Задача 1

Необходимо загрузить данные из указанного набора и произвести следующие действия.

Набор данных: Swiss.

Объясняемая переменная: Education.

Регрессоры: Agriculture, Fertility.

1. Оценить среднее значение, дисперсию и СКО переменных, указанных во втором и в третьем столбце.

* Education.
  + Среднее – 10.98.
  + Дисперсия – 92.46.
  + СКО – 9.62.
* Argiculture.
  + Среднее – 50.66.
  + Дисперсия – 515.80.
  + СКО – 22.71.
* Fertility.
  + Среднее – 70.14.
  + Дисперсия – 156.04.
  + СКО – 12.49.

1. Построить зависимости вида y = a + bx, где y – объясняемая переменная, x – регрессор.

Таблица 1.1. Характеристики модели зависимости параметра Education от параметра Fertility в наборе данных Swiss.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Параметр \ Характеристики | Значение | Std. Error | t value | Pr(>|t|) | Уровень значимости |
| (Intercept) | 46.81788 | 6.11244 | 7.659 | 1.08e-09 | \*\*\* |
| Fertility | -0.51095 | 0.08582 | -5.954 | 3.66e-07 | \*\*\* |

Education = .

Судя по полученной модели (см. Таблица 1.1), чем выше рождаемость, тем хуже образование, что действительно похоже на правду – люди, которые получают образование, не имеют времени на создание семьи.

Таблица 1.2. Характеристики модели зависимости параметра Education от параметра Argiculture в наборе данных Swiss.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Параметр \ Характеристики | Значение | Std. Error | t value | Pr(>|t|) | Уровень значимости |
| (Intercept) | 24.69527 | 2.68890 | 9.184 | 6.98e-12 | \*\*\* |
| Argiculture | -0.27076 | 0.04852 | -5.580 | 1.30e-06 | \*\*\* |

Education =

По данной полученной модели (см. Таблица 1.2), чем выше количество занимающихся сельских хозяйством, тем хуже образование, что тоже является правдивыми данными – сельским хозяйством, в основном, занималось крестьянство.

1. Оценить, насколько «хороша» модель по коэффициенту детерминации .

Education =

. Модель плохо объясняет взаимосвязь, необходим ввод новых переменных.

Education =

. Модель плохо объясняет взаимосвязь, необходим ввод новых переменных.

1. Оценить, есть ли взаимосвязь между объясняемой переменной и объясняющей переменной.

Education =

При обоих коэффициентах значение p-статистики определённо является «хорошим» (см. Таблица 1.1), взаимосвязь между объясняемой переменной и регрессором действительно существует. Данная взаимосвязь объясняется следующим образом: чем больше у человека детей, тем меньше он может тратить времени на учебу.

Education =

При обоих коэффициентах значение p-статистики определённо является «хорошим» (см. Таблица 1.2), взаимосвязь между объясняемой переменной и регрессором действительно существует. Данная взаимосвязь объясняется следующим образом: если человек является крестьянином, то он вряд ли может позволить себе обучение.

Вывод

С помощью среды R были построены две модели: Education от Agriculture и Education от Fertility. Обе полученные модели имеют действительно значимые элементы, а значит их можно адекватно описать, но имеют малый , поэтому необходимо вводить дополнительные переменные.

Существует взаимосвязь между Education и Fertility, а также между Education и Argiculture. Так, если у человека много детей, то он не имеет достаточно времени, чтобы тратить ее на учебу, а если человек является крестьянином, то у него недостаточно средств чтобы позволить себе обучение.

Приложение 1

Код задачи

require(stats); require(graphics)

plot(swiss)

head(swiss)

sprintf("Среднее значение: %.2f", mean(swiss$Education))

sprintf("Дисперсия: %.2f", var(swiss$Education))

sprintf("Среднеквадратическое отклонение: %.2f", sd(swiss$Education))

sprintf("Среднее значение: %.2f", mean(swiss$Agriculture))

sprintf("Дисперсия: %.2f", var(swiss$Agriculture))

sprintf("Среднеквадратическое отклонение: %.2f", sd(swiss$Agriculture)) sprintf("Среднее значение: %.2f", mean(swiss$Fertility))

sprintf("Дисперсия: %.2f", var(swiss$Fertility))

sprintf("Среднеквадратическое отклонение: %.2f", sd(swiss$Fertility))

install.packages("lmtest", dependencies=TRUE)

library("lmtest")

model = lm(Education~Fertility, swiss)

summary(model)

model = lm(Education~Agriculture, swiss)

summary(model)