Л.В. Дегтярева,
Ю.А. Семеняченко

Интеграция математики, информатики и маркетинга в процессе подготовки бакалавров бизнес-информатики

В статье обоснована необходимость интегрированного подхода при обучении математике, информатике и маркетинга в подготовке бакалавров бизнес-информатики и приведен пример реализации такого подхода.

Ключевые слова: маркетинг; информационно-коммуникационные технологии; бакалавр бизнес-информатики; статистические методы; дисперсионный анализ.

Современное предпринимательство, независимо от своих размеров, все больше вынуждено втягиваться в конкурентную борьбу, чтобы отстоять и закрепить свои позиции на рынке. В силу этих причин современный маркетинг, как деятельность, нацеленная на управление продвижением товаров и услуг от производителя к потребителю, играет ключевую роль в экономическом успехе или неуспехе каждой организации.

Маркетинговые исследования могут ответить на многие вопросы, возникающие в повседневной жизни организации либо появляющиеся в период стратегического планирования ее развития.

Безусловно, современный рынок требует от маркетинговых исследований качественных результатов, которые можно получить, используя творческий подход, проявляя инициативу и осваивая нетрадиционные пути исследования с акцентом на анализ большого объема данных. А все это имеет шанс реализоваться у специалистов различных видов бизнеса только в том случае, если они получат хорошую базовую подготовку в учебном заведении.

Цель данной статьи — показать возможность интегрированного подхода в обучении математике, информатике и маркетингу бакалавров бизнес-информатики на примере использования статистических методов и при обработке информации в маркетинговом анализе.

Тенденция использования различных методик анализа информации, собранной в ходе проведения маркетинговых исследований, стала возможной благодаря развитию рынка информационных технологий, который на сегодняшний момент предлагает широкий выбор статистических программных пакетов.

Программа Microsoft Excel является самым доступным инструментом для предпринимательства любых масштабов, поскольку может быть установлена практически на любом компьютере, между тем пользуются ею всего около 4 % представителей бизнеса. И связано это прежде всего с тем, что даже продвинутый пользователь, получив машинный результат анализа, не может его применить, так как не умеет его правильно интерпретировать.

Для подготовки бакалавров направления «Бизнес-информатика», по знаниям и умениям соответствующих современным требованиям рынка, профессионально разбирающихся в тонкостях результатов маркетингового анализа, полученных с помощью программных продуктов, следует рассматривать интеграцию математики, информатики и маркетинга.

Одним из путей реализации такого подходя является включение в процесс обучения бакалавров статистических методов обработки результатов наблюдений при проведении маркетинговых исследований. Такими статистическими методами могут служить корреляционный, регрессионный, дисперсионный анализ и т.п. Приведем пример такой интеграции на основе анализа достаточно распространенной ситуации, возникающей в бизнесе. Зачастую требуется выяснить, влияет ли, и если влияет, то в какой степени, какой-либо качественный фактор, присутствующий в работе организации, на результаты ее деятельности.

Известно, что влияние на результат бизнеса нередко оказывают самые разные качественные факторы, примерами которых могут быть торговая марка организации, дизайн продукции, сорт растений, свойства сырья, время года, погодные условия и т. п. Они не измеряются количественно, но их влияние может оказаться очень существенным. Для наиболее качественного и полного анализа сути таких влияний применяется методика дисперсионного анализа, реализация вычислений для которой может быть осуществлена с помощью программы MS Excel.

Итак, *дисперсионный анализ* — это статистический метод, предназначенный для оценки влияния конкретных мероприятий (факторов) в производственной, торговой, инвестиционной, сервисной или других хозяйственных сферах деятельности фирмы на изменение ее экономических показателей, а также для последующего планирования ее деятельности.

По числу факторов, влияние которых исследуется, различают однофакторный и многофакторный дисперсионный анализ.

Однофакторный дисперсионный анализ определяет влияние одного качественного явления (фактора) на изменение результативных количественных показателей хозяйственной деятельности организации. Также можно произвести количественную оценку степени этого влияния.

Все данные для анализа определяются опытным путем, т. е. берутся из текущей хозяйственной деятельности организации и группируются в виде матрицы наблюдений (табл. 1). Здесь $\varPhi$ — это исследуемый фактор, зафиксированный в своих разных вариациях, а x_{ji} — это показатель хозяйственной деятельности организации, соответствующий каждой вариации исследуемого фактора при многократном измерении.

Таблица 1

Матрица наблюдений однофакторного дисперсионного анализа

Номер измерения показателя при каждой вариации фактора	**Вариации исследуемого фактора $\varPhi$**				
	$\boldsymbol{\varPhi_1}$	$\boldsymbol{\varPhi_2}$	**…**	$\boldsymbol{\varPhi_j}$	$\boldsymbol{\varPhi_m}$
	Значение показателя x по каждой вариации исследуемого фактора				
1	x_{11}	x_{21}	…	x_{j1}	x_{m1}
2	x_{12}	x_{22}	…	x_{j2}	x_{m2}
3	x_{13}	x_{23}	…	x_{j3}	x_{m3}
…	…	…	…	…	…
i	$x_{1\,i}$	$x_{2\,i}$	…	$x_{j\,i}$	$x_{m\,i}$
n	$x_{1\,n}$	$x_{2\,n}$	…	$x_{j\,n}$	$x_{m\,n}$
Сумма	$\sum_{i=1}^{n} x_{1\,i}$	$\sum_{i=1}^{n} x_{2\,i}$	…	$\sum_{i=1}^{n} x_{j\,i}$	$\sum_{i=1}^{n} x_{m\,i}$
Групповое среднее	$\overline{x_1}$	$\overline{x_2}$	…	$\overline{x_j}$	$\overline{x_m}$

Для проведения дисперсионного анализа подсчитываются суммы показателей x по каждой вариации фактора $\varPhi$: $\sum_{i=1}^{n} x_{j\,i}$. Далее вычисляются групповые средние по формуле: $\overline{x_j} = \frac{1}{n}\sum_{i=1}^{n} x_{j\,i}$.

Они представляют собой средние значения показателя x, соответствующие каждой вариации исследуемого фактора $\varPhi$. Кроме того, необходимо найти общее среднее всех значений показателя: $\overline{x} = \frac{1}{m\,n}\sum_{j=1}^{m}\sum_{i=1}^{n} x_{j\,i}$.

Перед проведением анализа необходимо составить гипотезу (предположение) о том, что фактор $\varPhi$ влияет на изменение результативных количественных показателей x хозяйственной деятельности организации. Эту гипотезу проверяют на уровне значимости α. Уровень значимости — это вероятность ошибки, т. е. вероятность того, что эта гипотеза будет нами отвергнута, в то время как окажется верной. Обычно берут $\alpha \in (0; 0; 1)$.

Если фактор *Ф* не оказывает влияния на измеряемый показатель *x*, то различия между групповыми средними должны носить случайный характер, т. е. на приближенном уровне не должны отличаться друг от друга: $\overline{x_1} \approx \overline{x_2} \approx \ldots \approx \overline{x_m} \approx \overline{x}$. Степень рассеивания (отклонения) отдельных значений показателя x_{ji} относительно общего среднего $\overline{x}$ характеризуется общей суммой квадратов отклонений, вычисляемой по формуле $S_0 = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(x_{ji} - \overline{x}\right)^2$.

Это рассеивание может быть вызвано как действием фактора *Ф*, так и прочими случайными причинами, поэтому представим его в виде суммы: $S_0 = S_1 + S_2$, в которой S_1 — остаточная составляющая, а S_2 — факторная составляющая. Разложим общую сумму квадратов отклонений на факторную и остаточную составляющие, воспользовавшись формулами сокращенного умножения:

$$S_0 = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(x_{ji} - \overline{x}\right)^2 = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(x_{ji} - \overline{x_j} + \overline{x_j} - \overline{x}\right)^2 =$$

$$= \sum_{j=1}^{m}\sum_{i=1}^{n}\left(x_{ji} - \overline{x_j}\right)^2 - 2\sum_{j=1}^{m}\sum_{i=1}^{n}\left(x_{ji} - \overline{x_j}\right)\left(\overline{x_j} - \overline{x}\right) + \sum_{j=1}^{m}\sum_{i=1}^{n}\left(\overline{x_j} - \overline{x}\right)^2.$$

Так как $\sum_{j=1}^{m}\sum_{i=1}^{n}\left(x_{ji} - \overline{x_j}\right) = 0$ (эта сумма представляет собой центральный момент первого порядка, который равен нулю), то второе слагаемое в разложенной сумме квадратов становится равным нулю. Тогда

$$S_0 = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(x_{ji} - \overline{x_j}\right)^2 + \sum_{j=1}^{m}\sum_{i=1}^{n}\left(\overline{x_j} - \overline{x}\right)^2 = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(x_{ji} - \overline{x_j}\right)^2 + n\sum_{j=1}^{m}\left(\overline{x_j} - \overline{x}\right)^2.$$

В полученной сумме первое слагаемое $S_1 = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(x_{ji} - \overline{x_j}\right)^2$ — остаточная сумма квадратов отклонений, характеризующая рассеивание отдельных значений относительно групповых средних, которая не зависит от действия исследуемого фактора. Второе слагаемое $S_2 = n\sum_{j=1}^{m}\left(\overline{x_j} - \overline{x}\right)^2$ — факторная сумма квадратов отклонений, составленная на основе гипотезы о том, что изменение количественного показателя *x* является следствием влияния на него исследуемого фактора *Ф*.

Далее вычисляются общая σ_0^2, остаточная σ_1^2 и факторная σ_2^2 дисперсии по формулам: $\sigma_0^2 = \dfrac{S_0}{nm-1}, \sigma_1^2 = \dfrac{S_1}{m(n-1)}, \sigma_2^2 = \dfrac{S_2}{m-1}$.

Числа $k_0 = nm - 1, k_1 = m(n-1), k_2 = m - 1$ называют степенями свободы.

В качестве наблюдаемого показателя существенности различия дисперсий используется величина $F = \frac{\sigma_2^2}{\sigma_1^2}$, являющаяся отношением факторной дисперсии к остаточной дисперсии.

Для того чтобы сделать вывод о влиянии фактора *Ф* или отсутствии этого влияния, необходимо полученное значение F сравнить с критическим значением F-критерия Фишера-Снедекора на уровне значимости α при степенях свободы k_2 и k_1. Таблицы значений этого критерия приведены практически во всех учебных пособиях по теории вероятности и математической статистике. По таблице F-критерия необходимо найти F_{α,k_2,k_1}. Далее, если $F > F_{\alpha,k_2,k_1}$, то гипотеза принимается, то есть на уровне значимости *α* (с надежностью $(1 - \alpha) \cdot 100$ %) фактор *Ф* влияет на изменение результативных количественных показателей хозяйственной деятельности организации. В противном случае влияние фактора *Ф* на x отвергается.

Для количественной оценки данного влияния используют так называемый коэффициент детерминации: $k_{\text{det},2} = \frac{S_2}{S_0} \cdot 100, k_{\text{det},1} = \frac{S_1}{S_0} \cdot 100$, измеряемый в процентах. Чем ближе значение $k_{\text{det},2}$ к 100 %, тем существеннее влияние фактора *Ф* на измеряемые показатели x.

Пример. Необходимо определить, зависят ли, и если да, то в какой степени, объемы продаж в филиалах торговой компании от их места расположения (торговые центры, специализированные магазины). Для анализа собрана статистика объемов продаж в пяти филиалах компании, расположенных в специализированных магазинах и торговых центрах города за шесть месяцев текущего года. Статистика наблюдений представлена в таблице 2.

Таблица 2

Статистика объемов продаж филиалов торговой компании за шесть месяцев

Периоды работы филиала, взятые для исследования	Объем продаж в филиале, усл. ед.				
	Филиал 1	Филиал 2	Филиал 3	Филиал 4	Филиал 5
1-й месяц	1100	2000	1800	900	1200
2-й месяц	1700	1700	1200	800	1000
3-й месяц	900	1500	700	800	2000
4-й месяц	1200	750	900	1000	800
5-й месяц	1300	800	1400	900	1500
6-й месяц	800	700	1900	900	1600

Решение. Составим гипотезу: фактор *Ф* — месторасположение филиалов влияет на объемы их продаж. Проверим справедливость этой гипотезы на уровне значимости α = 0,05, т. е. с надежностью 95 %.

Имеем: $m = 5$ (пять филиалов компании), $n = 6$ (показатели сняты за шесть месяцев). Проведем подсчет групповых средних:

$$\overline{x_1} = \frac{1}{6}(1100 + 1700 + 900 + 1200 + 1300 + 800) \approx 1167;$$

$$\overline{x_2} = \frac{1}{6}(2000 + 1700 + 1500 + 750 + 800 + 700) \approx 1242;$$

$$\overline{x_3} = \frac{1}{6}(1800 + 1200 + 700 + 900 + 1400 + 1900) \approx 1317;$$

$$\overline{x_4} = \frac{1}{6}(900 + 800 + 800 + 1000 + 900 + 900) \approx 883;$$

$$\overline{x_5} = \frac{1}{6}(1200 + 1000 + 2000 + 800 + 1500 + 1600) \approx 1350.$$

Занесем все полученные значения в матрицу наблюдений (табл. 3).

Таблица 3

Матрица наблюдений однофакторного дисперсионного анализа

Периоды работы филиала, взятые для исследования	**Объем продаж филиала, усл. ед.**				
	Филиал 1	**Филиал 2**	**Филиал 3**	**Филиал 4**	**Филиал 5**
1-й месяц	1100	2000	1800	900	1200
2-й месяц	1700	1700	1200	800	1000
3-й месяц	900	1500	700	800	2000
4-й месяц	1200	750	900	1000	800
5-й месяц	1300	800	1400	900	1500
6-й месяц	800	700	1900	900	1600
Сумма	7000	7450	7900	5300	8100
Групповое среднее для филиала	1167	1242	1317	883	1350

Общую среднюю можно найти как среднюю арифметическую групповых средних:

$$\overline{x} = \frac{1}{5}(1167 + 1242 + 1317 + 883 + 1350) \approx 1192.$$

Вычислим общую, остаточную и факторную суммы квадратов отклонений:

$$S_0 = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(x_{ji} - \overline{x}\right)^2 = (1100-1192)^2 + (1700-1192)^2 + (900-1192)^2 +$$

$$+ (1200-1192)^2 + (1300-1192)^2 + (800-1192)^2 + \ldots + (1200-1192)^2 +$$

$$+ (1000-1192)^2 + (2000-1192)^2 + (800-1192)^2 + (1500-1192)^2 +$$

$$+ (1600-1192)^2 = 5\,060\,420;$$

$$S_1 = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(x_{ji} - \overline{x_j}\right)^2 = (1100-1167)^2 + (1700-1167)^2 + (900-1167)^2 +$$

$$+(1200-1167)^2+(1300-1167)^2+(800-1167)^2+(2000-1242)^2+...+$$
$$+(700-1242)^2+(1800-1317)^2+...+(1900-1317)^2+(900-883)^2+...+$$
$$+(900-883)^2+(1200-1350)^2+...+(1600-1350)^2=4\,227\,086;$$

$$S_2=n\sum_{j=1}^{m}\left(\overline{x_j}-\overline{x}\right)^2=6\cdot\left((1167-1192)^2+(1242-1192)^2+(1317-1192)^2+\right.$$
$$\left.+(883-1192)^2+(1350-1192)^2\right)=835\,170.$$

Далее вычислим общую, остаточную и факторные дисперсии:

$$\sigma^2{}_0=\frac{S_0}{nm-1}=\frac{5\,060\,420}{6\cdot5-1}\approx174\,497;$$

$$\sigma_1^2=\frac{S_1}{m\,(n-1)}=\frac{4\,227\,086}{5\cdot(6-1)}\approx169\,083\,;$$

$$\sigma^2{}_2=\frac{S_2}{m-1}=\frac{835\,170}{5-1}\approx208\,793\,.$$

Найдем степени свободы: $k_0=nm-1=29$, $k_1=m\,(n-1)=25$, $k_2=m-1=4$. Вычислим наблюдаемое значение фактора *Ф*: $F=\dfrac{\sigma_2^2}{\sigma_1^2}=\dfrac{208\,793}{169\,083}\approx1{,}235$. Для отыскания табличного значения *F*-критерия необходимо знать $\alpha=0{,}05$; $k_2=4$; $k_1=25$. По таблице критических значений Фишера-Снедекора находим $F_{0,05;4;25}=2{,}76$. Так как $F<F_{0,05;4;25}$, то с вероятностью 95 % можно считать, что в данном случае место расположения филиалов торговой компании не влияет на объемы их продаж.

Оценим количественно степень влияния исследуемого фактора *Ф*. Для этого вычислим коэффициент детерминации:

$$k_{\det,2}=\frac{S_2}{S_0}\cdot100=\frac{835\,170}{5\,060\,420}\cdot100=16{,}5\,\%.$$

Следовательно, на 16,5 % объемы продаж в филиалах торговой компании зависят от их расположения в торговых центрах или специализированных магазинах города.

Решение этого примера с помощью Excel. Введем данные в таблицу MS Excel (см. рис. 1).

Перейдем на вкладку «**Данные**» [1]. Далее «**Анализ данных».** В появившемся диалоговом окне «**Анализ данных**» в списке «**Инструменты анализа**» выбираем «**Однофакторный дисперсионный анализ**». В открывшемся диалоговом окне «**Однофакторный дисперсионный анализ**» (см. рис. 2) в разделе «**Входной интервал**» зададим диапазон ячеек B3:G7.

В разделе «**Группирование**» выберем положение «**по строкам**». В разделе «**Выходной интервал**» укажем ячейку для результатов анализа A13. После щелчка по кнопке «**ОК**» получим итоговую таблицу вычислений (см. рис. 3).

	A	B	C	D	E	F	G
1	Исследование влияния месторасположения филиалов торговой компании на их объем продаж						
2		Объем продаж филиала с 1-го по 6-й месяцы текущего года,усл. ед					
3	Филиал №1	1100	1700	900	1200	1300	800
4	Филиал №2	2000	1700	1500	750	800	700
5	Филиал №3	1800	1200	700	900	1400	1900
6	Филиал №4	900	800	800	1000	900	900
7	Филиал №5	1200	1000	2000	800	1500	1600

Рис. 1. Вид рабочего листа с введенными данными

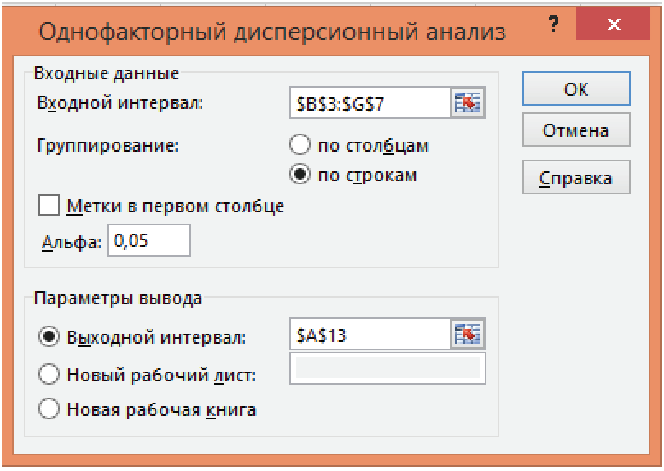

Рис. 2. Вид диалогового окна «Однофакторный дисперсионный анализ»

12							
13	Однофакторный дисперсионный анализ						
14	ИТОГИ						
15	*Группы*	*Счет*	*Сумма*	*Среднее*	*Дисперсия*		
16	Строка 1	6	7000	1166,66667	102666,667		
17	Строка 2	6	7450	1241,66667	316416,667		
18	Строка 3	6	7900	1316,66667	229666,667		
19	Строка 4	6	5300	883,333333	5666,66667		
20	Строка 5	6	8100	1350	191000		
21							
22	Дисперсионный анализ						
23	*Источник вариации*	*SS*	*df*	*MS*	*F*	*P-Значение*	*F критическое*
24	Между группами	833333,3	4	208333,333	1,23213406	0,32261496	2,75871047
25	Внутри групп	4227083	25	169083,333			
26	Итого	5060417	29				

Рис. 3. Вид рабочего листа с результатами дисперсионного анализа

В итоговой таблице дисперсионного анализа $F < F_{0,05;4;25}$. Следовательно, с надежностью 95 % можно утверждать, что место расположения филиалов торговой компании в торговых центрах или специализированных магазинах города не влияет на их объем продаж.

Коэффициенты детерминации при этом можно вычислить, взяв данные из столбца SS:

$$k_{\det,2} = \frac{S_2}{S_0} \cdot 100 = \frac{833\,333}{5\,060\,417} \cdot 100 = 16{,}5\%;$$

$$k_{\det,1} = \frac{S_1}{S_0} \cdot 100 = \frac{4\,227\,083}{5\,060\,417} \cdot 100 = 83{,}5\%.$$

Из чего следует вывод, что на объем продаж филиалов торговой компании на 83,5 % влияют прочие условия, влияние же исследуемого фактора (место расположения филиала) осуществляется на 16,5 %.

Проиллюстрированный пример обосновывает необходимость использования интегрированного подхода в подготовке бакалавров экономического направления. Для того чтобы грамотно проводить маркетинговые исследования, будущему специалисту помимо специальных знаний в области маркетинга, необходимо владеть, с одной стороны, информационно-коммуникационными технологиями для быстрого и точного проведения исследований на базе имеющихся программных продуктов и, с другой стороны, методами статистического анализа для правильной интерпретации полученных результатов. Такая интеграция при изучении математики, информатики и маркетинга для бакалавров бизнес-информатики дает безусловно положительный результат.

Литература

1. *Дегтярева Л.В.* Анализ в системе маркетинга: учебно-метод. пособие. М.: МГПУ, 2013. 52 с.

2. *Дегтярева Л.В., Семеняченко Ю.А.* Взаимосвязь маркетинга, математических методов и информационных технологий при подготовке бакалавров экономического направления // Наука и образование в XXI веке: проблемы и перспективы. Пенза: Приволжский дом знаний, 2014. С. 20–25.

3. *Лялин В.С., Зверева И.Г., Никифорова Н.Г.* Статистика: теория и практика в Excel: учебное пособие. М.: Финансы и статистика, 2010. 448 с.

Literatura

1. *Degtyareva L.V.* Analiz v sisteme marketinga: uchebno-metod. posobie. M.: MGPU, 2013. 52 s.

2. *Degtyareva L.V., Semenyachenko Yu.A.* Vzaimosvyaz’ marketinga, matematicheskix metodov i informacionny’x texnologij pri podgotovke bakalavrov e’konomicheskogo napravleniya // Nauka i obrazovanie v XXI veke: problemy’ i perspektivy’. Penza: Privolzhskij dom znanij, 2014. S. 20–25.

3. *Lyalin V.S., Zvereva I.G., Nikiforova N.G.* Statistika: teoriya i praktika v Excel: uchebnoe posobie. M.: Finansy’ i statistika, 2010. 448 s.

L.V. Degtyareva, Y.A. Semenyachenko

Integration of Mathematics, Informatics and Marketing in the Process of Training Bachelors in Business Informatics

The article substantiates the need for an integrated approach at teaching mathematics, computer science and marketing in the preparation of bachelors of Business Computer science and the authors give an example of implementation of such an approach.

Keywords: marketing; information and communication technologies; bachelor of Business informatics; statistical methods; variance analysis.