних  от общей средней . Показатель  вычисляется по формуле

, (13)

где  –групповые средние,

 – общая средняя,

–число единиц в j-ой группе,

***k*** – число групп.

Для расчета межгрупповой дисперсии  строится вспомогательная таблица 13 При этом используются групповые средние значения  из табл. 8 (графа 5).

Таблица 13

Вспомогательная таблица для расчета межгрупповой дисперсии

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Группы банков по размеру кредитных вложений,  млн руб. | Число банков, | Среднее значение  в группе |  |  |
| **1** | **2** | **3** | **4** | **5** |
| 40 – 90 | 3 | 16,800 | -40,103 | 4824,8320 |
| 90 – 140 | 6 | 40,216 | -16,687 | 1670,6690 |
| 140 – 190 | 12 | 59,283 | 2,380 | 67,9728 |
| 190 – 240 | 9 | 78,222 | 21,319 | 4090,4552 |
| Итого | 30 |  |  | 10653,9291 |

Расчет межгрупповой дисперсии  по формуле (11):



Расчет эмпирического коэффициента детерминации  по формуле (9):

 или 75,1%

**Вывод.** 75,1% вариации суммы прибыли банков обусловлено вариацией объема кредитных вложений, а 24,9% – влиянием прочих неучтенных факторов.

***Эмпирическое корреляционное отношение***  оценивает ***тесноту связи*** между факторным и результативным признаками и вычисляется по формуле

 (14)

Значение показателя изменяются в пределах . Чем ближе значение  к 1, тем теснее связь между признаками. Для качественной оценки тесноты связи на основе  служит шкала Чэддока (табл. 14):

Таблица 14

Шкала Чэддока

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***η*** | 0,1 – 0,3 | 0,3 – 0,5 | 0,5 – 0,7 | 0,7 – 0,9 | 0,9 – 0,99 |
| Характеристика  силы связи | Слабая | Умеренная | Заметная | Тесная | Весьма тесная |

Расчет эмпирического корреляционного отношения  по формуле (14):



**Вывод**. Согласно шкале Чэддока связь между объемом кредитных вложений и суммой прибыли банков является тесной.

**3. Оценка статистической значимости коэффициента детерминации .**

Показатели  и  рассчитаны для выборочной совокупности, т.е. на основе ***ограниченной информации*** об изучаемом явлении. Поскольку при формировании выборки на первичные данные могли иметь воздействии какие-либо случайные факторы, то есть основание полагать, что и полученные характеристики связи ,  несут в себе элемент случайности. Ввиду этого, необходимо проверить, насколько заключение о тесноте и силе связи, сделанное по выборке, будет правомерными и для генеральной совокупности, из которой была произведена выборка.

Проверка выборочных показателей на их неслучайность осуществляется в статистике с помощью ***тестов на статистическую значимость (существенность) показателя.*** Для проверки значимости коэффициента детерминации  служит ***дисперсионный F-критерий Фишера***, который рассчитывается по формуле

,

где n – число единиц выборочной совокупности,

******m – количество групп,

 – межгрупповая дисперсия,

 – дисперсия j-ой группы (j=1,2,…,m),

 – средняя арифметическая групповых дисперсий.

Величина рассчитывается, исходя из правила сложения дисперсий:

,

где  – общая дисперсия.

***Для проверки значимости показателя***  рассчитанное значение F-критерия **Fрасч** сравнивается с табличным **Fтабл** для принятого уровня значимости  и параметров **k1, k2,** зависящих от величин **n и m** : **k1=m-1, k2=n-m.** Величина **Fтабл** для значений , **k1, k2**определяется по таблице распределения Фишера, где приведены ***критические*** (предельно допустимые) величины F-критерия для различных комбинаций значений **, k1, k2.** Уровень значимости  в социально-экономических исследованиях обычно принимается равным 0,05 (что соответствует доверительной вероятности Р=0,95).

Если **Fрасч>Fтабл**, коэффициент детерминации признается ***статистически значимым***, т.е. практически невероятно, что найденная оценка  обусловлена только стечением случайных обстоятельств. В силу этого, выводы о тесноте связи изучаемых признаков, сделанные на основе выборки, можно распространить на всю генеральную совокупность.

Если **Fрасч<Fтабл**, то показатель  считается ***статистически незначимым*** и, следовательно, полученные оценки силы связи признаков относятся только к выборке, их нельзя распространить на генеральную совокупность.

Фрагмент таблицы Фишера критических величин F-критерия для значений =0,05; k1=3,4,5; k2=24-35 представлен ниже:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | k2 | | | | | | | | | | | |
| k1 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 |
| 3 | 3,01 | 2,99 | 2,98 | 2,96 | 2,95 | 2,93 | 2,92 | 2,91 | 2,90 | 2,89 | 2,88 | 2,87 |
| 4 | 2,78 | 2,76 | 2,74 | 2,73 | 2,71 | 2,70 | 2,69 | 2,68 | 2,67 | 2,66 | 2,65 | 2,64 |
| 5 | 2,62 | 2,60 | 2,59 | 2,57 | 2,56 | 2,55 | 2,53 | 2,52 | 2,51 | 2,50 | 2,49 | 2,48 |

Расчет дисперсионного F-критерия Фишера для оценки =75,1%, полученной при =473,0763, =355,1310:

**Fрас**ч

Табличное значение F-критерия при = 0,05:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| n | m | k1=m-1 | k2=n-m | Fтабл (,4, 26) |
| 30 | 4 | 3 | 26 | 2,74 |

**Вывод**: поскольку **Fрасч>Fтабл**, то величина коэффициента детерминации =75,1% ***признается значимой (неслучайной) с уровнем надежности*** **95%** и, следовательно, найденные характеристики связи между признаками *Объем кредитных вложений банков* и *Сумма прибыли банков* правомерны не только для выборки, но и для всей генеральной совокупности банков.

**Задание 3**

По результатам выполнения Задания 1 с вероятностью 0,954 необходимо определить:

1. ошибку выборки средней величины объема кредитных вложений банков и границы, в которых будет находиться генеральная средняя.
2. ошибку выборки доли банков с объемом кредитных вложений 175 млн руб. и выше, а также границы, в которых будет находиться генеральная доля.
3. необходимый объем выборки при заданной предельной ошибке выборки, равной 10 млн руб.

**Выполнение Задания 3**

Целью выполнения данного Задания является определение для генеральной совокупности коммерческих банков региона границ, в которых будут находиться величина среднего объема кредитных вложений банков и доля банков с объемом кредитных вложений не менее 175 млн руб.

**1. Определение ошибки выборки для среднего объема кредитных вложений банков и границ, в которых будет находиться генеральная средняя**

Применение выборочного метода наблюдения всегда связано с ***установлением степени достоверности оценок показателей генеральной совокупности***, полученных на основе значений показателей выборочной совокупности. Достоверность этих оценок зависит от репрезентативности выборки, т.е. от того, насколько полно и адекватно представлены в выборке статистические свойства генеральной совокупности. Как правило, генеральные и выборочные характеристики не совпадают, а отклоняются на некоторую величину **ε**, которую называют ***ошибкой выборки (ошибкой репрезентативности).***

Значения признаков единиц, отобранных из генеральной совокупности в выборочную, всегда случайны, поэтому и статистические характеристики выборки случайны, следовательно, и ошибки выборки также случайны. Ввиду этого принято вычислять два вида ошибок - среднюю  и предельную .

***Средняя ошибка выборки***  - это среднее квадратическое отклонение всех возможных значений выборочной средней от генеральной средней, т.е. от своего математического ожидания M[].

Величина средней ошибки выборки рассчитывается ***дифференцированно*** (по различным формулам) в зависимости от ***вида и способа отбора единиц*** из генеральной совокупности в выборочную.

Для собственно-случайной и механической выборки с бесповторным способом отбора средняя ошибка  выборочной средней  определяется по формуле

, (15)

где  – общая дисперсия выборочных значений признаков,

***N*** – число единиц в генеральной совокупности,

***n*** – число единиц в выборочной совокупности.

***Предельная ошибка выборки***  определяет границы, в пределах которых будет находиться генеральная средняя:

,

, (16)

где – выборочная средняя,

 – генеральная средняя.

Границы  задают ***доверительный интервал генеральной средней***, т.е. случайную область значений, которая с вероятностью **Р** гарантированно содержит значение генеральной средней. Эту вероятность **Р** называют ***доверительной вероятностью*** или ***уровнем надёжности***.

В экономических исследованиях чаще всего используются доверительные вероятности **Р= 0.954, Р= 0.997,** реже **Р= 0,683.**

В математической статистике доказано, что предельная ошибка выборки кратна средней ошибке ***µ*** с ***коэффициентом кратности*** ***t* (**называемым также ***коэффициентом доверия***), который зависит от значения доверительной вероятности **Р.** Для предельной ошибки выборочной средней  это теоретическое положение выражается формулой

 (17)

Значения ***t*** вычислены заранее для различных доверительных вероятностей ***Р*** и ***протабулированы*** (таблицы функции Лапласа **Ф**). Для наиболее часто используемых уровней надежности ***Р*** значения ***t*** задаются следующим образом (табл. 15):

Таблица 15

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Доверительная вероятность ***P*** | 0,683 | 0,866 | 0,954 | 0,988 | 0,997 | 0,999 |
| Значение ***t*** | 1,0 | 1,5 | 2,0 | 2,5 | 3,0 | 3,5 |

По условию демонстрационного примера выборочная совокупность насчитывает 30 банков, выборка 20% механическая, следовательно, ***генеральная совокупность включает 150 банков.*** Выборочная средняя , дисперсия  определены в Задании 1 (п. 3). Значения параметров, необходимых для решения задачи, представлены в табл. 16:

Таблица 16

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Р** | **t** | **n** | **N** |  |  |
| 0,954 | 2 | 30 | 150 | 160 | 2225 |

Расчет средней ошибки выборки по формуле (15):

,

Расчет предельной ошибки выборки по формуле (17):



Определение по формуле (16) доверительного интервала для генеральной средней:

160-15,406160+15,406,

144,594 млн руб. 175,406 млн руб.

**Вывод.** На основании проведенного выборочного обследования коммерческих банков региона с вероятностью 0,954 можно утверждать, что для генеральной совокупности банков средний объем кредитных вложений банка находится в пределах от 144,59 млн руб. до 175,41 млн руб.

**2. Определение ошибки выборки для доли банков с объемом кредитных вложений 175 млн руб. и выше, а также границ, в которых будет находиться генеральная доля**

Доля единиц выборочной совокупности, обладающих тем или иным заданным свойством, выражается формулой

, (18)

где ***m*** – число единиц совокупности, обладающих заданным свойством;

***n*** – общее число единиц в совокупности.

Для собственно-случайной и механической выборки с бесповторным способом отбора предельная ошибка выборки  доли единиц, обладающих заданным свойством, рассчитывается по формуле

, (19)

где ***w*** – доля единиц совокупности, обладающих заданным свойством;

***(1-w)*** – доля единиц совокупности, не обладающих заданным свойством,

***N*** – число единиц в генеральной совокупности,

***n***– число единиц в выборочной совокупности.

Предельная ошибка выборки  определяет границы, в пределах которых будет находиться генеральная доля ***р*** единиц, обладающих заданным свойством:

 (20)

По условию Задания 3 исследуемым свойством является *равенство или превышение объема кредитных вложений банка величины 175 млн руб.*

Число банков с заданным свойством определяется из табл. 3 (графа 3):

*m*=11

Расчет выборочной доли по формуле (18):



Расчет по формуле (19) предельной ошибки выборки для доли:



Определение по формуле (20) доверительного интервала генеральной доли:



0,21  0,524

или

21%  52,4%

**Вывод.** С вероятностью 0,954 можно утверждать, что в генеральной совокупности банков доля банков с объемом кредитных вложений 175 млн руб. и выше будет находиться в пределах от 21% до 52,4%.

**3. Определение необходимого объема выборки с заданным значением допустимой предельной ошибки выборки, равной 10 млн руб.**

Необходимый объем выборки для обеспечения заданной предельной ошибки выборки средней величины изучаемого признака в случае собственно-случайной и механической выборки с бесповторным способом отбора вычисляется по формуле

**** (21)

По условию демонстрационного примера ошибка выборки  не должна превышать 10 млн руб. Параметры ***t, N*** и  известны из решений предыдущих задач.

Расчет необходимой численности выборки по формуле (21):

****

**Вывод.** Для того, чтобы обеспечить для среднего объема кредитных вложений банков предельную ошибку выборки, равную 10 млн руб., необходимо из 150 банков, составляющих генеральную совокупность, отобрать в выборочную совокупность 56 банков.

1. Если в дискретном ряду все варианты встречаются одинаково часто, то в этом случае мода отсутствует. Могут быть распределения, где не один, а два (или более) варианта имеют наибольшие частоты. Тогда ряд имеет две (или более) моды, распределение является бимодальным (или многомодальным),что указывает на качественную неоднородность совокупности по изучаемому признаку. [↑](#footnote-ref-1)