

План лекции 7

«Типы статистических исследований»

1. Исследования на основе наблюдения
2. Выборочные обследования
3. Эксперимент
4. Полевые испытания
5. Гарвардский проект по физике
6. Эксперимент по улучшению жилищных условий
7. Преимущества рандомизированного эксперимента как вида исследования
8. Трудности в проведении эксперимента
9. Основные этапы планирования и анализа эксперимента
10. Критерии оптимальности планов

Виды статистических исследований

Классификация статистических исследований производится на основании возможности исследователя вмешиваться в изучаемый процесс и управлять факторами. Спектр участия исследователя очень широк - от невозможности вмешательства до воспроизводимости исследования с точностью до ошибки.

Роль исследователя

1. Управлять факторами (назначать уровни факторов экспериментальным единицам в случайном порядке);
2. Структурировать источники неоднородностей в блоки;
3. Рандомизировать случайные ошибки;
4. Воспроизвести эксперимент с точностью до ошибки.

Основные типы статистических исследований

1. Исследования на основе выборки
2. Выборочные обследования
3. Эксперимент
4. Полевые испытания

Исследования на основе выборки

- Исследователь является наблюдателем;
- «пассивный эксперимент»;
- Наблюдения – это сбор и регистрация фактов, событий, процессов;
- Если из наблюдений можно сделать заключение, то такие наблюдения называются данными;
- Данные классифицируются, представляется в виде таблиц, графиков, моделей;
- Изучение свойств массивов данных (меры центральной тенденции, разброс данных);
- Изучение как обстоят дела, а не оценивание эффектов;
- Анализ данных архивов (ретроспективная рандомизация);
- Рандомизация возможна только на стадии измерения откликов.

Выборочные обследования

- Предназначены для получения информации об изучаемой совокупности: изучение потребительского спроса, изучение общественного мнения, оценка организаторских способностей;
- Экономят время и деньги по сравнению со сплошными обследованиями;
- Однако в важных задачах (перепись населения, избирательные компании) используются сплошные обследования.

Совокупность

- Совокупность – это множество единиц, обладающих некоторыми общими свойствами, существенными для их характеристики. Например, совокупность учащихся, совокупность классов.
- Совокупность (конечная или бесконечная), относительно которой нужно сделать вывод, называется генеральной совокупностью.
- Выборка – это часть генеральной совокупности.
- Элемент совокупности в процедуре отбора называется единицей выборки.
- Если случайно выбираются школы из совокупности школ региона, то единицей выборки является школа.
- Единицей выборки может быть семья, учащийся, город и т.д.

Выборка

- Простой случайной выборкой называется такая вероятностная выборка, в которой любая единица имеет одинаковую и независимую возможность быть отобранной.
- В математической статистике понятия выборки и совокупности тесно связаны. Понятие выборки приобретает содержание только тогда, когда указывается, из какой совокупности она отобрана. В свою очередь, совокупность представляется только с помощью выборки. Выборка и совокупность связаны процедурой отбора.
- Для обоснованности вывода важно, чтобы выборка была случайной, т.е. Полученной с помощью случайного отбора.
- Элемент совокупности в процедуре отбора называется единицей выборки.
- Случайная выборка – это выборка, не взятая наобум, а полученная в соответствии с определенной вероятностной процедурой.

Совокупность и выборка

- Сравнение двух процедур – случайного отбора и рандомизации.
- В математической статистике понятия выборки и совокупности тесно связаны. Понятие выборки приобретает содержание только тогда, когда указывается, из какой совокупности она отобрана. В свою очередь, совокупность представляется только с помощью выборки. Выборка и совокупность связаны процедурой отбора.
- Для обоснованности вывода важно, чтобы выборка была случайной, т.е. Полученной с помощью случайного отбора.
- Элемент совокупности в процедуре отбора называется единицей выборки.
- Случайная выборка – это выборка, не взятая наобум, а полученная в соответствии с определенной вероятностной процедурой.

Процедуры случайного отбора и рандомизации

- Общим является искусственное внесение случайности.
- Отличие: в выборочных обследованиях речь идет о получении случайной выборки и случайность предназначена для уменьшения смещения. Здесь э.е. уже обработаны помимо воли исследователя. Поэтому невозможно установить причины изменения отклика.
- В процедуре рандомизации при планировании эксперимента исследователь распоряжается обработками, искусственное внесение случайности происходит при назначении обработок. Поскольку процесс назначения обработок полностью управляем, то это позволяет устанавливать причинные связи.

Эксперимент

- Исследование является экспериментом, если изучаемые факторы изменяются по воле исследователя в точно учитываемых условиях, позволяющих управлять ходом явления и воссоздавать его каждый раз при повторении с точностью до случайных ошибок.
- Именно эксперимент, как тип исследования, позволяет изучать причинные связи.
- Здесь рандомизация является принципиально необходимой.
- Социальный эксперимент связан с улучшением условий труда и быта, снижением заболеваемости и смертности населения, развитием науки и культуры.

Эксперимент в педагогике

В самом простом варианте эксперимент сводится к сравнению некоторого новшества (в педагогике – инновации) с прежним положением дел. Формируются две группы э. е., одна из которых «обрабатывается» предлагаемым нововведением, другая - служит контролем. Необходимо узнать, насколько отклики обрабатываемых единиц отличаются от откликов контрольных, т.е. провести сравнительный эксперимент с контролем для того, чтобы оценить эффективность обработки. Случайное назначение э. е. в контрольную и обрабатываемую группы является существенной чертой правильно организованного эксперимента. Это служит гарантией, что э. е. двух групп в среднем будут одинаковы по всем признакам.

Квазиэксперимент

- Часто встречаются такие ситуации, когда невозможно сделать рандомизированное назначение экспериментальных единиц для контрольной и экспериментальной групп.
- Пример – оценить вред от курения.
- Такие эксперименты называются «квазиэкспериментами». Они отличаются от «истинных экспериментов» отсутствием рандомизации при назначении единиц.
- Интерпретация эффектов обработок в нерандомизированных экспериментах столь же затруднительна, как и в наблюдениях вследствие смешивания этих эффектов с неучтенными эффектами, в частности с эффектами самих единиц.

Актуальность полевых испытаний

Перед тем, как широко внедрять новые программы, необходимо их тщательная апробация. Часто считают, что для оценки новшества в общественных исследованиях необязательно проводить широкую экспериментальную проверку, достаточно теоретического анализа или моделирования. К сожалению, для большинства экономических и социальных процессов отсутствуют хорошие теоретические модели и невозможно предвидеть, что случится, если в системе произойдут изменения. Отсутствие четкого понимания взаимосвязей между социальными и экономическими факторами не позволяет точно предсказать прямые и побочные эффекты социального и экономического вмешательства. Поэтому очень важно увеличить экспериментальный опыт, так как рискованно давать рекомендации о распространении новшества, не изучив все варианты на практике. С этой целью проводятся полевые испытания.

Особенности полевых испытаний

Термин «испытание» говорит об активном действии исследователя, направленном на прямое сравнение эффектов обработок. От понятия «эксперимент», который, обычно ассоциируется с лабораторным исследованием, «испытание» отличается прежде всего крупномасштабностью. Здесь участвует большое число э. е. (индивиды, семьи, районы, города и т.д.), которое может достигать многих тысяч и миллионов. Термин «полевой» говорит о том, что исследование проводится в реальных условиях («в поле»), а не с помощью моделирования или теоретически. Если результаты эксперимента - важное средство для ученого, то результаты полевых испытаний являются основой для принятия решений управленческим аппаратом и правительственными органами.

Особенности полевых испытаний

Проведение и оценка полевых испытаний новых социальных и экономических программ требуют громадных государственных расходов. Поэтому правильная организация полевых испытаний имеет особое значение. Они должны проводиться таким образом, чтобы можно было извлечь всю необходимую информацию и получить неискаженные оценки эффектов предлагаемых новшеств. В крупномасштабных испытаниях особенно важно иметь гарантию того, что будут устранены смещения, вызванные неучтенными факторами, в частности субъективными действиями исследователя. Если у исследователя есть свои фавориты среди проверяемых новых программ, то он может подобрать самые «хорошие» единицы и показать эффективность предлагаемого новшества. В действительности же положительные результаты могут относиться вовсе не к данной программе, а являться эффектом специально подобранных э.е.

Пример 1. Рандомизированное полевое испытание новой программы преподавания физики в вузах (Гарвардский проект по физике)

Целью испытаний являлось оценивание нового курса по физике в высшей школе по сравнению с традиционным изложением. Сравнение курсов производилось по нескольким показателям, характеризующим успеваемость студентов и их эмоциональное отношение к курсу. Новый курс выступал в роли обработки, а существующий служил контролем. В эксперименте участвовали преподаватели и студенты, которым эти преподаватели читали соответствующие курсы физики.

Пример 2. (выбор преподавателей)

Преподаватели выбирались случайным образом из 16911 преподавателей физики, зарегистрированных в Национальной ассоциации США. Тридцати четырем преподавателям случайно назначалась обработка, девятнадцати — контроль. Часть из преподавателей отсеялась вследствие болезней, изменения намерения участвовать в эксперименте. Проводилась специальная подготовка преподавателей, которая состояла в том, что одна группа слушала 6 недель новый курс, для другой проводились двухдневные занятия. Затем в течение учебного года каждый преподаватель читал свой курс студентам того вуза, в котором он работал. Эксперимент завершался экзаменом студентов по физике.

Пример 2. (результаты)

Знание предмета у студентов, прослушавших новый курс, оказалось примерно таким же, как и у контрольной группы. По этому показателю новый курс не превзошел старый. Однако «обработанные» студенты выразили большее удовлетворение новым курсом физики, нашли исторический подход более интересным, учебники более читабельными, а (физику более понятную без обширных знаний по математике. Студенты в контрольной группе считали физику одним из наиболее трудных курсов. Таким образом, по эмоциональным показателям новая программа получила высокую оценку. Проект по физике является примером полевого испытания новой программы по нескольким показателям. Результаты оценки по этим показателям оказались различными. Это говорит о важности многоаспектного оценивания новшеств.

Пример 2. (использование рандомизации)

В данном примере элемент случайности вносился на двух этапах: при выборе преподавателей (а следовательно, и студентов) и при назначении обработки. Это позволило, во-первых, распространить полученные выводы на исходные совокупности преподавателей и студентов, во-вторых, при нерандомизированном испытании неоднозначность интерпретации эффекта новой программы состояла бы в том, что экспериментальный курс могли согласиться читать более способные и энергичные преподаватели, а контрольный - пассивные и менее способные. Тогда эффект курса был бы смешан с эффектом личности преподавателя.

Пример 2. Нерандомизированный эксперимент по улучшению жилищных условий

Цель исследования - определить, как влияет улучшение жилищных условий на физическое и социальное здоровье семей с низкими доходами. Экспериментальная группа из 300 семей получила государственное жилье с улучшенными комнатами, отоплением, санузлом, холодильником и мусоропроводом. Контрольная группа состояла из 300 семей, которые были сопоставимы с экспериментальными по социальным демографическим признакам, но имели плохие жилищные условия. Этот эксперимент не был рандомизирован.

Пример 2. (результаты)

Жилье выделялось муниципалитетом по его усмотрению, а контрольная группа была взята из очереди. Через три года эти две группы сопоставлялись по заболеваемости, социопсихологической приспособляемости и успеваемости детей в школе. В экспериментальной группе смертность значительно снизилась, уменьшилась заболеваемость (для лиц не старше 35 лет), на 1/3 уменьшились несчастные случаи. Дети в экспериментальной группе не улучшили успеваемость, но они имели лучшую посещаемость в школе.

Пример 2. (причины провала)

Общий результат эксперимента можно было считать успешным, если бы эффект жилья не был бы связан с эффектом отбора семей в экспериментальную группу. Дело том, что неизвестно правило отбора, которым руководствовался муниципалитет. Если, например, при выделении жилья предполагалось, что возможности семьи разовьются в улучшенных условиях, то тогда исследование состояло бы в проверке этих способностей у сотрудников муниципалитета, а не в оценивании эффекта предоставления жилья. Такой неопределенности при интерпретации не было бы при проведении рандомизированного эксперимента.

Преимущества рандомизированного эксперимента (1)

1. Главным преимуществом эксперимента и его крупномасштабной разновидности - управляемого полевого испытания - является то, что различия между обработками целиком обусловлены эффектом обработок и случайной ошибкой, которую можно оценить. Такая достоверность не может быть достигнута в других типах исследования, даже если используются самые утонченные методы обработки данных.
2. Эксперимент может включать сложные обработки (комбинации уровней многих управляемых факторов), которые никогда не встречаются в обычных условиях.

Преимущества рандомизированного эксперимента (2)

3. Планирование эксперимента ставит перед нами определенные вопросы, которые в других видах исследований могут быть проигнорированы или оставаться неясными. Это относится к определению цели обработки, а следовательно, и к цели всей программы. Необходимо четко определить, в чем состоит обработка, какие действия и операции нужно совершить, чтобы на практике обработать э. е., как измерять эффект обработки. При этом вскрываются такие аспекты (как положительные, так и отрицательные), которые заранее нельзя было предвидеть. Проведение эксперимента поднимает действительные, а не надуманные проблемы.

4. Эксперимент является надежным источником знаний о причинных связях.

Трудности в проведении эксперимента (1)

1. Стоимость экспериментов, особенно крупномасштабных полевых испытаний, может быть очень велика. Например, семилетний эксперимент по исследованию влияния питания на умственное развитие детей, проведенный в Гватемале, обошелся в 3,5 млн. дол.
2. Чтобы обеспечить обработку э.е., в эксперименте требуются большой штат и ресурсы, совершенно отличные от тех с которыми имеет дело исследователь, занимающийся теоретическим анализом и моделированием. Возникает необходимость в дополнительном персонале для проведения эксперимента и контроля над ним. Могут понадобиться дополнительные помещения, транспортные средства и многое другое.

Трудности в проведении эксперимента (2)

3. Некоторые эксперименты требуют больших затрат времени. Для планирования, проведения и оценки крупномасштабных экспериментов необходимы многие месяцы и годы. Задержка во времени бывает столь велика, что исследователь предпочитает обращаться к ретроспективному анализу (к работе с архивами многолетней давности), если, конечно, накопленные данные содержат необходимую информацию. Как правило, информацию о новшествах они не содержат.

4. Трудности этического порядка, которые могут быть в любых исследованиях (например, наличие щепетильных вопросов при массовых обследованиях), в эксперименте особенно обострены. Сюда входят вопросы справедливости при назначении участников в обрабатываемые и контрольные группы, нарушение привычного хода жизни участников и др.

Рекомендуемая литература

1. Гласс Дж., Стэнли Дж. Статистические методы в педагогике и психологии. Пер. с англ. – М.: Издательство «Прогресс», 1976. -496 с.
2. Маслак А.А. Основы планирования и анализа сравнительного эксперимента в педагогике и психологии. – Курск: РОСИ, 1998. – 167 с.
3. Маслак А.А., Маркова Е.В., Самуйленко А.Я., Сергиенко А.И. Компьютерные системы биотехнологических исследований. – М.: ВНИИТИБП, 1993. – 432 с.
4. Маркова Е.В., Маслак А.А. Рандомизация и статистический вывод. – М.: Финансы и статистика, 1986. – 208 с.