РЕГРЕССИЯ ЛИНЕЙНАЯ МНОЖЕСТВЕННАЯ

Автор: словарь
28.06.2009 14:33 -

РЕГРЕССИЯ ЛИНЕЙНАЯ МНОЖЕСТВЕННАЯ - причинная модель статистической связи линейной (см.) между переменной зависимой (см.) y и переменными независимыми (см.) x1,x2,...,xk, представленная уравнением y = b1x1 + b2x2 + ... + bkxk + a = ∑ bixi + a (см. Анализ регрессионный). Коэффициенты b1,b2,...,bk называются нестандартизированными коэффициентами, a - свободным членом уравнения регрессии. Уравнение регрессии существует также в стандартизированном виде, когда вместо исходных переменных используются их z-оценки (см. Переменная стандартизированная): zy = ∑ βizi. Здесь zy - z-оценка переменной y; z1,z2,...,zk - z-оценки переменных x1,x2,...,xk; β1,β2,...,βk - стандартизированные коэффициенты регрессии (свободный член отсутствует).

Для того чтобы найти стандартизированные коэффициенты, необходимо решить систему линейных уравнений:

β1 + r12β2 + r13β3 + ... + r1kβk = r1y,

r21β1 + β2 + r23β3 + ... + r2kβk = r2y,

r31β1 + r32β2 + β3 + ... + r3kβk = r3y,

...

rk1β1 + rk2β2 + rk3β3 + ... + βk = rky,

в которой rij - коэффициенты линейной корреляции Пирсона (см.) для переменных xi и xj; riy - коэффициент корреляции Пирсона для переменных xi и y.

Нестандартизированные коэффициенты регрессии вычисляются по формуле bi = βi sy / si, где sy - стандартное отклонение (см.) переменной y; si - стандартное отклонение переменной xi. Свободный член уравнения регрессии находится по формуле a = y - ∑ bixi, где y - среднее арифметическое (см.) переменной y, xi - средние арифметические для переменных xi.

В настоящее время используются два подхода к интерпретации нестандартизированных коэффициентов линейной регрессии bi. Согласно первому из них, bi представляет собой величину, на которую изменится предсказанное по модели значение ŷ = ∑ bixi при увеличении значения независимой переменной xi на единицу измерения; согласно второму - величину, на которую в среднем изменяется значение переменной y при увеличении независимой переменной xi на единицу. Значения коэффициентов bi существенно зависят от масштаба шкал, по которым измеряются переменные y и xi, поэтому по ним нельзя судить о степени влияния независимых переменных на зависимую. Свободный член уравнения регрессии a равен предсказанному значению зависимой переменной ŷ в случае, когда все независимые переменные xi = 0.

Стандартизированные коэффициенты βi являются показателями степени влияния независимых переменных xi на зависимую переменную y. Они интерпретируются как "вклад" соответствующей независимой переменной в дисперсию (см.) (изменчивость) зависимой переменной.

Качество (объясняющая способность) уравнения множественной линейной регрессии измеряется коэффициентом множественной детерминации (см.), который равен квадрату коэффициента корреляции множественной (см.) R^2.

Предполагается, что все переменные (см.) в уравнении множественной линейной регрессии являются количественными. При необходимости включить в модель номинальные переменные используется техника dummy-кодирования (см.).

О.В. Терещенко