

Интерпретация данных в прикладной социологии

Программа курса 12 часов

к.филос.н. М.Ф.Черныш

Тема 1

Теоретическая база интерпретации данных

1. В прикладной социологии существуют различные уровни анализа эмпирических данных. Самым высоким из них является уровень теоретической парадигмы. В арсенале социологии присутствуют несколько общих теоретических парадигм – культурологическая (веберовская, постмодернистская), конфликтологическая (марксистская), функциональная (Парсонс, Мертон), а также ряд смешанных, вариантов. Одним из примеров использования общей теоретической парадигмы при построении возможных вариантов интерпретации данных является гендерная социология, где одинаково успешно используется как функциональный, так и конфликтологический подходы.

Литература

1. **Бергер П.** Приглашение в социологию. М.: Аспект Пресс, 1996.
2. **Смелсер Н.** Социология. Пер. с англ. М.: Феникс, 1994.
3. **Lazarsfeld P. and Rosenberg M.** The language of Social Research, Glencoe: Free Press, 1955.
4. **Smelser N.** Comparative methods in the social sciences. Prentice: Hall, 1976.

2. Теория среднего уровня является основным инструментом интерпретации данных. К.Поппер рассматривал теорию среднего уровня как оптимальный уровень анализа прикладной информации. Р.Мертон определял теорию среднего уровня как основной рабочий инструмент социолога. В отличие от общей теоретической парадигмы теория среднего уровня оперирует ограниченным числом признаков. Теория среднего уровня может быть определена как представление исследователя о характере взаимосвязи между тремя или несколькими переменными. Прежде чем приступить к построению теории среднего уровня исследователь должен обозначить границы "мира переменных". В этом мире отбираются переменные, дающие возможность построения объяснительной стратегии. Пример: теории девиантного поведения – теория "блокирования ценностей", теория "уклонения от боли", теория "наследования порочных качеств". На основании переменных, включаемых в рамки анализа, разрабатывается ряд гипотез о характере возможной взаимосвязи между ними. Возможно выделить следующие типы гипотез: гипотезы о направленности влияния переменных друг на друга и гипотезы об интенсивности взаимовлияния переменных.

Литература

1. **Chavetz J.** A primer of the Construction and Testing of Theories in Sociology. Itasca, Illinois: Peacock, 1978.
2. **Freidheim E.** Sociological Theory in Research Practice. Cambridge: Schenkman, 1976.
3. **Merton. R.** Social Theory and Social Structure. New York: Free Press, 1968.
4. **Rose. G.** Deciphering Sociological Research. London: Macmillan, 1978.
5. **Seiber J.** Planning Ethically Responsible Research: A Guide for Students and Internal Review Boards. Beverly Hills: Sage, 1992.
6. **Stinchcombe A.** Constructing Social Theories. N.Y.: Harcourt Brace Jovanovich, 1968.

Тема 2

Интерпретация элементарных статистических распределений

1. На первом этапе данные прикладного исследования представлены в форме элементарных частотных распределений. Уже тогда возможно уловить некоторые закономерности, характерные для наблюдаемого объекта. Отдельные его свойства выступают в форме эмерджентных свойств или, иными словами, в форме упорядоченных аналитических характеристик. Первый тип эмерджентного свойства – это численность группы. Пример: численность "средних слоев населения". В разных теоретических парадигмах – численность этой группы будет интерпретироваться по-разному. Для социологов марксистской ориентации наличие в обществе широкой прослойки говорит об "обуржуазивании пролетариата", разложении революционного движения. Для социологов структурного направления мощный средний класс являет собой один из признаков современности, перехода общества в стадию стабильной эволюции. Второй эмерджентное свойство, проявляющееся на этапе анализа элементарных частотных распределений, – это состав социальной совокупности, выражающейся в различиях между ее элементами. Пример: анализ генезиса капитализма в работах М. Вебера. Сравнительный анализ группы предпринимателей с населением в целом позволил Веберу определить группу протестантов как носителей особой системы ценностей, свойственной капиталистической форме хозяйствования.

Уже на этапе анализа элементарных статистических распределений становится возможным говорить об ограничениях интерпретации данных. Первый тип ограничений связан с характером интерпретируемой переменной. Так возраст или доход, воплощенные в числовых переменных, скрывают закономерности за множеством значений. Прежде, чем приступить к их анализу, необходимо выполнить процедуру преобразования переменных. Для упрощения интерпретации возраст или доход разбиваются на категории, наполненность которых определяется целями исследования. Второй тип ограничений связан с

необходимостью учитывать основные параметры выборочной совокупности. Даже самая лучшая выборка – это только модель населения, страдающая многими неточностями. При интерпретации данных следует с особой тщательностью относиться к такой характеристике выборки как стандартная статистическая ошибка. Предположим, что анализируются данные об электоральных предпочтениях населения. Как правило, случайная вероятностная выборка российского населения объемом в 3000 человек дает стандартную статистическую ошибку не меньше 3,5%. Любые различия между группами, определяемые на уровне стандартной статистической ошибки не могут считаться значимыми. Третий тип ограничений при интерпретации линейных статистических распределений связан с неточностью операционального определения показателей. Очень часто в многонациональной среде группы, принадлежащие к разным культурам, по-разному трактуют понятия, содержащиеся в вопросах анкеты. С особой остротой эта проблема встает в кросс-культурных сравнительных исследованиях. Пример: международное исследование системы ценностей. Вопрос: насколько вы счастливы? В США быть счастливым – это нормативная ценность. В России, наоборот, не принято признавать за собой счастье или удачу.

Литература

1. **Freeman I.** Elementary Applied Statistics. N.Y.: Wiley.
2. **Fowler F.** Survey Research Methods. Beverly Hills: Sage, 1988.
3. **Mueller J., Schuessler K. and Cosner H.** Statistical Reasoning in Sociology. Boston: Houghton Mifflin, 1977.
4. **Tabular Presentation.** In: International Encyclopaedia of the Social Sciences. N.Y. MacMillan and Free Press, 1968.

2. Типичные ошибки при интерпретации линейных статистических распределений. Наибольшее распространение имеет так называемое "экологическое" заблуждение. Пример: "экологическое" заблуждение в работе Дюргейма "Самоубийство". Вторая типичная ошибка – интерпретации данных, полученных на статистически незначимых совокупностях. Если в исследовании выделена группа, объем которой менее тридцати респондентов, то интерпретация распределений в этой группе имеет целый ряд статистических ограничений и, в частности, находится в зависимости от метода, при помощи которого была сформирована выборка исследования. Третье типичная ошибка связана с интерпретацией временных рядов, когда изменения рейтинговых значений рассматривается без учета ошибки выборки, различий в формулировках вопросов и т.д.

Литература

1. **Ноэль Е.** Массовые опросы. Введение в методику демоскопии. М.:Авва-Эстра, 1993.
2. **Johnson A.** Social Statistics Without Tears. N.Y.: McGraw Hill, 1977.

3. **Kalton G.** Introduction to survey sampling. Beverly Hills: Sage, 1983.
4. **Loether H. and McTavish D.** Descriptive Statistics for Sociologists. Boston: Allyn and Bacon, 1974.

Тема 3

Анализ и интерпретация взаимосвязей

1. Как правило, построение сложных статистических моделей начинается с анализа и интерпретации взаимосвязи переменных. При этом тестирование уровня взаимосвязи между переменными должно следовать в русле общей логики исследования. Первоначально замеряется взаимосвязь первого порядка, т.е. взаимосвязь между двумя переменными, включенными в набор гипотез исследования. Пример: взаимосвязь между возрастом и доходом. Как правило, даже наличие тесной взаимосвязи между переменными не дает возможности адекватно интерпретировать полученную информацию. Основная проблема интерпретации состоит в том, чтобы адекватно отразить в выводах не две или три переменные, а все возможное признаковое пространство, которым оперирует исследователь. Пример: по данным многочисленных исследований возраст и доход – это две тесно коррелирующие между собой переменные. Однако, прежде чем делать какие-либо выводы, следует включить в цепь рассуждений целый ряд других переменных, таких как пол, образование, сфера занятости. Подобное включение осуществляется посредством введения в контекст анализа контрольных переменных. Под контрольной понимается переменная, способная повлиять на характер взаимосвязи между двумя другими переменными. Пример: взаимосвязь таких переменных как доход и возраст существенно слабеет тогда, когда в качестве контролирующей переменной используется такая переменная как образование. Вместе с тем, взаимосвязь таких переменных как доход и образование обнаруживает высокий уровень устойчивости и лишь незначительно слабеет под влиянием другим переменных, описывающих статус респондента. На этом этапе полученные данные позволяют отбросить ряд гипотез, не получивших подтверждения, но ни в коем случае не дают возможности делать выводы позитивного порядка. В прикладном исследовании необходимо избегать беспредметных спекуляций, строя выводы лишь на взаимосвязи двух переменных. Действительно корректные выводы могут быть сделаны только по результатам многомерного анализа.

Литература

1. **Ehrenberg. A.** Data reduction. Analysing and Interpreting Statistical Data. London: Wiley, 1975.
2. **Fowler F.** Survey Research Methods. Beverly Hills: Sage, 1988.

3. **Freeman L.** Elementary Applied Statistics. N.Y.: Wiley.

2. Анализ взаимосвязей первого порядка позволяет внести коррективы в уже имеющиеся гипотезы исследования и выдвинуть ряд новых взаимосвязанных гипотез. Коррекция старых гипотез и надстраивание новых гипотез над уже имеющимися это не что иное, как сужение поля интерпретации. На этом этапе возникает острая потребность в процедурах многомерного анализа. Имеющиеся статистические методы многомерного анализа открывают возможности для а) выявления новых латентных переменных, б) определение системы взаимосвязи между переменными. Наиболее часто используемый способ выявления латентных переменных – факторный анализ. Из всех методов факторного анализа чаще всего используется метод главных компонент. Его главное преимущество в том, что он позволяет представить систему взаимосвязей внутри целой группы переменных в компактной форме. Основным продуктом факторного анализа является типология данных. Особенности статистического метода налагают ограничения на возможности интерпретации результатов факторного анализа. Так, веса переменных в факторе оказываются зависимыми от характера использовавшихся в анкете шкал. Особенно сложно интерпретировать результаты, полученные после ротации факторной матрицы. Каждый метод ротации гипертрофирует те или иные характеристики матрицы, принуждая исследователя к осторожной интерпретации.

В практике анализа данных прикладного исследования часто используется метод регрессионного анализа. На этом этапе анализа следует сделать выбор между зависимыми и независимыми переменными. Следует помнить, что адекватная интерпретация данных линейной регрессии возможна лишь в том случае, если переменные носят числовой характер и не связаны между собой. Данные линейной регрессии указывают лишь на линейную взаимосвязь между переменными и эффективно работают лишь в том случае, если используемые переменные характеризуются нормальным распределением. Данные линейной регрессии открывают возможности для создания каузальных моделей, т.е. моделей, воспроизводящих причинно-следственные отношения между переменными.

Литература

1. **Blalock H.** Social Statistics. New York: McGraw Hill, 1972.
2. **Norouisis M.** SPSSX: Introductory Statistics Guide. N.Y.: McGraw-Hill, 1983.
3. **Hellevik O.** Introduction to Caysak Analysis. London: George Allen and Unwin, 1984.
4. **Huck S. and Sandier H.** Rival Hypotheses: Alternative Interpretation of Data Based Conclusions, N.Y.: Harper and Row.

3. Наиболее сложным и всеобъемлющим способом тестирования гипотез является построение структурных уравнений. Структурные уравнения включают в себя совокупность переменных и указания на характер взаимосвязи между ними. Структурные уравнения создаются тогда, когда уже есть данные о парных взаимосвязях переменных. Для вычисления структурных уравнений с использованием базы данных прикладного исследования широко используется метод логлинейного анализа. Получаемые в ходе выполнения данной процедуры данные позволяют максимально приблизиться к пониманию системы каузальных взаимосвязей между переменными и, следовательно, к адекватной интерпретации данных.

Литература

1. **Joreskog and Sorbom.** Lisrel 7. A Guide to the Program and Applications, 1988.

Тема 4

Аналитический отчет и презентация данных

Аналитический отчет по результатам исследования – это документ, включающий в себя характеристику самого исследования и основных его результатов. В аналитический отчет должны с необходимостью входить следующие основные разделы:

Краткое изложение результатов исследования. Подобный раздел часто бывает необходим не только в аналитических отчетах, но и научных публикациях, особенно публикациях в западных научных журналах. Краткое изложение результатов не должно превышать одной–двух страниц.

Описание условий проведения исследования. В отчете по данным количественного исследования следует охарактеризовать его выборку, расчет стандартной статистической ошибки выборки, метод интервьюирования, количество уклонений от ответа, других нарушений принципов отбора респондентов.

Изложение основных результатов исследования. В этом разделе воспроизводится не только результаты, но и логика анализа данных. Как правило, этот раздел состоит из подразделов, отражающих логику анализа данных.

Заключение. В заключении по результатам исследования содержатся указания на проблемы, не получившие своего разрешения в ходе исследования, а также возможные пути достижения решения данных проблем.

В последнее время для адекватной презентации данных исследования все чаще используются методы построения графиков и диаграмм. Успешное применение этого метода опирается на закономерности восприятия форм и цветов человеческим сознанием. В основе лежит знание законов перспективы.

Литература

1. **Ноэль Е.** Массовые опросы. Введение в методику демоскопии. М.:Авва–Эстра, 1993.
2. **Рабочая книга социолога.** 2-е издание дополненное. М.: Наука, 1983.