**Федеральное государственное автономное учреждение высшего профессионального образования «Национальный исследовательский университет «Высшая школа экономики»**

МИЭМ НИУ ВШЭ

Кафедра «Кибернетика»

**ОТЧЕТ**

**К ЛАБОРАТОРНОЙ РАБОТЕ №3**

**по дисциплине**

**«Регрессионные модели сложных систем»**

Работу выполнил

студент группы МС-91

А. С. Ковалевский

Москва, 2014

**Часть 1. Первичная статическая обработка.**

Таблица исходных данных выглядит следующим образом:

По результатам построения диаграмм размаха вокруг среднего с размахом в одно стандартное отклонение, можем увидеть ниже, что точки исключать нет необходимости.







После того как обнаружили, что аномальных данных нет можем приступить к построению следующих диаграмм рассеяния:



Видно, что PI и FOOD, PI и CLOT, скорее всего связаны линейной зависимостью. GASOL связана с переменными PI и FOOD чуть более нелинейной чем CLOT. В то же время не похоже, что переменная TAB связана с другими переменными.



**Часть 2. Построение модели множестенной линейной регрессии.**

В качестве зависимой переменной выбираем PI. По итогам множественной регрессии, построенной в пакете STATISTICA, находим значимые переменные и свободный член (выделены красным):





В результате удаления незначимой переменной TOBAC, коэффициенты, свободный член и стандартные ошибки уменьшились, а следовательно модель улучшилась.

Итоговые статистики множественной регрессии:



Исходя из данных таблицы, видим, что модель описывает данные адекватно.

**Часть 3. Анализ адекватности полученной модели.**

Поскольку по полученным параметрам регрессии модель признана адекватной, анализируем остатки.

Распределение остатков приблизительно напоминает нормальное:





Точки распределены хаотично, а следовательно остатки не зависят от предсказанных и наблюдаемых значений





Значение Дарбина-Уотсона находится между dL = 1,12 и dU = 1,66 (n = 25, k = 3 при 5%-ном уровне значимости) –







**Заключение**

Исходя из проведенных исследований, можно считать, что модель адекватна.

Итоговое уравнение регрессии можно записать следующим образом:

**PI = 0.184855 FOOD + 0.545646 CLOT + 0.340840 GASOL – 394.467**

Основные характеристики ее качества:

* R2 = 0.996
* Стандартная ошибка оценки = 15.314