

В.В. Бабич

*Алматынський Филиал Санкт-Петербурзького Гуманитарного університету профсоюзів,
г. Алматы, Казахстан*

ПРИМЕНЕНИЕ ЭКОНОМЕТРИЧЕСКИХ МЕТОДОВ В ОПЕРАЦИОННОМ ПЛАНИРОВАНИИ

Аннотация

В статье рассматривается прикладной аспект использования статистических моделей в операционной деятельности компании. Автор предлагает статистическую модель, позволяющую определить функцию спроса, на основании которой можно прогнозировать объём продаж и включать прогноз в текущие операционные планы. Автор приводит алгоритм вычислений, позволяющий построить модель, с определенной оценкой качества и достоверными результатами, используя эконометрические методы. Целью модели является отбор независимых факторных признаков, оказывающих наиболее сильное влияние на изменение объёма продаж. Строится модель множественной регрессии. Автор выдвигает гипотезы и с помощью тестов и определенных статистических коэффициентов проверяет их на определенном доверительном интервале. Учитывая, что перед любым предпринимателем стоит задача максимизации прибыли, в определенной конкурентной среде, задача выявления факторов, оказывающих значительное влияние на объем спроса является чрезвычайно важной.

Ключевые слова: экономическая модель, функция спроса, эконометрика, множественная регрессия, мультиколлинеарность, проверка гипотезы, предельный показатель, интервал прогноза, качество модели.

Аңдатпа

В.В. Бабич

Санкт-Петербург кәсіподақтар гуманитарлық университетінің Алматы филиалы, Алматы, Қазақстан
ОПЕРАЦИЯЛЫҚ ЖОСПАРЛАУДА ЭКОНОМЕТРИКАЛЫҚ ӘДІСТЕРДІ ҚОЛДАНУ

Мақалада компанияның операциялық қызметінде статистикалық модельдерді қолданудың қолданбалы аспектілері қарастырылады. Автор сізге сұраныс функциясын анықтауға мүмкіндік беретін статистикалық модель ұсынады, соның негізінде сіз сату көлемін болжай аласыз және болжамды ағымдағы операциялық жоспарларға енгізе аласыз. Автор эконометриялық әдістерді қолдана отырып, белгілі бір сапа бағалары мен сенімді нәтижелері бар үлгіні құруға мүмкіндік беретін есептеу алгоритмін ұсынады. Модельдің мақсаты сатылымдағы өзгерістерге күшті әсер ететін тәуелсіз факторлық сипаттамаларды таңдау болып табылады. Бірнеше регрессиялық модель салынууда. Автор гипотезаларды алға тартады және тесттер мен белгілі бір статистикалық коэффициенттерді қолдана отырып, оларды белгілі бір сенімді аралықта тексереді. Кез келген кәсіпкердің белгілі бір бәсекелестік ортада кірісті ұлғайту міндеті тұрғанын ескере отырып, сұранысқа айтарлықтай әсер ететін факторларды анықтау міндеті өте маңызды.

Түйін сөздер: экономикалық модель, сұраныс функциясы, эконометрика, бірнеше регрессия, көп сызықтық, гипотезаны тексеру, шекті индикатор, болжам интервалы, модель сапасы.

Abstract

APPLICATION OF ECONOMETRIC METHODS IN OPERATIONAL PLANNING

Babich V.V.

Almaty Branch of the St. Petersburg Humanitarian University of Trade Unions, Almaty, Kazakhstan

The applied aspect of the use of statistical models in the operating activities of the company is discussed in the article. The author offers a statistical model that allows you to determine the demand function, based on which you can predict the volume of sales and the forecast to be included in the current operational plans. The author provides a calculation algorithm that allows you to build a model with a certain quality assessment and reliable results using econometric methods. The purpose of the model is the selection of independent factor characteristics that have the strongest effect on changes in sales. A multiple regression model is being constructed. The author puts forward hypotheses and, using tests and certain statistical coefficients, checks them at a certain confidence interval. Given that any entrepreneur is faced with the task of maximizing profits in a certain competitive environment, the task of identifying factors that have a significant impact on demand is extremely important.

Keywords: economic model, demand function, econometrics, multiple regression, multicollinearity, hypothesis testing, marginal indicator, forecast interval, model quality.

Анализ спроса и планирование объема продаж, обеспечивающего компании финансовую устойчивость, является одним из ключевых направлений в разработке операционного годового плана. Для того чтобы вероятность прогноза была более высока, необходимо провести ряд измерений, позволяющих утверждать о достоверности функции спроса и построенных на её основании прогнозов продаж [1]. Цель нашего исследования – разработка регрессионной модели прогнозирования функции спроса. Нами были проанализированы данные по объемам продаж торговой точки, занимающейся реализацией запасных частей для автомобилей в городе Алматы. Данные собраны ежемесячно за период 1 полугодия 2020 года. При анализе взаимосвязи объема продаж и факторов, влияющих на него, мы использовали линейную модель множественной регрессии. Согласно эконометрической практике «использование методологии от общего к конкретному, позволяет привести создание модели множественной регрессии с одним уравнением авторегрессионного запаздывания» [2, с.248]. Для построения модели множественной регрессии функции спроса были рассмотрены следующие переменные:

Зависимая переменная: Y – объём продаж в день (тенге).

Независимые факторные признаки:

X_1 – возраст автомобиля (лет),

X_2 – проходимость в день через торговую точку (человек),

X_3 – тип запасной части (оригинальная и не оригинальная 1 или 0),

X_4 – средний чек (тенге).

Для построения функции спроса и модели прогноза объема продаж, на основе множественной регрессии использовался следующий алгоритм. На первом этапе проводился корреляционный анализ, позволяющий оценить уровень качества модели в целом. Далее изучалось наличие факторов мультиколлинеарности. Мультиколлинеарность – это коррелированность двух или нескольких объясняющих переменных в уравнении регрессии [3]. После проверки наличия взаимосвязи между объясняющими факторами и зависимой переменной, проводилась корректировка модели и строилась модель прогнозирования. На следующем этапе оценивалось качество полученной модели. В заключении на основании модели определялся точечный и интервальный прогноз продаж.

Проведенный корреляционный анализ позволил определить наличие связи между зависимым фактором и объясняющими переменными. В модели определялась связь между объемом продаж (зависимая переменная) и количеством посетителей магазина в день, возрастом автомобиля, видом запасных частей и средним чеком в день (независимые переменные). Практическое значение корреляционного анализа это выявление причинно-следственной связи между факторными признаками и результивным признаком [4]. На примере данных, взятых для исследования, мы определили корреляционную матрицу и величину коэффициента корреляции. Коэффициент корреляции характеризует направление и силу связи между признаками. Рассмотрим полученный коэффициент корреляции, который приведен в таблице 1.

Таблица 1. Регрессионная статистика

Коэффициент корреляции	0,839724
Коэффициент детерминации	0,705136
Нормированный коэффициент детерминации	0,696888
Стандартная ошибка	23805,6
Число наблюдений	148

В рассматриваемой нами выборке данных коэффициент корреляции равен 0,8397. Данный коэффициент, согласно шкале близок к 1, что говорит о наличии положительной и высокой связи между объемом продаж (зависимая переменная) и факторными признаками: возраст автомобиля, проходимость человек в день, оригинальность запасной части и средний чек.

Зная, что необходимо исполнение допущения о независимости факторных признаков, мы провели тест Фаррара-Глоубера в три этапа [5]. На первом этапе проверили наличие мультиколлинеарности всего массива данных. На втором этапе определили мультиколлинеарность между объясняющими переменными. На третьем этапе проверили парную мультиколлинеарность.

Для проверки наличия мультиколлинеарности всех факторов была построена матрица межфакторных корреляций (таблица 2) и по определителю матрицы проверена гипотеза о наличии или об отсутствии мультиколлинеарности.

Таблица 2. Матрица межфакторных корреляций

Факторы	Объем продаж	Возраст авто	Проходимость в день чел	Оригинал/не оригинал	Средний чек
объем продаж	1,000	0,189	0,409	-0,176	0,793
возраст авто	0,189	1,000	0,051	-0,835	0,357
проходимость в день чел	0,409	0,051	1,000	-0,021	0,197
оригинал/не оригинал	-0,176	-0,835	-0,021	1,000	-0,346
средний чек	0,793	0,357	0,197	-0,346	1,000

Определитель: 0,073754

Согласно полученному результату, определитель близок к нулю, что позволяет сделать вывод о наличии мультиколлинеарности. Подтвердим это предположение оценкой статистики Фаррара-Глоубера. Вычислим наблюдаемое значение статистики Фаррара – Глоубера по формуле:

$$FG = -[n - 1 - \frac{1}{6} * (2 - k + 5)] - \ln(\det[R])$$

Оценка статистики Фаррара– Глоубера дала следующий результат:

$$FG = -[148 - 1 - \frac{1}{6}(2 - 4 + 5)] - \ln 0,073 = 381,928$$

Табличное значение статистики χ^2 равно 12,59. Так как $FG > \chi^2$ ($381,9 > 12,59$), то в массиве независимых переменных имеет место мультиколлинеарность. Вывод о наличии мультиколлинеарности потребовал проверки гипотезы о наличии мультиколлинеарности между объясняющими факторами [6].

Согласно выше изложенному алгоритму, была определена матрица обратную корреляционной матрице, она приведена в таблице 3. И на основании данных обратной матрицы рассчитаны F-критерии.

$$F_{ij} = (c_{jj} - 1) * \frac{(n - k - 1)}{k}$$

Таблица 3. Обратная матрица

3,3914	0,2226	-0,8803	-0,1530	-2,6484
0,2226	3,3893	-0,1234	2,7226	-0,4211
-0,8803	-0,1234	1,2728	-0,0687	0,4676
-0,1530	2,7226	-0,0687	3,3583	0,3238
-2,6484	-0,4211	0,4676	0,3238	3,2704

Вычисление F-критериев (таблица 5) с заданными степенями свободы $v_1=4$, $v_2 = n-k-1 = 148-4-1 = 143$ и уровнем значимости $\alpha=0,05$ дало результаты, приведенные в таблице 5. Следует отметить, что уровень значимости определяет вероятность ошибки первого рода. Ошибку первого рода часто называют ложной тревогой, ложным срабатыванием.

Таблица 4. Значения F-критериев

Критерий	Возраст авто	Проходимость в день чел	Оригинал/не оригинал	Средний чек
F	85,419	9,753	84,310	81,167

Сравнивая фактические значения F-критериев с табличным значением этого критерия $F_{таб} = 2,37$, мы пришли к следующим выводам. Согласно данным таблицы 4, все значения F-критериев выше табличного, что позволяет сделать вывод о наличии мультиколлинеарности между всеми объясняющими факторами. В рассматриваемой модели наиболее влияющими на мультиколлинеарность факторами являются: возраст автомобиля, оригинальность запасных частей, средний чек. Наименее влияющим является фактор проходимости потенциальных покупателей через торговую точку. Поскольку имеет место зависимость между объясняющими переменными, следует определить наличие связи по парам переменных. Вычисленные частные коэффициенты корреляции приведены в таблице 5.

Таблица 5. Матрица коэффициентов частных корреляций

Факторы	Возраст авто	Пройодимость в день чел	Оригинал/не оригинал	Средний чек
возраст авто		-0,066	0,424	0,045
проходимость в день чел	-0,066		0,059	-0,807
оригинал/не оригинал	0,424	0,059		0,033
средний чек	0,045	-0,807	0,033	

Чтобы сделать выводы о парной мультиколлинеарности были определены t-критерии (таблица 6).

Таблица 6. T-критерии для коэффициентов частной корреляции

t_{12}	-0,78695
t_{13}	5,59331
t_{14}	0,54254
t_{23}	0,71153
t_{24}	-16,34032
t_{34}	0,39737

Сравнивая фактические значения t-критериев с табличным значением равным 1,96 при степени свободы 143 и уровне значимости $\alpha=0,05$, мы пришли к следующему заключению.

Две пары факторов:

- 1) возраст автомобиля и оригинальность запасных частей;
- 2) проходимость в торговой точке и средний чек имеют высокую статистически значимую корреляцию.

Поэтому в целях устранения мультиколлинеарности исключим одну из этих пар объясняющих переменных. В паре факторов возраст автомобиля и оригинальность запасной части, достаточно высокий F- критерий у фактора возраст, следовательно, эта пара более сильно влияет на мультиколлинеарность выбранных объясняющих переменных. Поэтому мы удалили фактор возраст автомобиля из модели прогнозирования.

Таким образом, в результате проверки теста Фаррара-Глоубера в модели остается три объясняющие переменные: средний чек, оригинальность запасной части и проходимость торговой точки в день. Коэффициенты модели множественной регрессии приведены в таблице 7.

Таблица 7. Параметры множественной регрессии

Параметры	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение
Y- среднее значение	-39698,5	7079,61	-5,607	1,01E-07
проходимость в день чел	4307,5	772,43	5,576	1,17E-07
оригинал/не оригинал	8762,09	4315,35	2,03	0,04
средний чек	5,76	0,36	15,72	4,5E-33

Согласно расчетам, получена следующая модель множественной регрессии:

$$Y = -39\,698,5 + 4\,307,51X_1 + 8\,762 X_2 + 5,76 X_3$$

Логическая проверка характера взаимосвязи между зависимой переменной (объем продаж) и влияющими на него факторами (проходимость, оригинальность, средний чек) позволила заключить следующее. Если на объем продаж не влияет ни один из факторных признаков, то объем продаж в день равен – 39 698 тенге, что не логично экономически, поскольку отрицательным объем продаж быть не может. Таким образом, константа не может использоваться при планировании фактических объемов продаж. Коэффициенты, стоящие перед факторными признаками X_1 , X_2 , X_3 характеризуют предельные показатели, позволяющие определить насколько изменится зависимый показатель при изменении факторного признака на единицу. В полученной нами модели наиболее сильно на объем продаж влияет тип запасной части (оригинальная или не оригинальная). Если запасная часть оригинальная, то объем продаж возрастает на 8 762 тенге в день. Достаточно сильное влияние оказывает проходимость человек в день. Увеличение проходимости на 1 человека в день ведет к увеличению объема продаж на 4 307,5 тенге в день. Изменение среднего чека также влияет на объем продаж в день. При увеличении среднего чека на 1000 тенге объем продаж в день возрастет на 5 760 тенге. Все факторные признаки имеют прямую связь с зависимой переменной, что соответствует экономической сущности рассматриваемых данных. Для достоверности оцененной модели и возможности её использования, была проведена оценка качества модели. Для оценки качества рассчитаны коэффициенты регрессионной статистики, которые приведены в таблице 8.

Таблица 8. Регрессионная статистика

Коэффициент корреляции	0,83896303
Коэффициент детерминации	0,70385896
Нормированный коэффициент детерминации	0,69768936
Стандартная ошибка	23774,1133
Число наблюдений	148

Основным показателем, оценивающим качество регрессионной модели и величину взаимосвязи между зависимой и независимыми факторами, является коэффициент детерминации. Он характеризует долю вариации зависимой переменной от факторов, включенных в модель и факторов, оставшихся за пределами модели [7].

В целом рассматриваемая модель значима, доля объяснённой дисперсии в объясняемой составляет 70,38%. Однако, так как коэффициент детерминации ниже 80% мы не можем сказать о высоком качестве модели. В рассматриваемой модели 30% факторов, влияющих на объем продаж, осталось за пределами модели. Поэтому маркетологам, определяющим спрос, необходимо изучить другие факторы, оказывающие значительное влияние на изменение динамики спроса. Мерой для оценки включения факторного признака в модель служит коэффициент Фишера. Таблица дисперсионного анализа (таблица 9), показывает значимость последовательного добавления к уравнению регрессии соответствующего фактора.

Таблица 9. Дисперсионный анализ

Источники вариации	df	SS	MS	F	Значимость F
Регрессия	3	1,93445E+11	64481770025	114,084933	7,18225E-38
Остаток	144	81390018886	565208464,5		
Итого	147	2,74835E+11			

Сравнивая расчетное значение коэффициента Фишера $F=114$ с табличным значением при 5% уровне значимости и степенями свободы $\nu_1=3$ (количество факторных признаков) и $\nu_2=148-3-1=144$, проверим, является ли значимой совместная объясняющая способность трех независимых переменных, то есть проверим нулевую гипотезу: $H_0: \beta_1: \beta_2: \beta_3=0$. Отметим, что нулевая гипотеза соответствует состоянию «по умолчанию» естественному, наиболее ожидаемому положению вещей.

В данном случае мы предполагаем, что объясняющие переменные не влияют на зависимую переменную.

В рассматриваемой модели F критическое (табличное), равное 2,6 меньше F расчетного 114, следовательно, нулевая гипотеза об отсутствии влияния независимых переменных отвергается и принимается альтернативная гипотеза, которая указывает на значимый уровень объяснения. В целом модель на 95% доверительном интервале может применяться. С помощью t-теста оценим значимость каждого факторного признака. Аналогично примем нулевые гипотезы об отсутствии влияния, то есть предельные коэффициенты в модели равны нулю: $H_0: \beta_1: \beta_2: \beta_3=0$.

Критический уровень t-статистики при 5% уровне значимости и степенями свободы $n-4$ ($148-4=144$) равен 1,96. Расчетные t-статистики приведены в таблице 10. Мы видим, что все расчетные значения t-статистики больше $|1,96|$. Следовательно, нулевая гипотеза об отсутствии влияния каждого факторного признака на результирующий показатель отвергается и принимается альтернативная гипотеза о влиянии. Таким образом, корректировка модели позволила определить достоверную модель для прогнозирования продаж с учетом 95% доверительного интервала

Таблица 10. Регрессионная статистика

Признаки	Стандартная ошибка	t-статистика
Y-пересечение	7079,615914	-5,607443136
проходимость в день чел	772,4375029	5,576599035
оригинал/не оригинал	4315,350576	2,030448966
средний чек	0,366477028	15,72730066

Используя полученное уравнение множественной регрессии,

$$Y = -39\,698,5 + 4\,307,51X_1 + 8\,762 X_2 + 5,76 X_3$$

был определён точечный и интервальный прогноз объема продаж в день. Уравнение регрессии отражает только общую тенденцию для выборки, поэтому мы ввели допущение, что

X_1 – проходимость возрастет с 7 человек в день до 10 (на 3 человека);

X_2 – оригинальность запчастей, в 80% будут использоваться оригинальные запаные части, возрастет на 0,2;

X_3 – средний чек с учетом инфляции возрастет на 10% или на 900 тенге.

При таких допущениях, объём продаж в день возрастет на 19 858 тенге:

$$\Delta Y = 4307,51 * 3 + 8762 * 0,2 + 5,76 * 900 = 19\,858.$$

Если не считать константу верной (модель не подтвердила её достоверность), а взять среднюю текущую величину продаж в день 48 556 тенге, то точечный прогноз объема продаж составит $48\,556 + 19\,858 = 68\,414$ тенге в день. Для определения интервального прогноза объёма продаж была использована стандартная ошибка. Стандартная ошибка в модели равна 23 774 (таблица 8). Критический уровень t-статистики равен $|1,96|$, в таком случае 95% доверительный интервал прогноза объема продаж составит:

$$48\,556 - 1,96 * 23\,774 < Y < 48\,556 + 1,96 * 23\,774$$

$$1958 < Y < 95153$$

Используя результаты анализируемой модели, мы можем дать следующие рекомендации компании.

Во-первых, следует обращать внимание на динамику трех основных факторов: проходимость торговой точки, оригинальность запасных частей и средний чек. Во-вторых, наиболее влияющим фактором на динамику объёма продаж в стоимостном выражении является фактор оригинальность запасной части. Поэтому компании следует провести четкую дифференциацию ценовой политики в соответствии с целевой аудиторией. В-третьих, 30% факторов, влияющих на объём продаж, остались за пределами модели, что снижает качество модели и требует доработки данных для построения

новой модели лучшего качества. В-четвертых, средний объем продаж в день может изменяться в диапазоне от 1 958 тенге до 95 153 тенге, что необходимо учитывать при планировании выручки компании.

Список использованной литературы:

- 1 Молчанов Н.Н., Пецольт К. Выбор метода прогнозирования объема продаж малого предприятия. [Электронный ресурс] – URL: <https://pure.spbu.ru> (дата обращения 19.08.2020)
- 2 Wojciech W. Charemza. Defek F. New directions in econometric practice: general to specific modeling, cointegration and vector autoregression. Biddles Limited. 1997. 435 с.
- 3 Мультиколлинеарность и методы ее устранения. [Электронный ресурс] – URL: <https://moscow-stud.com> (дата обращения 19.08.2020)
- 4 Множественная регрессия. [Электронный ресурс] – URL: <http://mylektsii.ru> (дата обращения 19.08.2020)
- 5 Эконометрическое моделирование стоимости автомобиля Toyota на вторичном рынке. [Электронный ресурс] – URL: <https://top-technologies.ru> (дата обращения 19.08.2020)
- 6 Выбор экзогенных факторов в модель регрессии при мультиколлинеарности данных. [Электронный ресурс] – URL: <http://applied-research.ru> (дата обращения 19.08.2020)
- 7 Оценка значимости множественной регрессии. [Электронный ресурс] – URL: <https://economy-ru> (дата обращения 19.08.2020)

References

- 1 Molchanov N. N., Pecol'dt K. (2020) Vybor metoda prognozirovaniya ob'ema prodazh malogo predpriyatija [Choosing a method for predicting the sales volume of a small enterprise]. [Jelektronnyj resurs]. URL: <https://pure.spbu.ru>. (In Russian)
- 2 Wojciech W. Charemza. Defek F. (1997) New directions in econometric practice: general to specific modeling, cointegration and vector autoregression. Biddles Limited. 435. (In English)
- 3 Mul'tikollinearnost' i metody ee ustraneniya (2020) [Multiple regression]. [Jelektronnyj resurs]. URL: <https://moscow-stud.com> (In Russian)
- 4 Mnozhestvennaja regressija (2020) [Multiple regression]. [Jelektronnyj resurs]. URL: <http://mylektsii.ru> (In Russian)
- 5 Jekonometricheskoe modelirovanie stoimosti avtomobilja Toyota na vtorichnom rynke (2020) [Econometric modeling of the cost of a Toyota car in the secondary market]. [Jelektronnyj resurs]. URL: <https://top-technologies.ru> (In Russian)
- 6 Vybor jekzogennyh faktorov v model' regressii pri mul'tikollinearnosti dannyh (2020) [Selection of exogenous factors in the regression model with multicollinearity of data]. [Jelektronnyj resurs]. URL: <http://applied-research.ru> (In Russian)
- 7 Ocenka znachimosti mnozhestvennoj regressii (2020) [Assessing the significance of multiple regression]. [Jelektronnyj resurs]. URL: <https://economy-ru> (In Russian)