

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

им. М.В.Ломоносова

---

ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И КИБЕРНЕТИКИ

Кафедра оптимального управления

**Соломко Дмитрий Владимирович**

группа 413

ЗАВИСИМОСТЬ СРЕДНЕЙ ПРОДОЛЖИТЕЛЬНОСТИ ЖИЗНИ  
ОТ НЕСКОЛЬКИХ ФАКТОРОВ ПО ДАННЫМ.

Отчёт по исследованию  
по спецкурсу эконометрика

Москва 2009 г.

# Содержание

<b>1 Введение</b>	<b>2</b>
<b>2 Формулировка модели(первая часть исследования)</b>	<b>2</b>
<b>3 Вторая часть исследования</b>	<b>3</b>
<b>4 Вывод</b>	<b>4</b>

## 1 Введение

Интересно было бы изучить зависимости средней ожидаемой продолжительности жизни от разного рода параметров: от ВВП на душу населения, темпов прироста населения по сравнению с предыдущим годом, темпов прироста рабочей силы по сравнению с предыдущим годом, от коэффициента младенческой смертности и от общего показателя ВВП страны.

## 2 Формулировка модели(первая часть исследования)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5.$$

Где  $y$  - средняя продолжительность жизни при рождении, лет

$x_1$  - ВВП в паритетах покупательской способности на душу населения (США берётся за 100, остальные относительно США вычисляются)

$x_2$  - темпы прироста населения по сравнению с предыдущим годом,

$x_3$  - темпы прироста рабочей силы по сравнению с предыдущим годом,

$x_4$  - коэффициент младенческой смертности,

$x_5$  - общий показатель ВВП (трлн. долларов)

Очевидно, что общий объём ВВП никак не должен влиять на продолжительность жизни, так как если сравнивать продолжительность жизни, например, в Китае и в Швеции, то в Швеции она будет больше, несмотря на то, что ВВП Китая не сравним с ВВП Швеции.

Покажем это, используя доверительный интервал,  $t$ -статистику,  $p$ -значение.

**То есть модификацией мы пришли к следующей модели:**

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$$

Где  $y$  - средняя продолжительность жизни при рождении, лет

$x_1$  - ВВП в паритетах покупательской способности на душу населения (США берётся за 100, остальные относительно США вычисляются)

$x_2$  - темпы прироста населения по сравнению с предыдущим годом,

$x_3$  - темпы прироста рабочей силы по сравнению с предыдущим годом,

$x_4$  - коэффициент младенческой смертности.

и в новой модели с помощью доверительного интервала,  $t$ -статистики,  $p$ -значения покажем, что все переменные значимы.

Далее показываем с помощью  $F$ -значения, что уравнение значимо. И то, что во второй модели значение  $F$  получилось больше, чем в первой, то есть вторая модель оказалась лучшей.

Показываем, что  $R_{adj}^2$  для нашей модели достаточно близок к единице.

Далее следует объяснение знаков коэффициентов модели, исходя из здравого смысла.

Затем начертили график остатков. Два ряда - значения  $y$  с помощью таблицы и значения  $y$  с помощью модели.

### **3 Вторая часть исследования**

Так как все регрессоры в той или иной мере характеризуют социальные аспекты общества, можно предположить, что регрессоры как-то связаны друг с другом, поэтому проведем тест на мультиколлинеарность.

Получив в Excel-файле таблицу корреляций, заключаем, что темп прироста рабочей силы и темп прироста населения зависят друг от друга.

Однако удаляя какой-либо из этих регрессоров из модели, мы получаем другую модель с меньшим значением  $R_{adj}^2$ .

Из этого можно сделать вывод о том, что исправление мультиколлинеарности может хуже повлиять на модель. То есть можно в какой-то мере согласиться с теми эконометристами, которые говорят, что ничего с мультиколлинеарностью делать не надо

"так как так устроен мир".

## 4 Вывод

Средняя ожидаемой продолжительности жизни зависит от разного рода параметров: от ВВП на душу населения, темпов прироста населения по сравнению с предыдущим годом, темпов прироста рабочей силы по сравнению с предыдущим годом и от коэффициента младенческой смертности.

Причём от ВВП на душу населения и от темпов прироста рабочей силы средняя ожидаемая продолжительность жизни зависит с положительным коэффициентом, а от темпов прироста населения и коэффициента младенческой смерти - с отрицательным.

Во второй части исследования был проведён тест на мультиколлинеарность и выяснилась зависимость между собой двух регрессоров. Однако при попытке удаления одного из регрессоров с целью избежания мультиколлинеарности мы получили модель, у которой  $R_{adj}^2$ . То есть получили вывод, что исправлять мультиколлинеарность в-общем-то необязательно, так как это исправление не всегда улучшает точность модели.