

3.2. ПРОГНОЗИРОВАНИЕ ПРОИЗВОДСТВА МЕТРОВАГО- НОВ МЕТОДАМИ ЭКОНОМЕТРИКИ

Галочкин В.Т., к.ф.-м.н., доцент
кафедры «Моделирование экономических
и информационных систем»

Финансовый университет при
Правительстве РФ

[Перейти на Главное МЕНЮ](#)
[Вернуться к СОДЕРЖАНИЮ](#)

В статье предложена линейная множественная модель, предназначенная для оценки объемов производства метровагонов на Мытищинском заводе «Метровагонмаш». Проверена статистическая значимость качества и надежности полученного уравнения и коэффициентов регрессии. На примере предложенной модели выполнен прогноз объемов производства метровагонов в 2013 г.

Машиностроительный завод в Мытищах был основан в 1897 г. для производства железнодорожных пассажирских и товарных вагонов для Транссибирской магистрали. В 1903 году началось производство трамвайных вагонов. С 1934 года одной из основных продукций завода стали метровагоны. На учредительной конференции трудового коллектива объединения 20 апреля 1992 г. было принято решение о преобразовании завода в открытое акционерное общество (ОАО) «Метровагонмаш». С 2005 г. ОАО «Метровагонмаш» входит в состав группы «Трансмашхолдинг», крупнейшей в Российской Федерации компании на рынке транспортного машиностроения. Предприятия группы выпускают маневровые и магистральные тепловозы, промышленные и магистральные электровозы, грузовые и пассажирские вагоны, вагоны метро и электропоездов, тепловозные и судовые дизели, стрелочные переводы, тяговые агрегаты, вагонное литье и другую продукцию.

Линии метрополитена в Москве и других городах интенсивно развиваются, требуется все больше и больше

метровагонов. Нынешние мощности двух поставщиков вагонов метро – ОАО «Метровагонмаш» и Завода им. Егорова – с трудом смогут справиться с планируемым объемом заказов без масштабной реконструкции, считают в мэрии Москвы. «Московский метрополитен планирует ежегодно закупать до 400 новых метровагонов, чтобы к 2020 г. полностью обновить весь подвижной состав», – сообщил мэр столицы С. Собянин. (Подробнее см.: Москва будет ежегодно закупать до 400 новых вагонов метро – С.С. Собянин. URL: /inmsk.ru/transport_news/20130603/3641_9162.html).

Прогноз выпуска метровагонов в 2013 г. опирается на результат математического решения аналитического уравнения, построенного по данным годовых отчетов ОАО «Метровагонмаш». Все данные, приведенные в табл. 1, взяты из официальных годовых отчетов ОАО «Метровагонмаш» (подробнее см. www.tmholding.ru/investor/quarterly_report/) и Закрытого акционерного общества (ЗАО) «Трансмашхолдинг» (подробнее см. www.metrowagonmash.ru/for_investors/all/, с 2005 г. ОАО «Метровагонмаш» входит в ЗАО «Трансмашхолдинг»).

В табл. 1 представлены:

- колонка 1 – год выпуска метровагонов;
- колонка 2 – количество выпущенных метровагонов;
- колонка 3 – общая численность работающих в ОАО «Метровагонмаш»;
- колонка 4 – численность производственных рабочих;
- колонка 5 – доля работников, не имеющих профессионального образования, %;
- колонка 6 – доля работников, имеющих среднее профессиональное образование (профессионально-техническое училище, техникум, колледж), %;
- колонка 7 – доля работников, имеющих высшее образование, включая кандидатов наук, %;
- колонка 8 – доля работников в возрасте до 30 лет, %;
- колонка 9 – доля работников в возрасте 30-60 лет (женщин до 55 лет), %;
- колонка 10 – доля работников пенсионного возраста, %.

Таблица 1

ИСХОДНЫЕ ДАННЫЕ

Год	Вагоны, шт.	Численность, чел.		Образование, %			Возраст, %		
		Всего	Произв. рабочие	До 10 кл.	Проф.	Высшее	До 30 лет	До 60 лет	Свыше 60 лет
1	2	3	4	5	6	7	8	9	10
2005	102	5 250	2 295	62,2	19,5	18,3	22	69	9
2006	232	5 180	2 274	61	20,5	18,5	23	69	9
2007	258	5 110	2 244	58,5	22	19,5	21	68,5	10,5
2008	284	5 090	2 235	56	23,00	21,00	20	68	12
2009	285	4 630	2 033	50	24	26	18,5	67,9	13,6
2010	293	4 053	1 781	51,8	23,4	24,8	17	68	15
2011	356	4 558	1 999	49	26,9	24,1	18	68	14
2012	503	4 463	1 951	46,9	26,8	26,3	19	68	13

Таблица 2

ДАнные для РАСЧЕТА

у	x ₁	x ₂	x ₃	x ₄	x ₅	x ₆	x ₇
1	2	3	4	5	6	7	8
102	1	5,250	2,295	0,195	0,183	0,69	0,09
232	2	5,180	2,274	0,205	0,185	0,69	0,09
258	3	5,110	2,244	0,22	0,195	0,685	0,105
284	4	5,090	2,235	0,23	0,21	0,68	0,12
285	5	4,630	2,033	0,24	0,26	0,679	0,136
293	6	4,053	1,781	0,234	0,248	0,68	0,15
356	7	4,558	1,999	0,269	0,241	0,68	0,14
503	8	4,463	1,951	0,268	0,263	0,68	0,13

Первоначально была попытка построить уравнение множественной линейной регрессии, учитывая все известные из годовых отчетов нефинансовые характеристики работы ОАО «Метровагонмаш». Данные приведены в табл. 2.

В табл. 2 обозначены:

- x_1 – год выпуска метровагонов (с 2005 по 2012 гг.);
- x_2 – общая численность работающих в ОАО «Метровагонмаш», тыс. чел.;
- x_3 – численность производственных рабочих, тыс. чел.;
- x_4 – доля сотрудников имеющих среднее профессиональное образование, %;
- x_5 – доля сотрудников, имеющих высшее образование, включая кандидатов наук, %;
- x_6 – доля сотрудников в возрасте от 30 до 60 лет, %;
- x_7 – доля сотрудников пенсионного возраста, %;
- y – количество выпущенных метровагонов в отчетном году.

Попытка построить уравнение линейной множественной регрессии $y = f(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$ не увенчалась успехом. Это означает сильную линейную зависимость столбцов матрицы экзогенных (предопределенных) переменных или, что матрица $(X^T X)^{-1}$ не имеет полный ранг. Здесь $X_{6,7}$ – матрица столбцов всех объясняющих переменных. Ее определитель $\det(X^T X)^{-1}$ стремится к нулю. Т.е. в первоначально предложенной модели присутствует совершенная мультиколлинеарность. (Мультиколлинеарность – линейная взаимосвязь двух или более экзогенных переменных).

Чтобы устранить в предлагаемой модели совершенную мультиколлинеарность, были объединены в одну группу данные о сотрудниках, имеющих образование (колонки 6 и 7 табл. 1), и в другую группу – сотрудники старше 30 лет (колонки 9 и 10 табл. 1). Результаты такого объединения представлены в табл. 3.

Таблица 3

ПЕРЕСТРОЕННЫЕ ДАННЫЕ ДЛЯ РАСЧЕТА

у	x_1	x_2	x_3	x_4	x_5
102	1	5,25	2,295	0,378	0,78
232	2	5,18	2,274	0,39	0,78
258	3	5,11	2,244	0,415	0,79
284	4	5,09	2,235	0,44	0,8
285	5	4,63	2,033	0,5	0,815
293	6	4,053	1,781	0,482	0,83
356	7	4,558	1,999	0,51	0,82
503	8	4,463	1,951	0,531	0,81

В табл. 3 обозначены:

- x_1 – год выпуска метровагонов (с 2005 по 2012 гг.);
- x_2 – общая численность работающих в ОАО «Метровагонмаш», тыс. чел.;
- x_3 – численность производственных рабочих, тыс. чел.;
- x_4 – доля сотрудников, имеющих среднее и высшее образование, %;
- x_5 – доля сотрудников старше 30 лет, %;
- y – количество выпущенных вагонов в отчетном году.

Строилось уравнение множественной линейной регрессии $y = f(x_1, x_2, x_3, x_4, x_5)$. Все ниже выполненные расчеты и выводы относятся к данным, приведенным в табл. 3.

Для решения задачи построения модели множественной линейной регрессии использовался пакет «анализ данных» из программы Microsoft Excel.

Шаг 1

По методу наименьших квадратов (МНК) получили уравнение множественной линейной регрессии:

$$\hat{y} = 5839,39 + 67,75x_1 - 2792,61x_2 + 6197,85x_3 + 413,23x_4 - 7079,53x_5 \quad (1)$$

Полученная модель (1) имеет высокие величины множественного коэффициента корреляции:

$$R(y, x_1, x_2, x_3, x_4, x_5) = \sqrt{1 - \frac{\sum (y_i - \hat{y}(x_1, x_2, x_3, x_4, x_5))^2}{\sum (y - y_{cp})^2}} = 0,994, \text{ здесь } y_{cp} = (\sum y_i) / 8,$$

и коэффициента детерминации $R^2 = Q_R / Q = 0,988$, здесь:

$$Q = \sum (y - y_{cp})^2, \quad Q_R = Q - Q_e, \quad Q_e = \sum (y_i - \hat{y}(x_1, x_2, x_3, x_4, x_5))^2.$$

Все формулы для расчетов взяты из [3, с. 82-106].

Скорректированный коэффициент детерминации (зависит от числа объясняющих переменных) также высок:

$$\check{R}^2 = 1 - (n - 1) * (1 - R^2) / (n - p - 1) = 0,958,$$

здесь

n – число наблюдений (восемь наблюдений, соответствует числу лет с 2005-2012 гг., по которым приведены данные в табл. 3),

p – число экзогенных переменных (пять переменных – x_1, x_2, x_3, x_4 и x_5).

Статистическая значимость уравнения, оцененная с помощью F -критерия Фишера:

$$F = ((R^2 / (1 - R^2)) * (n - p - 1) / p) = 33,184,$$

также высока.

Полученная модель в целом имеет хорошие показатели коэффициентов качества и значимости.

Посмотрим значимость каждого коэффициента регрессии. Для того чтобы полученная модель множественной линейной регрессии адекватно описывала действительные процессы, необходимо, чтобы коэффициенты регрессии были значимы и имели правильное экономическое толкование.

Статистическую значимость t каждого коэффициента оцениваем по формуле:

$$t = b_i / S_{b_i},$$

здесь

b_i – i -коэффициент уравнения регрессии,

S_{b_i} – дисперсия оцениваемого коэффициента.

Если наблюдаемое значение статистики по модулю больше критического, исследуемый коэффициент статистически значим.

Получена следующая статистическая значимость коэффициентов:

$$|t_{b1}| = 4,20; |t_{b2}| = 2,42; |t_{b3}| = 2,41; |t_{b4}| = 0,56; |t_{b5}| = 3,56.$$

Критическое значение статистики Стьюдента для доверительной вероятности $\gamma = 0,95$ на первом шаге $t_{кр} = 4,3$. Такое высокое значение критической величины объясняется малым объемом выборки (восемь групп наблюдений) и большим числом объясняющих переменных (пять переменных). Наименее значима переменная x_4 – доля сотрудников имеющих среднее и высшее образование. Посмотрим матрицу корреляции эндогенной и экзогенных переменных; приведена в табл. 4.

Таблица 4

КОРРЕЛЯЦИОННАЯ МАТРИЦА ЭНДОГЕННОЙ И ЭКЗОГЕННЫХ ПЕРЕМЕННЫХ УРАВНЕНИЯ (1)

Показатель	у	x_1	x_2	x_3	x_4	x_5
у	1	0,9110	-0,6116	-0,6145	0,8561	0,5896
x_1	0,9110	1	-0,8335	-0,8361	0,9717	0,8500
x_2	-0,6116	-0,8335	1	0,9997	-0,8210	-0,9389
x_3	-0,6145	-0,8361	0,9997	1	-0,8233	-0,9353
x_4	0,85619	0,9717	-0,8210	-0,8233	1	0,8653
x_5	0,58964	0,8500	-0,9389	-0,9353	0,8653	1

Видно, что между всеми объясняющими переменными наблюдается сильная линейная связь. Знаки у объясняющих переменных соответствуют объективной реальности. Для улучшения уравнения регрессии (1) исключим из дальнейшего рассмотрения переменную x_4 , имеющую наименьшую статистику (0,56) и наибольший коэффициент корреляции с объясняющей переменной x_1 . Метод исключения объясняющих переменных подразумевает последовательное по одной исключение объясняющих переменных, пока статистика оставшихся коэффициентов станет больше критической. После исключения x_4 рассматриваем набор значений объясняющих переменных, представленный в табл. 5.

Таблица 5

ИСХОДНЫЕ ДАННЫЕ ПОСЛЕ УДАЛЕНИЯ ПЕРЕМЕННОЙ x_4

y	x_1	x_2	x_3	x_5	\hat{y}
102	1	5,25	2,295	0,78	104,1055
232	2	5,18	2,274	0,78	240,327
258	3	5,11	2,244	0,79	258,9978
284	4	5,09	2,235	0,8	269,2619
285	5	4,63	2,033	0,815	273,2832
293	6	4,053	1,781	0,83	291,9914
356	7	4,558	1,999	0,82	382,4306
503	8	4,463	1,951	0,81	492,6026

Шаг 2

По МНК получили уравнение множественной линейной регрессии:

$$\hat{y} = 5489,39 + 75,76x_1 - 2602,55x_2 + 5796,22x_3 - 6538,48x_5. \quad (2)$$

Модель (2) имеет высокие величины множественного коэффициента корреляции

$$R(y, x_1, x_2, x_3, x_5) = 0,993$$

и коэффициента детерминации $R^2 = 0,986$. Скорректированный коэффициент детерминации также высок $\tilde{R}^2 = 0,967$.

Статистическая значимость уравнения, оцененная с помощью F -критерия Фишера $F = 53,528$. Полученная модель (2) имеет хорошие показатели коэффициентов качества и значимости.

Посмотрим значимость каждого коэффициента регрессии. Получена следующая статистическая значимость коэффициентов

$$|t_{b1}| = 11,24;$$

$$|t_{b2}| = 2,68;$$

$$|t_{b3}| = 2,87;$$

$$|t_{b5}| = 4,27.$$

Видно, что в пределах ошибки на шаге 2 исследования все коэффициенты статистически значимы. Посмотрим матрицу корреляции эндогенной и экзогенных переменных, которая приведена в табл. 6.

Таблица 6

КОРРЕЛЯЦИОННАЯ МАТРИЦА ЭНДОГЕННОЙ И ЭКЗОГЕННЫХ ПЕРЕМЕННЫХ УРАВНЕНИЯ (2)

Показатель	y	x_1	x_2	x_3	x_5
y	1	0,9110	-0,6116	-0,6145	0,5896
x_1	0,9110	1	-0,8335	-0,8361	0,8500
x_2	-0,6116	-0,8335	1	0,9997	-0,9389
x_3	-0,6145	-0,8361	0,9997	1	-0,9353
x_5	0,5896	0,8500	-0,9389	-0,9353	1

Из табл. 6 видно, что, несмотря на статистическую значимость коэффициентов уравнения регрессии, в модели все еще присутствует мультиколлинеарность.

Так как задача, поставленная автором, состоит в прогнозировании количества выпускаемых в 2013 г. метровагонов при неизменных производственных ресурсах и численности работающих, то при больших коэффициентах детерминации и множественной корреляции и статистической значимости уравнения (2) мультиколлинеарность не сказывается на прогнозных качествах полученного уравнения [2, с. 85]. (Если бы целью исследования было изучение влияния каждой из экзогенных переменных на эндогенную, то пренебрегать мультиколлинеарностью было бы нельзя.)

Оценим прогноз производства метровагонов на 2013 г. по полученному уравнению (2). (Обычно годовые отчет предприятия готов не раньше апреля следующего за отчетным года. Т.е. на момент написания этой статьи официальные результаты работы в 2013 г. ОАО «Метровагонмаш» еще не известны.)

По нашим оценкам, эта величина составит примерно 569 вагонов метро при среднем значении ошибки аппроксимации 9,46%, что удовлетворяет требованиям эконометрических расчетов. При эконометрических расчетах средняя ошибка аппроксимации не должна превышать 10%.

Представим исходные данные производства метровагонов и полученные по уравнению множественной линейной регрессии на графике (рис. 1). Расчетные значения получены по уравнению (2) и представлены в колонке 6 табл. 5. На графике также отражено прогнозное значение на 2013 г. – 569 метровагонов.

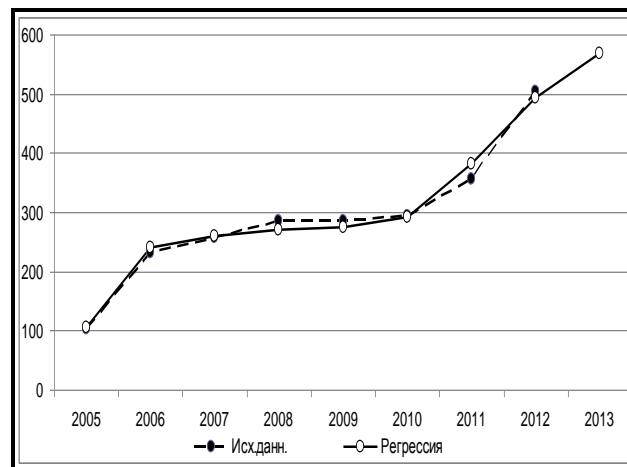


Рис. 1. Расчетные данные (горизонтальная ось – год выпуска метровагонов, вертикальная ось – количество метровагонов в отчетном году, включая прогноз производства метровагонов в 2013 г.)

На представленном графике видна хорошая корреляция исходных данных и величин, рассчитанных по уравнению множественной линейной регрессии (2).

Интерпретация результатов. Мерой эффективности множественной линейной регрессионной модели являются множественный коэффициент детерминации ($R^2 = 0,986$) и корреляции ($R(y, x_1, x_2, x_3, x_5) = 0,993$). Полученное нами уравнение характеризуется весьма высокой статистической значимостью $F = 53,528$. На основании величины показателя множественной корреляции (0,993) делаем вывод в соответствии со шка-

лой Чеддока [1, с. 119] о высокой взаимосвязи количества произведенных метровагонов y с факторами:

- x_1 – год выпуска метровагонов (с 2005 по 2012 гг.);
- x_2 – общая численность работающих в ОАО «Метровагонмаш», тыс. чел.;
- x_3 – численность производственных рабочих, тыс. чел.;
- x_5 – доля сотрудников старше 30 лет, %.

На основании значения множественного коэффициента детерминации утверждается, что на 98,6% факторные показатели (x_1, x_2, x_3, x_5) объясняют вариацию результивного показателя (y). Принято считать, что если $R^2 > 0,95$, это говорит о высокой точности модели, модель хорошо описывает реальный производственный процесс.

Посмотрим значимость регрессионной модели. Статистическая значимость F в нашем случае имеет вероятность $p = 0,004$. Общепринято, высказывания, имеющие вероятность ошибки $p < 0,05$, называются значимыми; высказывания, имеющие вероятность ошибки $p < 0,01$, – очень значимыми, а высказывания, имеющие вероятность ошибки $p < 0,001$, – максимально значимыми. Вывод – построенная модель очень значима.

Относительно знаков коэффициентов уравнения (2). Знаки уравнения соответствуют экономической реальности. Увеличение общего числа сотрудников не увеличивает производство метровагонов, скорее, увеличивая численность административно-управленческого персонала, уменьшаем выпуск метровагонов. Так как нас интересует производственный процесс, увеличение доли старшего поколения работающих замедляет ускорение выпуска метровагонов.

ВЫВОДЫ

Полученные в статье результаты сформулированы для конкретных исходных данных. При других исходных данных возможны несколько отличные результаты и, возможно, выводы. Если, несмотря на значимость оставленных коэффициентов, попробовать выполнять дальнейшие шаги по удалению переменных, это может привести к тому, что удаленной окажется очень важная переменная, значительно влияющая на производство метровагонов, что может существенно снизить прогнозную точность модели.

Кроме данной модели могут быть получены и другие. Например, если на 1-м шаге принять решение об удалении не переменной x_4 , а переменной x_2 , которая очень сильно коррелирует с переменной x_3 . Но в силу их сильной коррелированности модель получится близкой по точности и значимости. Автором была проделана такая работа. В этом случае прогноз на 2013 г. – 568 вагонов, что в пределах средней ошибки аппроксимации.

Галочкин Валерий Тимофеевич

Литература

1. Гладиллин А.В. и др. Практикум по эконометрике [Текст] / А.В. Гладиллин, А.Н. Герасимов, Е.И. Громов. – Ростов н/Д : Феникс, 2011.
2. Ежеманская С.Н. Эконометрика [Текст] / С.Н. Ежеманская. – Ростов н/Д : Феникс, 2003.
3. Кремер Н.Ш. Эконометрика [Текст] / Н.Ш. Кремер, Б.А. Путко. – М. ЮНИТИ, 2010.

Ключевые слова

Эконометрика; линейная регрессия; метод наименьших квадратов; множественная линейная регрессия; факторы регрессии; корреляция; прогнозирование экономических явлений; метровагон.

РЕЦЕНЗИЯ

Актуальность проблемы. Рецензируемая статья предлагает прогноз производства метровагонов Мытищинским заводом «Метровагонмаш» в 2013 г. Прогноз опирается на результат математического решения аналитического уравнения, построенного по данным годовых отчетов Открытого акционерного общества (ОАО) «Метровагонмаш». За базу автор взял восемь лет – с 2005 по 2012 гг..

Актуальность поставленных автором задач обусловлена проблемой внедрения и практического использования в реальной экономике математических методов эконометрического прогнозирования. Теоретической основой методов эконометрического прогнозирования являются математические дисциплины (прежде всего теория вероятностей и математическая статистика, дискретная математика и исследование операций), а также экономическая теория, экономическая статистика и другие социально-экономические науки. Учет нежелательных тенденций, выявленных при прогнозировании, позволяет принять необходимые меры для их предупреждения на стадии бизнес-планирования.

Среди различных моделей прогнозирования наиболее употребительными на практике являются регрессионные модели: линейные, нелинейные и системы одновременных уравнений. Автор, опираясь на данные годовых отчетов ОАО «Метровагонмаш», строит уравнение множественной линейной регрессии, исследует его качественные и количественные характеристики и дает прогноз объемов производства метровагонов на 2013 г. Фактические данные по 2013 г. ОАО «Метровагонмаш» появятся в открытом доступе только к середине текущего года.

Научная новизна и практическая значимость. Выполненная автором оценка работы конкретного производства встречается в экономической литературе нечасто. Это делает работу новой, актуальной и приближенной к реальной жизни. Безусловно, предложенная автором модель может обладать определенной схематичностью, поскольку в ней не учтены все параметры производства, так или иначе влияющие на итоговый прогноз. Следует добавить также некоторую сложность в получении статистических данных по определенной отрасли, сегменту экономики или конкретному предприятию. Однако при всех недостатках таких моделей принятие решений на их основе в конечном итоге создает основу для выбора наилучшего экономического решения на предприятии, оптимальной тактики и стратегии его руководства.

Заключение. Статья написана грамотным математическим языком, приведенный список литературы адекватно отражает содержание статьи.

Рассматриваемая статья может быть рекомендована для опубликования в журнале «Аудит и финансовый анализ».

Бельчук А.И., д.э.н., проф. кафедры «Мировая и национальная экономика» Всероссийской академии внешней торговли

[Перейти на Главное МЕНЮ](#)
[Вернуться к СОДЕРЖАНИЮ](#)