Министерство образования Республики Беларусь

Учреждение образования

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

Факультет Информационных технологий и управления

Кафедра Интеллектуальных информационных технологий

**ОТЧЁТ**

по дисциплине «Статистические основы индуктивного вывода»

ИПР

Тема: Построение множественной линейной регрессионной модели

Выполнил: Самута Д. В.

гр. 221703

Проверил: Ефремов А. А.

Минск 2024

**Постановка задачи:** известны статистические данные наблюдений за некоторым

количеством однородных экономических объектов.

**Требуется:**

1. Осуществить выбор факторных признаков для построения регрессионной модели.

2. Построить линейное уравнение регрессии, описывающее зависимость между

факторами и результатом.

3. Оценить качество уравнения регрессии с экономической и математической точки

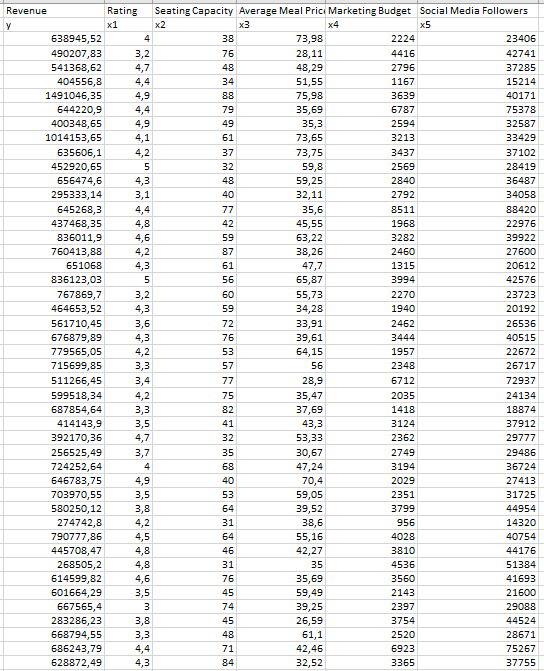
зрения.

4. Найти прогнозное значение результата.

**Описание работы:**

Известны статистические данные ресторанов. Постройте линейное уравнение регрессии, описывающее зависимость между факторами и результатом для предсказания месячной выручки, если прогнозный объем факторных признаков должен составить 115% от их среднего уровня.

**Шаг 1: Таблица с исходными данными:**



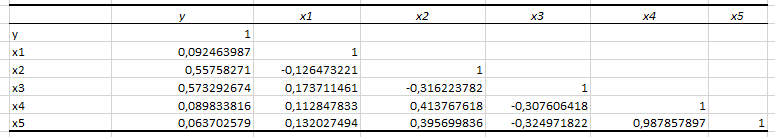
Зависимая переменная: Revenue

Факторами являются:

1. Rating
2. Seating Capacity
3. Average Meal Price
4. Marketing Budget
5. Social Media Followers

**Шаг 1. Осуществить выбор факторных признаков для построения регрессионной модели.** Для подбора факторных признаков используем инструмент КОРРЕЛЯЦИЯ. Вызываемкоманду ДАННЫЕ – АНАЛИЗ ДАННЫХ (рис. 3.2). В меню АНАЛИЗ ДАННЫХ выбираем инструментКОРРЕЛЯЦИЯ.

Результат выполнения команды:



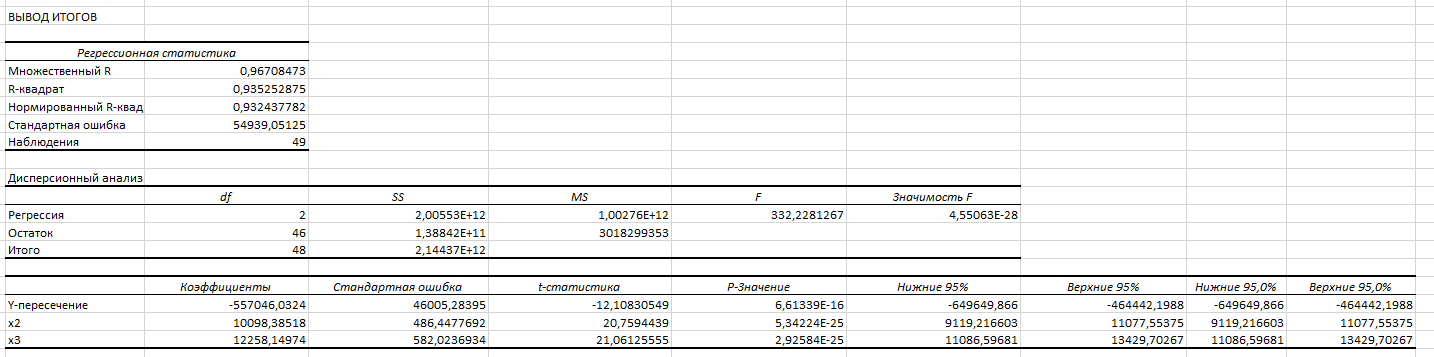
Корреляционная матрица

Анализ коэффициентов парной корреляции: зависимая переменная у имеет

тесную связь с переменной и . Остальные факторы можно исключить из модели.

**Шаг 2. Построить линейное уравнение регрессии, описывающее зависимость между факторами и результатом.**

Регрессия:



Таким образом, мы можем записать полученную модель, описывающую зависимость между выручкой, количество посадочных мест и средней ценой блюд:

**Шаг 3. Оценить качество уравнения регрессии с экономической и математической точки зрения.**

α=0,05 – уровень значимости, n=49 – число наблюдений, число независимых переменных: m=2.

2,013

3,2

Сравним t-статистики с :

Коэффициент = **>**  – значима

Коэффициент (20,8) > – значима

Коэффициент (21) > – значима

Сравним F-статистики:

F = ; F > – уравнение значимо

**Математическая значимость**:

Все коэффициенты (включая константу) статистически значимы, так как их t-статистики превышают ​. Уравнение в целом значимо, так как F-статистика значительно превышает

**Экономическая значимость**:

означает, что 93,5% вариации зависимой переменной y объясняется независимыми переменными ​, что указывает на хорошее качество модели.

**Шаг 4. Рассчитайте эластичность по каждому фактору и ранжируйте факторы по их силе влияния на результативный показатель.**

Средние значения:

56,8

48,45

611443

10098

12258

Таким образом при увеличении x на 1% y увеличится на E%.

Ранжирование факторов по силе влияние: чем выше коэффициент, тем больше влияние данного фактора на результативный показатель. Следовательно наибольше влияние оказывает фактор , затем .

Таким образом можно сделать общий вывод о качестве уравнения и возможности его использования для прогноза:

1) Все коэффициенты в модели статистически значимы (t-значения превышают критическое значение​), а также сама модель в целом статистически значима (F=> =3.2). Это говорит о том, что факторы​ оказывают значительное влияние на результативную переменную y.

2) Коэффициент детерминации показывает, что 93,5% изменчивости зависимой переменной y объясняется изменениями факторов и ​. Это высокий уровень объяснения, особенно для экономических моделей, где коэффициент выше 0.8 считается хорошим результатом.

**Шаг 5. Найти прогнозное значение результата.**

Построенная в примере 3 модель имеет хорошее качество, является адекватной и ее

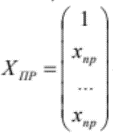
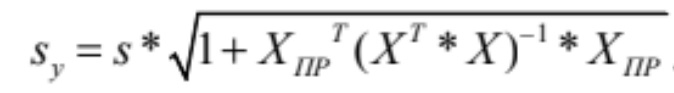
параметры соответствуют экономической теории, следовательно, она может быть использована

для прогноза. Рассчитаем прогноз объема выручки, если прогнозный объем

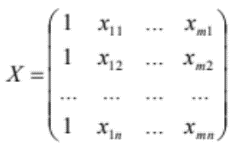
независимых факторов составит 115% от их среднего уровня. Для расчета средних значений

воспользуемся встроенной функцией СРЗНАЧ.

Прогнозный объем выручки составит . Рассчитаем ошибку прогноза. Для этого сформируем таблицу (рис. 3.7), в которой сформируем столбец R64:T112 из единиц и столбцы значений факторов , , а также вектор прогнозных значений AL17:AN17. Ошибка прогноза находится по формуле:

**

- - вектор-столбец прогнозных значений независимых факторов;



* матрица наблюдаемых значений m независимых переменных; s – стандартная ошибка.

Предельная ошибка прогноза:

Дополнительный интервал прогноза:

Произведем необходимые расчеты (см. рис. 3.7). Выделим ячейки AL17:AN17 и вызовем математическую функцию МУМНОЖ. В окошке Массив 1 введем формулу ТРАНСП(R64:T112), а в окошке Массив 2 – диапазон ячеек R64:T112. Далее нажмем сочетание клавиш Сtrl-Shift-ОК, которое всегда используется при работе с массивами. В результате будет рассчитан массив ячеек AK19:AM21. Выделим диапазон ячеек AK23:AM25 и вызовем математическую функцию МОБР. В качестве аргумента в окошке Массив введем адреса ячеек AK19:AM21. Снова воспользуемся сочетанием клавиш Сtrl-Shift-ОК. выделим ячейку AK28 и вызовем функцию МУМНОЖ, в которой в окошке Массив 1 введем диапазон AL17:AN17, а в окошке Массив 2 – формулу МУМНОЖ(AK23:AM25;ТРАНСП(AL17:AN17)). Не забудем про сочетание клавиш Сtrl-Shift-ОК. И, наконец, в ячейке AK30 найдем ошибку прогноза, воспользовавшись формулой=стандартная ошибка \*КОРЕНЬ(1+AK28). В ячейке AK32 рассчитаем предельную ошибку прогноза. В результате получим, что прогнозное значение фактора y находится в пределах 112,545, т.е. между значениями 899,262 и 674172.

