

Л. Л. Соловьева, Н. В. Снопок
г. Гомель, ГГТУ им. П. О. Сухого

ОБЕСПЕЧЕНИЕ РЕПРЕЗЕНТАТИВНОСТИ ВЫБОРКИ ПРИ ПРОВЕДЕНИИ МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЙ

Маркетинговые исследования – это важная составляющая общей деятельности любой фирмы, которая возникла с переходом от экономики, ориентированной на производство к экономике, ориентированной на потребление. Маркетинговые исследования обеспечивают связь потребителя, покупателя и обществу с продавцом с помощью информации, которая используется для определения благоприятных и неблагоприятных возможностей использования маркетинга. Кроме того маркетинговые исследования ассоциируются со всеми фазами информационно-управленческого процесса: уточнения необходимой информации; сбором и анализом информации; интерпретации информации в соответствии с целями при проведении маркетинговых исследований.

При проведении маркетингового исследования решаются следующие вопросы: выбор методов и инструментов проведения исследования; определение объема необходимой информации; способов ее сбора; проектирование выборки, а также выявление особенностей проведения исследования и принципов анализа данных. Исследование потребителей является основным направлением при исследовании рынка. При его проведении следует принимать во внимание, что получить точные данные, поддающиеся количественной оценке затруднительно.

Маркетинговые исследования часто требуют оценки некоторых характеристик интересующей совокупности. Например, это может быть уровень известности торговой марки среди жителей города. Для решения такой задачи при соблюдении требований эффективности затрат и точности данных, отбирается часть элементов, представляющих всю исследуемую совокупность, то есть формируют выборку. Выборка применяется, когда размеры совокупности велики или когда для получения информации от всей совокупности необходимо затратить слишком много времени и средств. Понятие выборки в маркетинге рассматривается в двух значениях. Во-первых, это элементы генеральной совокупности, подлежащие изучению, то есть выборочная совокупность. Во-вторых, выборка – это процесс (соответственно: методы, приемы, процедуры) формирования выборочной совокупности при необходимом условии обеспечения репрезентативности.

Репрезентативность (представительность) выборки – это свойство выборки, предоставляющей исследователю возможность сделать на ее основании правильные выводы о свойствах всей проверяемой совокупности. То есть, выборка, исследование которой позволяет проецировать выводы на всю совокупность, из которой эта выборка сделана, является репрезентативной. Чтобы обеспечить репрезентативность выборки, необходимо выполнить следующие условия: 1) выбор адекватного метода отбора; 2) правильное определение объема выборочной совокупности; 3) добросовестная работа интервьюеров; 4) проверка репрезентативности выборки.

Первое условие касается правильности выбора вида выборки. Методы построения выборки должны быть адекватны объекту, целям и задачам исследования. Выборки делятся на два типа: вероятностные и детерминированные (невероятностные). К вероятностным относят следующие виды: простой случайный отбор; систематический отбор; серийный отбор; стратифицированный отбор и пр. К детерминированным относят: стихийный отбор; отбор по методу «снежного кома»; отбор на основе квот и пр. При маркетинговых исследованиях чаще используются следующие виды отбора.

Простой случайный отбор – один из основных, используемых в маркетинговых исследованиях. Главный принцип подобного отбора – обеспечение возможности каждой единице генеральной совокупности попасть в выборочную. С этой целью используются таблицы случайных чисел, лотерейный подбор, механический отбор.

Стихийный отбор, т. е. отбор по принципу «добровольности» вхождения единиц генеральной совокупности в выборочную. Он используется довольно часто, в частности в почтовых и прессовых опросах. Основной недостаток подобного отбора – невозможность качественной репрезентации генеральной совокупности. Тем не менее, стихийная выборка используется и с учетом ее экономичности, и с учетом возможности корректировки выборки (добора или исключения части полученной совокупности с целью достижения требуемых в соответствии с определенными объемами и элементами выборочной совокупности), и, наконец, в некоторых исследованиях, когда формирование выборочной совокупности просто невозможно. Размер

и состав стихийных выборок заранее не известен, и определяется только одним параметром – активностью респондентов.

Квотный отбор, в основе которого лежит построение качественной модели генеральной совокупности, затем – отбор единиц наблюдения в выборочную совокупность, исходя из имеющейся модели. Главное для обеспечения репрезентативности квотной выборки – обеспечить элементам генеральной совокупности равные шансы попасть в выборку. Квотная выборка репрезентативна только по квотирующим признакам (а их, как правило, не может быть больше 2–3, иначе квотные задания становятся слишком сложными). Тем не менее, практика показывает репрезентативность квотной выборки и по тем признакам, которые обычно не являются квотирующими (семейное положение, вероисповедание, размер семьи и др.).

Вторая проблема касается выбора способа определения объема выборки. В социологических исследованиях объем выборки определяют двумя способами: по таблицам зависимости от объема генеральной совокупности [2] или по статистическим формулам, обеспечивающим получение требуемой точности результатов с заданной вероятностью.

Объем выборки определяется по формуле (1) [1, с. 241].

$$n = \frac{t^2 \sigma^2}{\Delta^2}, \quad (1)$$

где n – объем выборки; t – коэффициент доверия (t -критерий Стьюдента для определенного уровня вероятности); σ – среднее квадратичное отклонение изучаемого признака; Δ – допустимая ошибка. Значение показателей t и Δ определяются исследователем в зависимости от желаемой достоверности полученных результатов. Высокая степень вероятности и маленькая ошибка приводят к увеличению объема выборки. В маркетинговых исследованиях предельная ошибка обычно берется равной 3–5 %, а вероятность 0,954 ($t = 2$).

Проблема использования данной формулы состоит в определении среднее квадратичного отклонения изучаемого признака. Для начала необходимо выбрать показатели, характеризующие изучаемый признак. Изучаемый признак может иметь физическую величину (например, средний размер покупки, интенсивность потребления, размер заказа, количество времени, затраченное на посещение данного магазина) или выражаться в долях единицы (например, доля постоянных приверженцев марке, доля «странников», доля предпочитающих импортных или отечественных производителей). Если значение дисперсии неизвестно, то при расчете объема выборки используются приближенная оценка, полученная, либо при пробных исследованиях, а при ограниченных ресурсах с помощью экспертной оценки среднего значения признака.

σ для признака, имеющего физическую величину, рассчитывается по правилу «трех сигм» [1, с. 241].

$$\sigma = \frac{1}{3} x_{cp}, \quad (2)$$

где x_{cp} – среднее значение показателя (например, размера покупки).

Для признака, выраженного в долях, σ определяется по формуле (3).

$$\sigma^2 = p(1-p), \quad (3)$$

где p – доля потребителей, имеющих изучаемый признак (например, доля приверженцев марке, любителей отечественной продукции). Значение p берется равным 0,5, и в этом случае σ^2 будет максимальным.

Если подставить значения показателей, имеющих физическую величину, то формулу (1) можно упростить следующим образом:

$$n = \frac{t^2 \sigma^2}{\Delta^2} = \frac{t^2 \left(\frac{1}{3} x_{cp}\right)^2}{(x_{cp} \times 0,05)^2} = \frac{t^2}{9 \times 0,05^2} = 177 \text{ чел.} \quad (4)$$

Таким образом, если изучаемый признак имеет физическую величину, и у нас нет предварительно рассчитанной дисперсии, то объем выборки будет равен 177 человек при вероятности 0,954 и 5 % ошибке независимо от значения самого признака.

Величина объема выборки показывает, сколько человек необходимо опросить, чтобы получить достоверные данные. Если использовать разные подходы при определении объема

выборки, то результаты могут отличаться в десятки раз. Например, изучаемая генеральная совокупность 250 тыс. чел. По табличным значениям [2, с. 234] объем выборки составит 398 человек. Если изучаемый признак – количество приобретенных единиц товара, то по формуле (5) получим 177 человек, а если использовать признак доля какой-либо марки и рассчитывать по формулам (1) и (3), то получим 1600 человек. Как правило, при использовании нескольких методов расчета объема выборки выбирается наибольшее число, или выбирают метод исходя из средств и целей проведения исследования.

Третье условие касается кадровых вопросов выбора интервьюеров и контроля их работы.

Однако кроме выполнения вышеперечисленных условий необходимо оценить степень репрезентативности выборки. Задача исследователя – обеспечить соответствие структуры выборки структуре генеральной совокупности по характеристикам потребителей. Как правило, для этого используются такие показатели как возраст, пол, уровень дохода, т. е. основные социально-демографические характеристики генеральной совокупности. Обоснование репрезентативности выборки используют статистические методы и внестатистические.

Статистические методы можно использовать при использовании вероятностных методов отбора. Приблизительно эти методы можно использовать и при стихийном отборе. Для обобщения результатов исследования на генеральную совокупность применяются строгие индуктивные процедуры статистического вывода, оценивается ошибка выборки с заданной вероятностью. *Внестатистическое обоснование* репрезентативности предполагает теоретическое доказательство того, что выборка достаточно хорошо представляет генеральную совокупность. При использовании этого подхода статистическое оценивание ошибок выборки не производится. Квотные выборки уже предполагают, что они репрезентативны по формирующим признакам и использование здесь статистических расчетов нецелесообразно. Для статистической оценки репрезентативности определяются доверительные интервалы и ошибка выборки. Для определения ошибки выборки по средней и доле используются следующие формулы.

Пределы среднего значения признака – формула (5).

$$x_{\text{выб}} - \Delta x \leq \tilde{x} \leq x_{\text{выб}} + \Delta x \quad (5)$$

где \tilde{x} – среднее значение признака по генеральной совокупности; Δx – предельная ошибка средней; $x_{\text{выб}}$ – среднее значение признака по выборочной совокупности.

Предельная ошибка средней рассчитывается по формуле (6).

$$\Delta x = \pm t \sqrt{\frac{\sigma^2}{n}} \quad (6)$$

Например, предположим, что руководитель туристического агентства задался целью узнать среднегодовые расходы рыбаков на питание и проживание во время их походов. Некоторое число рыбаков не расходует никаких средств, т. к. выезжают только на день. Другие совершают несколько недельных поездок в год. Предположим, что 30 дней в году – это предел пребывания на рыбалке. При этом ежедневные расходы составляют около \$10, т. е. верхний денежный предел равен \$300. Область изменения этой величины также равна \$300 (т. к. потратить меньше \$0 невозможно), а среднее значение – \$150 и расчетное среднееквадратическое отклонение составляет $150/3=50$. Расчетный объем выборки в нашем случае составляет 177 чел.

Оценку репрезентативности проведем по возрасту исследуемой совокупности. Предположим, что произведенные наблюдения дадут выборочное среднее $\bar{X} = 45$ лет. Тогда доверительный интервал определяется:

$$45 \pm 2 * 50 / \sqrt{177} = 45 \pm 7,5 \text{ или } 37,5 \leq \tilde{x} \leq 52,5 \text{ лет.}$$

Исходя из расчетов, если средний возраст генеральной совокупности попадет в заданный интервал, то выборка будет репрезентативной и результаты исследования можно переносить на генеральную совокупность. Выполнение перечисленных условий репрезентативности выборки позволят обеспечить достоверность результатов маркетинговых исследований.

Список использованных источников

- 1 Елисеева, И. И. Общая теория статистики : учебник / И. И. Елисеева, М. М. Юзбашев ; под общ. ред. И. И. Елисеевой. – 5-е изд., перераб. и доп. – М. : Финансы и статистика, 2004. – 656 с.
- 2 Ядов, В. А. Стратегия социологического исследования: описание, объяснение, понимание социальной реальности : учебник для вузов / В. А. Ядов. – М. : Добросвет, 2000. – 596 с.