

Частные коэффициенты корреляции между двумя случайными величинами используются для того, чтобы исследовать зависимость между ними при условии устранения влияния другой переменной. Рассмотрим три признака: X_1 – детская смертность в городах США, X_2 – процент афроамериканцев, X_3 – уровень безработицы. Коэффициенты корреляции равны соответственно $r = r_{12} = 0.84$, $r_{13} = 0.32$, $r_{23} = 0.37$ (их можно вычислить в программе VOA.Statistic). Для проверки значимости отклонения от нуля коэффициента корреляции r , вычисленного по парной выборке из n пар, используется статистика Стьюдента вида

$$t = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2},$$

доверительный уровень вероятности можно вычислить в *Excel* при помощи функции

$$p = \text{СТЮДРАСП}(\text{abs}(t); n-2; 2).$$

В данном примере соответствующие p -значения равны $2.45 \cdot 10^{-21}$, 0.004, 0.001. Все они меньше уровня значимости 0.05, поэтому гипотеза о равенстве нулю коэффициента корреляции на генеральном уровне отвергается, и можно сказать, что какая-то зависимость между уровнем детской смертности, с одной стороны, и процентом афроамериканцев с безработицей, с другой. И эта зависимость не может быть объяснена случайностью. Однако между процентом афроамериканцев и безработицей тоже существует значимая корреляция. Не может ли одна зависимость являться следствием другой? Для уточнения этого вычисляются частные коэффициенты корреляции, которые являются корреляциями между двумя признаками при устранении влияния третьего признака (в общем случае в качестве условия может выступать любое количество переменных).

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{23}^2}} = 0.82,$$

$$r_{13.2} = \frac{r_{13} - r_{12}r_{23}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{23}^2}} = 0.03.$$

Значимость отклонения от нуля частного коэффициента корреляции вычисляется также при помощи статистики Стьюдента при уменьшении числа степеней свободы

$$t = \frac{r}{\sqrt{1-r^2}} \sqrt{n-3},$$

доверительный уровень вероятности можно вычислить в *Excel* при помощи функции

$$p = \text{СТЮДРАСП}(\text{abs}(t); n-3; 2).$$

Значимость корреляции между безработицей и детской смертностью ($p = 0.808$) при условии устранения влияния процента афроамериканцев исчезает, а между процентом афроамериканцев и детской смертностью остается даже если бы уровень безработицы не влиял ни на один из этих признаков.