МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ ОРЛОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

ИНСТИТУТ БИЗНЕСА И ПРАВА

ФИНАНСОВО-ЭКОНОМИЧЕСКИЙ ФАКУЛЬТЕТ

Кафедра "Экономика и менеджмент"

А.В. Чернова

И.А. Краснобокая

СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ВЗАИМОСВЯЗИ

СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ И ПРОЦЕССОВ

Методические указания по выполнению лабораторной работы

Дисциплина - "Статистика"

Специальность - 060400 "Финансы и кредит"

060500 "Бухгалтерский учет, анализ и аудит"

060800 "Экономика и управление на предприятиях

туризма и гостиничного хозяйства"

061000 "Государственное и муниципальное управление"

061100 "Менеджмент организации"

061500 "Маркетинг"

351000 "Антикризисное управление"

351200 "Налоги и налогообложение"

Печатается по решению редакционно-издательского совета Орел ГТУ

Орел 2003

Авторы: профессор кафедры экономики и менеджмента, доктор экономических наук А.В. Чернова, старший преподаватель кафедры экономики и менеджмента, кандидат экономических наук И.А. Краснобокая

Рецензент: заведующий кафедрой экономики и менеджмента, профессор, доктор экономических наук С.А. Никитин.

Методические указания по выполнению лабораторной работы содержат рекомендации и задания по установлению степени тесноты и характера направления зависимости между признаками. Предназначены для студентов специальностей 060400 "Финансы и кредит", 060500 "Бухгалтерский учет, анализ и аудит", 060800 "Экономика и управление на предприятиях туризма и гостиничного хозяйства", 061000 "Государственное и муниципальное управление", 061100 "Менеджмент организации", 061500 "Маркетинг", 351000 "Антикризисное управление", 351200 "Налоги и налогообложение" при изучении дисциплины "Статистика".

Редактор.

Технический редактор.

Орловский государственный технический университет

Лицензия ИД №00670 от 05.01.2000 г.

Подписано к печати. .03 г. Формат 60х84 1/16.

Печать офсетная. Уч. изд. л. Усл. печ. л.,. Тираж 300 экз.

Заказ №

Отпечатано с готового оригинал-макета на полиграфической базе ОрелГТУ, 302030, г. Орел, ул. Московская, 65.

Орел ГТУ, 2003

Чернова А.В., Краснобокая И.А., 2003

Содержание

1. Методические указания по выполнению лабораторной работы

2. Пример выполнения лабораторной работы

2.1 Задание на лабораторную работу

3. Порядок выполнения лабораторной работы

Рекомендуемая литература

Приложение

## 1. Методические указания по выполнению лабораторной работы

Наиболее разработанной в теории статистики является методология корреляционно-регрессионного анализа парной корреляции, которая исследует связь между одним признаком-фактором (х) и одним признаком-результатом (у).

В основу выявления и установления аналитической формы связи положено применение в анализе исходной информации математических функций, для чего применяют различного вида уравнения прямолинейной и криволинейной связи.

Это уравнение называется уравнением регрессии (или уравнение парной зависимости).

Например, уравнение парной линейной корреляционной зависимости имеет следующий вид:

, (1)

где ух - теоретические значения результативного признака, полученные по уравнению регрессии;

a0, a1 - коэффициенты (параметры) уравнения регрессии.

Коэффициент парной линейной регрессии а1 показывает изменение результативного признака у под влиянием изменения факторного признака х. Уравнение (1) показывает среднее значение изменения результативного признака у при изменении факторного признака х на одну единицу его измерения, т.е. вариацию у, приходящуюся на единицу вариации х. Знак а1 указывает направление этого изменения.

Параметры уравнения a0, a1 определяют путем решения системы нормальных уравнений, полученной на основе метода наименьших квадратов.

В основу этого метода положено требование минимальности сумм квадратов отклонений фактических данных (уi) от выровненных (yxi):

Σ (уi - yxi) 2 = Σ (уi - а0 - а1⋅хi) 2 → min, (2)

Так, для уравнения парной линейной зависимости система уравнений имеет следующий вид:

 (3)

 (4)

Параметры уравнения прямой будут иметь следующий вид:

 (5)

. (6)

Определив значения а0, а1 и подставив их в уравнение связи , находим значение ух, зависящее только от заданного значения х.

Для прямолинейных зависимостей измерителем тесноты связи между признаками является коэффициент парной корреляции, который рассчитывается по формуле:

, (7)

где - среднее произведение факторного и результативного признака:

; (8)

 - среднее значение факторного признака:

; (9)

 - среднее значение результативного признака:

; (10)

 - среднее квадратическое отклонение результативного признака:

; (11)

 - среднее квадратическое отклонение факторного признака:

. (12)

Квадрат линейного коэффициента корреляции называется линейным коэффициентом детерминации:

r2 = d. (13)

Коэффициент детерминации показывает, какая часть общей вариации результативного признака (y) объясняется влиянием изучаемого фактора (x).

Для получения выводов о практической значимости синтезированных в анализе моделей, показаниям тесноты связи дается качественная оценка. Это осуществляется на основе шкалы Чеддока.

Таблица 1 - Шкала Чеддока

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Показаниятесноты связи | 0,1 - 0,3 | 0,3 - 0,5 | 0,5 - 0,7 | 0,7 - 0,9 | 0,9 - 0,999 |
| Характеристикасилы связи | слабая | умеренная | заметная | высокая | весьмавысокая |

При r = 1 связь является функциональной, при r= 0 связь отсутствует. Если коэффициент корреляции со знаком "+", то связь прямая, если со знаком "-", то связь обратная.

Для практического использования моделей регрессии важна оценка их адекватности, т.е. соответствия фактическим статистическим данным.

Поскольку корреляционно-регрессионный анализ связи между признаками проводится для ограниченной по объему совокупности, то параметры уравнения регрессии, коэффициенты корреляции и детерминации могут быть искажены действием случайных факторов. Чтобы проверить насколько эти показатели характерны для всей генеральной совокупности, не являются ли они результатом стечения случайных обстоятельств, необходимо проверить адекватность построенной статистической модели.

При численности объектов анализа до 30 единиц возникает необходимость проверки значимости (существенности) коэффициента регрессии. При этом выясняют насколько вычисленные параметры характерны для отображения условий: не являются ли полученные значения параметров результатом действия случайных причин.

Значимость параметров простой линейной регрессии осуществляется с помощью t-критерия Стьюдента. При этом вычисляют фактические (расчетные) значения t-критерия:

для параметра а0:

, (14)

где - средне квадратическое отклонение результативного признака

у от выровненных значений уx, которые рассчитываются по уравнению регрессии:

. (15)

для параметра а1:

. (16)

Вычисленные по формулам (13) и (15) значения, сравниваются с критическими tк, которые принимаются согласно данным таблицы Стьюдента с учетом заданного уровня значимости (α) и числа степеней свободы (k = n - 2). В социально-экономических исследованиях уровень значимости α обычно принимают равным 5%, т.е. α = 0,05, что соответствует доверительной вероятности 95%. Параметр признается существенным при условии, если tф > tк. В таком случае практически невероятно, что найденные значения параметров обусловлены только случайными совпадениями.

Показатели тесноты связи, исчисленные по данным сравнительно небольшой статистической совокупности, также могут искажаться действием случайных причин. Это вызывает необходимость проверки их существенности, дающей возможность распространять выводы по результатам выборки на генеральную совокупность.

**Для оценки значимости линейного коэффициента корреляции r применяется t-критерий Стьюдента. При этом определяется фактическое (расчетное) значение критерия (trф):**

, (17)

где n-2 - число степеней свободы при заданном уровне значимости α и объеме выборки n.

Вычисленное значение trфсравнивается с критическим tk, которое берется из таблицы Стьюдента с учетом заданного уровня значимости α и числа степеней свободы k = n - 2.

Если trф > tk, то это свидетельствует о значимости линейного коэффициента корреляции r и существенности связи между признаком-фактором и признаком-результатом.

Поскольку не все фактические значения результативного признака лежат на линии регрессии, более справедливо для записи уравнения корреляционной зависимости воспользоваться следующей формулой:

, (18)

где ε - отражает случайную составляющую вариации результативного признака.

В некоторых случаях рассеяние точек корреляционного поля настолько велико, что для принятия решений в управлении не целесообразно пользоваться уравнением регрессии, так как погрешность в оценке анализируемого показателя будет чрезвычайно велика. Для всей совокупности наблюдаемых значений рассчитывается средняя квадратическая ошибка уравнения регрессии, которая представляет собой среднее квадратическое отклонение фактических значений результативного признака у относительно значений, рассчитанных по уравнению регрессии ух:

. (19)

Среднюю квадратическую ошибку уравнения регрессии Sε сравнивают со средним квадратическим отклонением результативного признака σу. Если Sε < σу, то использование уравнения регрессии в статистическом анализе является целесообразным.

Таким образом, опираясь на оценку существенности параметров уравнения регрессии и значений линейного коэффициента корреляции, а также на основании оценки надежности уравнения регрессии, дают заключение об адекватности построенной регрессионной модели и возможности распространения выводов, полученных по результатам малой выборки на всю генеральную совокупность.

После проверки адекватности, установления точности и надежности регрессионной модели необходимо ее проанализировать, т.е. дать экономическую интерпретацию параметров регрессии.

Для уравнения парной линейной зависимости прежде всего необходимо проверить согласуется ли знак параметра а1 с теоретическими представлениями и соображениями о направлении влияния признака-фактора на результативный признак. Для удобства интерпретации параметра а1 следует использовать коэффициент эластичности:

. (20)

Коэффициент эластичности показывает среднее изменение результативного признака при изменении факторного признака на 1% и вычисляется в% -ах.

Уравнение регрессионной зависимости является базой для расчета прогнозных значений результативного признака, стоящих за пределами изучаемого ряда. Для осуществления прогноза значений результативного признака по уравнению регрессии используют не дискретные (точечные), а интервальные оценки.

Средняя квадратическая ошибка уравнения регрессии дает возможность в каждом отдельном случае с определенной вероятностью указать, что величина результативного признака расположена в определенном интервале относительно значения, вычисленного по уравнению регрессии.

Зная дисперсию результативного показателя у и задаваясь уровнем доверительной вероятности, определяют доверительные границы прогнозного значения результативного признака упрогноз при значении факторного признака хо по формуле:

**, (21)**

**где ухо - дискретная (точечная) оценка прогнозного значения результативного признака у, рассчитанная по уравнению регрессии, при заданном значении факторного признака хо;**

**tα - критерий Стьюдента, который для линейной зависимости определяется в соответствии с уровнем значимости α по распределению Стьюдента с k = n - 2 степенями свободы;**

При практическом использовании уравнения регрессии следует помнить, что экстраполяция, т.е. нахождение прогнозируемых уровней за пределами изучаемого ряда, допускается только тогда, когда существенно не изменяются условия формирования уровней признаков, которые лежат в основе определения параметров уравнения регрессии. В противном случае использование уравнений для составления прогнозов должно быть отвергнуто.

## 2. Пример выполнения лабораторной работы

## 2.1 Задание на лабораторную работу

На основе ранжированных данных о производительности труда и стаже работы двадцати рабочих бригады (таблица) необходимо:

2.1 Установить результативный и факторный признаки.

2.2 Определить наличие и форму корреляционной связи между производительностью труда рабочих бригады и стажем работы.

2.3 Построить на графике поле корреляции и эмпирическую линию корреляционной связи.

2.4 Построить регрессионную модель парной корреляционной зависимости и определить её параметры.

2.5 Построить на графике теоретическую кривую корреляционной зависимости.

2.6 Рассчитать показатели тесноты связи между выработкой рабочего и стажем работы. Дать качественную оценку степени тесноты связи.

2.7 Оценить существенность параметров регрессивной модели и показателей тесноты связи. Дать оценку надёжности уравнения регрессии.

2.8 Дать экспериментальную интерпретацию параметров построенной регрессионной модели.

2.9 На основании регрессионной модели парной зависимости указать доверительные границы, в которых будет находиться прогнозное значение уровня производительности труда рабочего бригады, если стаж его работы составит 10,5 лет при уровне доверительной вероятности 95%.

**Решение:**

Установим результативный и факторный признаки: результативный признак (y) - выработка, факторный (x) - стаж работы, лет.

Определим наличие и форму корреляционной связи между производительностью труда рабочих бригады и стажем работы. Так как увеличение значений признака-фактора влечёт за собой увеличение величины результативного признака. То можно предположить наличие прямой корреляционной связи между выработкой и стажем работы. Проведём группировку работников бригады по признаку-фактору - стажу работы. Результаты оформим в таблицу 2. Сравнив средние значения результативного признака по группам, можно сделать вывод о наличии связи между выработкой и стажем работы. Причём она будет являться прямой, так как рост значений признака фактора влечёт рост средних значений признака результата.

Построим поле корреляции.

Рисунок 1. Поле корреляции

х

8 9 10 11 12

1200

1100

1000

900

800

700

Теоретическая кривая корреляционной зависимости

Построим регрессионную модель парной корреляционной зависимости и определим её параметры: - уравнение парной линейной корреляционной зависимости (регрессионная модель).

→, →

Таблица 2 - Расчётная таблица.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |
| 8 | 800 | 6400 | 640000 | 64 | 789,02 | -1,95 | 3,8025 | 152,5 | 23256,25 | 10,98 | 120,56 |
| 8 | 850 | 6800 | 722500 | 64 |  |  |  | 102,5 | 10506,25 | 60,98 | 3718,56 |
| 8 | 720 | 5760 | 518400 | 64 |  |  |  | 232,5 | 54056,25 | -69,02 | 4763,76 |
| 9 | 850 | 1650 | 722500 | 81 | 872,86 | -0,95 | 0,9025 | 102,5 | 10506,25 | -22,86 | 622,57 |
| 9 | 800 | 7200 | 640000 | 81 |  |  |  | -152,5 | 23256,3 | -72,86 | 5308,57 |
| 9 | 880 | 7920 | 774400 | 81 |  |  |  | -72,5 | 5256,25 | 7,14 | 50,98 |
| 9 | 950 | 8550 | 902500 | 81 |  |  |  | 2,5 | 6,25 | 77,14 | 5950,57 |
| 9 | 820 | 7380 | 672400 | 81 |  |  |  | -132,5 | 17556,25 | -52,86 | 2794,17 |
| 10 | 900 | 9000 | 810000 | 100 | 956,7 | 0,05 | 0,0025 | -52,5 | 2756,25 | -56,7 | 3114,89 |
| 10 | 1000 | 10000 | 1000000 | 100 |  |  |  | 47,5 | 2256,25 | 43,3 | 1874,89 |
| 10 | 920 | 9200 | 846400 | 100 |  |  |  | -32,5 | 1056,25 | -36,7 | 1346,89 |
| 10 | 1060 | 10600 | 1123600 | 100 |  |  |  | 107,5 | 11556,25 | 103,3 | 10670,89 |
| 10 | 950 | 9500 | 902500 | 100 |  |  |  | 2,5 | 6,25 | -6,7 | 44,89 |
| 11 | 900 | 9900 | 810000 | 121 | 1040,54 | 1,05 | 1,1025 | -52,5 | 2756,25 | -140,54 | 975,15 |
| 11 | 1200 | 13200 | 1440000 | 121 |  |  |  | 247,5 | 61256,25 | 159,46 | 25421, 19 |
| 11 | 1150 | 12650 | 1322500 | 121 |  |  |  | 197,5 | 39006,5 | 109,46 | 11981,49 |
| 11 | 1000 | 11000 | 1000000 | 121 |  |  |  | 47,5 | 2256,25 | -40,54 | 1643,49 |
| 12 | 1200 | 14400 | 1440000 | 144 | 1124,38 | 2,05 | 4, 2025 | 247,5 | 6156,25 | 75,62 | 5718,38 |
| 12 | 1100 | 13200 | 1210000 | 144 |  |  |  | 147,5 | 21756,25 | -24,38 | 594,38 |
| 12 | 1000 | 12000 | 1000000 | 144 |  |  |  | 47,5 | 2256,25 | -124,38 | 5470,38 |
| 199 | 19050 | 192310 |  | 2013 | 19050,16 |  | 32,95 |  | 358275 |  | 12969,33 |

Найдём среднее произведение факторного и результативного признака по формуле (8):

.

Рассчитаем средние значение факторного и результативного признака:

факторного по формуле (9):

.

результативного, по формуле (10):

; .

Подставим значения результативного и факторного признака в уравнение парной линейной корреляционной зависимости получим регрессионную модель парной корреляционной зависимости: - регрессионная модель зависимости выработки от стажа работы.

; .

5. Построим на графике теоретическую кривую корреляционной зависимости.

6. Рассчитаем показатели тесноты связи между выработкой рабочего и стажем работы. Для прямолинейных зависимостей измерителем тесноты связи между признаками является коэффициент парной корреляции, который рассчитывается по формуле (7).

Для расчёта коэффициента парной корреляции рассчитаем среднее квадратическое отклонение факторного и результативного признака:

результативного признака, по формуле (11)

 (штук)

факторного признака, по формуле (12)

 (лет)

Подставим полученные значения в формулу (7) рассчитаем показатель тесноты связи:

Дадим качественную оценку степени тесноты связи. Для этого рассчитаем коэффициент детерминации, который показывает какая часть общей вариации результативного признака (y) объясняется влиянием изучаемого фактора (x).

; .

На основе шкалы Чеддока можно сделать вывод о том, что между выработкой т стажем работы существует прямая высокая связь.64% изменения выработки обусловлено изменением стажа работы рабочих.

7. Оценим существенность параметров регрессионной модели и показателей тесноты связи и дадим оценку надёжности уравнения регрессии.

Значимость параметров простой линейной регрессии осуществляется с помощью t-критерия Стьюдента. Рассчитаем значения t-критерия Стьюдента для параметра a0 и a1: для параметра а0, по формуле (14). Для этого рассчитаем средне квадратическое отклонение результативного признака у от выровненных значений уx по формуле (15):

,

для параметра a1 по формуле (16):

Для оценки значимости линейного коэффициента корреляции r применяется t-критерий Стьюдента. При этом определяется фактическое (расчетное) значение критерия (trф). Рассчитаем это значение по формуле (17):

Для всей совокупности наблюдаемых значений рассчитаем среднюю квадратическую ошибку уравнения регрессии по формуле (19):

 (штук).

Так как <, то уравнение регрессии целесообразно и может быть использовано в дальнейшем статистическом анализе.

81,98 < 133,8423.

Так как (фактическое) > (критическое), то значение параметра признаётся существенным, то есть оно не является результатом стечения случайных обстоятельств.

Так как > , то также признаётся существенным.

Так как > , то связь между произвольностью труда и стажем работы признаётся существенной.

8. Дадим экспериментальную интерпретацию параметров построенной регрессионной модели. Так как коэффициент регрессии > 0, то это подтверждает теоретические представления о прямой зависимости между выработкой и стажем работы. Значение = 83,84 шт. можно интерпретировать так: при увеличении стажа на 1 год выработка увеличивается на 83,84 шт.

Рассчитаем коэффициент эластичности по формуле (20), который показывает среднее изменение результативного признака при изменении факторного признака на 1%:

%.

То есть при увеличении стажа на 1% их выработка увеличивается на 0,88%.

9. Укажем доверительные границы, в которых будет находиться прогнозное значение уровня производительности труда рабочего бригады, если стаж его работы составит 10,5 лет при уровне доверительной вероятности 95% по формуле (21):

 штук

Таким образом, с вероятностью 95% можно ожидать, что при стаже работы работника 10,5 лет составит не менее 956 штук и не более 1040 штук.

## 3. Порядок выполнения лабораторной работы

На основе ранжированных данных о производительности труда и стаже работы двадцати рабочих бригады (таблица 2) необходимо:

3.1 Установить результативный и факторный признаки.

3.2 Определить наличие и форму корреляционной связи между производительностью труда рабочих бригады и стажем работы.

Построить на графике поле корреляции и эмпирическую линию корреляционной связи.

Построить регрессионную модель парной корреляционной зависимости и определить ее параметры.

Построить на графике теоретическую кривую корреляционной зависимости.

Рассчитать показатели тесноты связи между выработкой рабочего и стажем работы. Дать качественную оценку степени тесноты связи.

Оценить существенность параметров регрессионной модели и показателей тесноты связи. Дать оценку надежности уравнения регрессии.

Дать экономическую интерпретацию параметров построенной регрессионной модели.

На основании регрессионной модели парной зависимости указать доверительные границы, в которых будет находиться прогнозное значение уровня производительности труда рабочего бригады, если стаж его работы составит 10,5 лет при уровне доверительной вероятности 95%.

Таблица 2 - Распределение рабочих бригады по выработке и стажу работы

|  |  |  |
| --- | --- | --- |
| № п/прабочего | Стаж работы,лет | Выработано изделий одним рабочим, шт.  |
| 1234567891011121314151617181920 | 88899999101010101011111111121212 | 80085072085080088095082090010009201060950900120011501000120011001000 |

## Рекомендуемая литература

1. Гусаров В.М. Статистика: Учеб. пособие для вузов. - М.: ЮНИТИ-ДАНА, 2002. - 463 с.

2. Ефимова М.Р. и др. Практикум по общей теории статистики: Учеб. пособие / М.Р. Ефимова, О.И. Ганченко, Е.В. Петрова. - М.: Финансы и статистика, 2000. - 280 с.

3. Практикум по статистике: Учеб. пособие для вузов / Под ред.В.М. Симчеры / ВЗФЭИ. - М.: ЗАО "Финстатинформ", 1999. - 259 с.

4. Статистика: Учеб. пособие / Под ред. проф. М.Р. Ефимовой. - М.: ИНФРА-М, 2002. - 336 с.

5. Теория статистики: Учебник / Под ред. Г.А. Громыко. - М.: ИНФРА-М, 2000. - 414 с.

6. Теория статистики: Учебник / Под ред.Р.А. Шмойловой. - 2-е изд., доп. и перераб. - М.: Финансы и статистика, 1998. - 576 с.

7. Чернова А.В. Теория статистики: Учебное пособие по курсу "Статистика". Ч.1 (для студентов экономических специальностей) / Орел ГТУ, 1997. - 101 с.

8. Чернова А.В. Статистика промышленности: Учебное пособие по курсу "Статистика". Ч.2 (для студентов экономических специальностей) / Орел ГТУ. Каф. "ЭиМ". - Орел, 1998. - 114 с.

## Приложение

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ОРЛОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

ИНСТИТУТ БИЗНЕСА И ПРАВА

Кафедра "Экономика и менеджмент"

**ЛАБОРАТОРНАЯ РАБОТА**

по дисциплине "Статистика"

на тему: " Статистическое изучение взаимосвязи социально-экономических явлений и процессов"

Выполнил студент: \_\_\_\_\_\_\_\_\_\_

Группа: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Специальность: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Руководитель лабораторной работы: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Оценка защиты: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Орел 2003