

УДК 330.4

Жадан М.О.¹, Ткач В.М.¹, Шишканова Г.А.²

¹ студ. гр. БАД-911 НУ «Запорізька політехніка»

² канд. фіз.-мат. наук, доц. НУ «Запорізька політехніка»

РЕГРЕСІЙНІ МОДЕЛІ З ФІКТИВНОЮ ЗМІННОЮ

Обговорюючи моделі багатofакторної регресії, часто припускають, що кожна незалежна змінна є числовою. Однак у багатьох ситуаціях модель необхідно включати категорійні змінні. Наприклад, можна спробувати врахувати моделі розташування товару всередині магазину (наприклад, на вітрині чи ні). Щоб врахувати в регресійній моделі категорійні змінні, слід включити до неї фіктивні змінні. Наприклад, якщо якась категорійна змінна, що пояснює, має дві категорії, для їх представлення достатньо однієї фіктивної змінної X_d : $X_d = 0$, якщо спостереження належить першій категорії, $X_d = 1$, якщо спостереження належить другій категорії.

Для ілюстрації фіктивних змінних розглянемо модель для передбачення середньої оцінної вартості нерухомості Y на основі вибірки з 25 будинків. Як пояснюють змінних виберемо житлову площу будинку (m^2) та наявність каміна. Фіктивна змінна X_2 (наявність каміна) визначається наступним чином: $X_2 = 0$, якщо каміна в будинку немає, $X_2 = 1$, якщо в будинку є каміні.

Після знаходження коефіцієнтів лінійної регресії приходимо до наступних висновків, що якщо фіктивна змінна має незмінне значення, збільшення житлової площі на $10 m^2$ призводить до збільшення передбаченої середньої оцінної вартості на 6,2 тис. умовних грошових одиниць. Якщо житлова площа стала, наявність каміна збільшує середню оцінку вартість будинку на 0,75 тис. умовних грошових одиниць.

Досить часто вважається, що вплив відгуку на змінну, що пояснює, є статистично незалежним від впливу відгуку на інші пояснюючі змінні. Якщо ця умова не виконується, виникає взаємодія між залежними змінними. Наприклад, цілком імовірно, що реклама дуже впливає на обсяг продажу товарів, що мають низьку ціну. Однак, якщо ціна товару надто висока, збільшення витрат на рекламу не може суттєво підвищити обсяг продажу. У цьому випадку спостерігається взаємодія між ціною товару та витратами на його рекламу. Інакше висловлюючись, не можна робити загальних тверджень про залежність обсягу продажу від витрат за рекламу. Вплив рекламних витрат обсяг продажів залежить від ціни. Цей вплив враховується у моделі багатофакторної регресії за допомогою ефекту взаємодії.

У розробленій нами регресійній моделі передбачалося, вплив розміру будинку на його вартість не залежить від того, чи є в будинку каміні. Інакше висловлюючись, вважалося, що відхилення оцінної вартості, що залежить від житлової площі будинку, однаковий у будинків, що мають каміні і не мають його. Якщо ці відхилення відрізняються один від одного, між розміром будинку та наявністю каміна існує взаємодія.

Перевірка гіпотези про рівність відхилень зводиться до оцінки вкладу, який вносить у модель регресії добуток змінної X_1 , що пояснює, і фіктивної змінної X_2 . Якщо цей вклад є статистично значущим, вихідну модель регресії застосовувати не можна. В нашому випадку він виявився несуттєвим.