Лекция 12

Проверка гипотез

Введение

При проведении исследований специалистам бывает необходимо получить ответ на вопросы различного типа. Например, исследователю может быть необходимо, установить, действительно ли Земля нагревается. Перед физиологом может возникнуть вопрос, понижает ли предложенная схема лечения кровяное давление или нет. Преподавателю бывает необходимо проверить, дает ли новый метод обучения лучшие результаты, чем прежний. Управляющему магазином необходимо решить, предпочтут ли покупатели новой коллекции одежды вещи определенного цвета. Производителей автомобилей может интересовать, уменьшают ли установленные подушки безопасности серьезность ранений при авариях. Ответы на эти вопросы могут быть получены в рамках раздела статистики, занимающегося *проверкой гипотез* — проверкой истинности утверждений относительно генеральной совокупности, осуществляемой с помощью результатов выборочного наблюдения.

Аппарат проверки гипотез при первом знакомстве является сложным для понимания, в первую очередь из-за единовременного введения большого количества новых понятий. Для того чтобы овладеть ими, рекомендуется при первом чтении прежде всего обратить внимание и подробно разобрать предлагаемые примеры, а при повторном - попытаться осознать суть подхода в целом.

Основные этапы проверки гипотез

Формулировка гипотез

Проверка гипотез начинается с формулировки утверждения.

Определение. ***Статистической гипотезой*** называется утверждение относительно показателя или типа распределения. Это утверждение может быть или не быть правильным.

Первые несколько лекций будут посвящены гипотезам о показателях распределений.

Для того чтобы осуществить проверку, необходимо сформулировать две гипотезы - основную (нулевая) и альтернативную.

Определение. ***Основная гипотеза*** (обозначается H_0) является статистической гипотезой, утверждающей, что нет различий между значением показателя и конкретным числом или двумя различными значениями показателя.

Определение. ***Альтернативная гипотеза*** (обозначается H_1) является статистической гипотезой, утверждающей, что есть различия между значением показателя и конкретным числом или двумя различными значениями показателя.

Для иллюстрации того, как формулируются статистические гипотезы, рассмотрим три ситуации.

Ситуация А. Разработчик лекарственного препарата хочет проверить, будет ли его использование вызывать нежелательные побочные эффекты и, в частности, изменится ли частота пульса у принявших лекарство пациентов. Так как среднее значение пульса у пациентов известно и составляет 82 удара в минуту, то гипотезы формулируются следующим образом:

$$H_0 : m = 82; \quad H_1 : m \neq 82.$$

Основная гипотеза в данном случае утверждает, что среднее значение пульса не изменилось, а альтернативная - что изменилось. Альтернативная гипотеза такого типа называется двусторонней, так как побочные эффекты могут как повысить, так и понизить пульс.

Ситуация Б. Исследователь-химик изобрел добавку, увеличивающую срок службы автомобильных аккумуляторов. Если средний срок службы батарей выпускаемой сейчас модели составляет 36 месяцев, то пара гипотез формулируется так:

$$H_0 : m = 36; \quad H_1 : m > 36.$$

В этой ситуации исследователя интересует только вопрос увеличения срока службы, поэтому альтернативную гипотезу он формулирует, как средний срок службы больше 36 месяцев. Альтернативная гипотеза такого типа называется односторонней.

Ситуация В. Руководитель строительной компании хочет добиться снижения расходов на отопление возводимых коттеджей за счет применения новых теплоизолирующих материалов. Если до начала применения средний счет за отопление составлял 480 рублей, то утверждения относительно эффективности изоляции формулируются так:

$$H_0 : m = 480; \quad H_1 : m < 480 .$$

И в этом случае альтернативная гипотеза называется односторонней, так как вопрос ставится только о снижении расходов.

Общие формулировки трех типов рассмотренных гипотез (гипотез о среднем) представлены в табл. 1.

Таблица 1. *Формулировки гипотез различного типа*

Проверка правосторонней гипотезы	Проверка левосторонней гипотезы	Проверка двусторонней гипотезы
$H_0 : m = m_0$	$H_0 : m = m_0$	$H_0 : m = m_0$
$H_1 : m > m_0$	$H_1 : m < m_0$	$H_1 : m \neq m_0$

Для того чтобы корректно сформулировать гипотезы, исследователь должен уметь транслировать словесные утверждения в стандартные математические отношения. В табл. 2. приведен ряд выражений, часто встречающихся в задачах на проверку гипотез, и соответствующие им математические символы.

Таблица 2. *Выражения, часто встречающиеся при проверке гипотез*

Типы утверждений	
>	<
больше, чем длинней, чем шире, чем	меньше, чем короче, чем уже, чем

выше, чем лучше, чем эффективнее, чем дороже, чем	ниже, чем хуже, чем менее эффективно дешевле, чем
=	≠
Равен точно соответствует точно совпадает не изменился	не равен отличается не совпадает изменился
≥	≤
больше или равен по меньшей мере не меньше	меньше или равен не превосходит не больше

Пример. Сформулировать основную и альтернативную гипотезы.

1) Исследователь полагает, что разработанный им витаминный комплекс для беременных позволит повысить средний вес новорожденных. Средний вес для генеральной совокупности на данный момент составляет 3.3 кг.

2) Инженер считает, что применение роботов вместо людей на ряде операций по изготовлению компакт-дисков позволит уменьшить долю брака. В настоящее время среднее количество бракованных дисков составляет 18 на 1000.

3) У психолога появилось предположение, что если при проведении теста в помещении играет спокойная музыка, то это оказывает влияние на результат. По предшествующим тестам средний результат составляет 74 балла.

Решение.

1) $H_0: m = 3.3;\quad H_1: m > 3.3$;

2) $H_0: m = 18;\quad H_1: m < 18$;

3) $H_0: m = 74;\quad H_1: m \neq 74$.

Виды ошибок при проверке гипотез

После того как гипотезы сформулированы, исследователь на следующем этапе должен провести выборочное наблюдение и собрать данные о значении признака в выборке. Например, в ситуации А исследователь формирует группу пациентов, дает им препарат и через некоторое время, когда он подействует, измеряет значение пульса и рассчитывает среднее по выборке значение $\overline{x}$. Однако в большинстве случаев выборочное среднее будет отличаться от генерального среднего, даже если в действительности препарат не меняет частоту пульса и генеральное среднее остается неизменным - 82 удара в минуту. Поэтому возникает вопрос — если выборочное среднее всегда отличается от генерального, то как же решить вопрос о влиянии препарата? Как понять, чем обусловлены отличия в данном случае - случайностью или действием лекарства?

Если средняя величина пульса по выборке составит, скажем, 83 удара в минуту, то исследователь, вероятно, решит, что различия вызваны случайностью, и примет основную гипотезу. Если, с другой стороны, средний пульс окажется 90 ударов, то исследователь предположит, что увеличение связано с действием препарата и нулевую гипотезу нужно отклонить. Возникает, однако, следующий вопрос — где же провести границу между двумя ситуациями? Решение осуществляется не с помощью интуиции, а с помощью статистического аппараты. Для того чтобы отклонить основную гипотезу $H_0 : m = 82;$, различия между m и $\overline{x}$ должны быть велики или ***статистически значимы***.

При проверке гипотез с точки зрения правильности принятого решения может осуществиться один из четырех случаев, табл. 3. С одной стороны, гипотеза H_0 в действительности, в генеральной совокупности, может быть или не быть верной (заголовки колонок). С другой стороны, на основании выборочной совокупности может быть выбрано одно из двух решений - отклонить или принять гипотезу H_0 (заголовки строк). На “пересечении” этих возможностей и возникают 4 случая, в двух из которых принятое решение оказывается правильным, а в двух - ошибочным.

***Таблица**.3. Возможные результата проверки гипотез*

	H_0 верна	H_0 не верна
H_0 отклоняется	Ошибка I рода	Правильное решение
H_0 принимается	Правильное решение	Ошибка II рода

Если основная гипотеза верна, но мы ее отклоняем, тогда совершается *ошибка I рода.* В ситуации А это случай, когда в среднем по генеральной совокупности препарат не вызывает серьезных изменений пульса у пациентов, но, в силу случайного совпадения, значительно изменит его у тех из них, кто попал в выборку. В этом случае исследователь отклонит основную гипотезу, тогда как на самом деле она верна, тем самым совершив *ошибку I рода*.

С другой стороны, может сложиться и так, что у пациентов в выборке препарат существенных изменений пульса не вызовет, тогда как на самом деле, если бы были рассмотрены все пациенты в генеральной совокупности, были бы выявлены его существенное снижение или увеличение. Однако, базируясь на выборке, исследователь принимает решение – гипотезу H_0 принять (считать, что препарат на пульс не действует), тем самым совершая *ошибку II рода.*

Определение.

1) *Ошибка I рода* совершается, когда основная гипотеза верна, но отклоняется.

2) *Ошибка II рода* совершается, если основная гипотеза неверна, но принимается.

Важно хорошо понять, что принятое на основании выборки решение никоим образом не является доказательством. Единственный способ в статистике что-то доказать - это исследовать всю генеральную совокупность, что в большинстве случаев (в том числе во всех трех рассмотренных ситуациях) невозможно. Во всех остальных ситуациях решение носит вероятностный характер и всегда остается возможность ошибиться, поэтому до принятия решения необходимо задать, какую вероятность ошибки мы можем или хотим допустить.

Определение. *Уровень значимости* α - это максимальная вероятность совершить ошибку первого рода.

Наиболее часто в статистике применяют уровни значимости $\alpha = 0.01$(однопроцентный) и $\alpha = 0.05$ (пятипроцентный). Их использование означает, что вероятность отклонить основную гипотезу, когда нас самом деле она верна, будет составлять 1 или 5%.

Статистический критерий принятия решения

Ключевым элементом при проверке гипотез является критерий.

Определение. *Критерий* - величина, рассчитываемая по выборочным значениям, с помощью которой принимается решение относительно верности гипотез.

Именно критерий является количественной мерой того, насколько значимым является отличие выборочного значения показателя от его предполагаемой генеральной величины. Конкретная функциональная форма критерия строится для каждого типа гипотез в отдельности. Основной характеристикой критерия служит его *закон распределения.* Интерпретация критерия основана на понятии критических значений и критических областей.

Определение. *Критическое значение* - значение критерия, разделяющее критическую и некритическую область. Величина критического значения определяется уровнем значимости.

Определение. *Критическая область* - диапазон значений критерия, при которых отличия выборочного и тестируемого значения считаются значимыми и основная гипотеза отклоняется.

Определение. *Некритическая область* - диапазон значений критерия, при которых отличия выборочного и тестируемого значения полагаются случайными и основная гипотеза не отклоняется.

Проиллюстрируем все введенные понятия на примере гипотез о среднем, проверяемые в ситуациях А, Б, В, и предположим, что величина среднего квадратического отклонения генеральной совокупности σ известна. В этом случае критерием для принятия решения является уже знакомая по разделу "Оценки" величина $z=\frac{\overline{x}-m_0}{\sigma/\sqrt{n}}$, где $\overline{x}=\frac{1}{n}\left(x_1+x_2+x_3+\ldots+x_n\right)$ - выборочное среднее, m_0 - предполагаемое значение генеральной средней. Для различных выборок значения z будут приближенно распределены по стандартному нормальному закону $N(0;1)$, а критическими значениями для двусторонней (односторонних) гипотез являются величины $z_{\alpha/2}$ (z_α), определяемые по уровню значимости α.

Возьмем уровень значимости $\alpha=0.01$ и осуществим проверку гипотез в ситуации А. Критические значения, расположение критических и некритической областей показаны на рис. 1. Если выборочное значение далеко отклоняется от предполагаемого то $m_0=82$ в большую или меньшую сторону, то значение критерия попадет в критическую область $|z|>2.58$, и основная гипотеза будет отклонена в пользу альтернативной. Какова вероятность при таком решении совершить ошибку первого рода, то есть отклонить основную гипотезу, когда она верна? Рассуждаем следующим образом. Предположим, что нулевая гипотеза верна и истинное значение генерального среднего равно 82. Тогда критерий z имеет стандартное нормальное распределение и вероятность его попадания в критическую область $|z|>2.58=z_{\alpha/2}$, равна $\alpha=0.01$. Так как при попадании в критическую область основная гипотеза отклоняется, вероятность совершить ошибку первого рода также равна α. Следовательно, критическая область построена правильно.

При том же уровне значимости рассмотрим правую одностороннюю гипотезу, ситуация Б, рис. 2. Величина z_α определяется из условия $P(z > z_\alpha) = \alpha$. Если выборочное значение намного больше $m_0 = 36$, то значение критерия попадет в критическую область $z > 2.33 = z_\alpha$, основная гипотеза отклоняется и принимается альтернативная. Во всех остальных случаях основная гипотеза не отклоняется. Какова вероятность отклонить истинную основную гипотезу? Рассуждая аналогично предыдущему случаю, получаем, что она, по определению z_α будет равна уровню значимости $\alpha = 0.01$.

При проверке левой односторонней гипотезы с тем же уровнем значимости, ситуация В, критической будет область $z < -2.33 = z_\alpha$, рис. 3. В случае, если выборочное значение среднего будет намного меньше 480, критерий окажется в критической области и основная гипотеза будет отклонена.

Общая схема рассуждений для гипотез всех трех типов представлена на блок-схеме рис. 4.

Следует пояснить дополнительно один вопрос, связанный с односторонними гипотезами. При проверке, например, правосторонней гипотезы критическая область слева отсутствует, то есть даже для выборочных средних намного меньших, чем m_0, все равно будет принята основная гипотеза $H_0 : m = m_0$. Объясняется это тем, что хотя при истинном среднем $m = m_0$ вероятность уклониться от m_0 далеко влево мала, но при любом из значений $m > m_0$ вероятность будет еще меньше. Для односторонних гипотез во многих случаях удобнее вести рассуждение относительно альтернативной гипотезы: если утверждение $H_1 : m > m_0$, m <u>больше</u> m_0 отклоняется, то принимается противоположное ему утверждение, m <u>не больше</u> m_0, $m \le m_0$.

В заключение следует сделать одно важное замечание. Случайно ли, что при проверке гипотезы о среднем мы столкнулись с той же самой величиной

z характеризующей распределение выборочной средней, что и при построении доверительного интервала? На самом деле между двумя типами задач - построением доверительных интервалов и проверкой гипотез существует тесная связь.

С чего начинается статистическое исследование - с теоретического построения гипотез или эмпирического выборочного исследования, организации наблюдения? В действительности теория и практика тесно связаны друг с другом и подобные вопросы не возникают. С одной стороны, можно предположить, что сначала формулируются гипотезы и выборочное исследование проводится с целью его проверить. В этом случае мы пользуемся схемой проверки гипотез. Если же сначала проведено выборочное исследование, то затем, построив доверительный интервал, легко определить, с какими из предполагаемых теоретических значений данные эксперимента совместимы, а с какими нет.

Рис. 1. *Проверка двусторонней гипотезы*

Рис. 2. *Проверка правосторонней гипотезы*

Рис. 3. *Проверка левосторонней гипотезы*

Рис. 4. *Схема рассуждений при проверке гипотез*

Алгоритм проверки гипотез

Проверку гипотез большинства типов удобно разделить на следующие этапы:

1) По условиям задачи выбрать тип гипотез, сформулировать основную и альтернативную гипотезы.

2) Задать уровень значимости и найти критические значения для критерия, соответствующего типу гипотез.

3) Рассчитать значение критерия по данной выборке.

4) Принять решение, отклоняется или нет основная гипотеза.

5) Сделать вывод.

Пример. Исследователь должен проверить утверждение, что среднее время (в неделю), которое студенты колледжа смотрят телевизор, меньше, чем среднее время по стране, составляющее $m = 29.4$ часа при среднем квадратическом отклонении $\sigma = 2$ часа. Для проверки опрошено 25 студентов, у которых среднее время просмотра телевизора составило $\overline{x} = 27$ часов. Проверить утверждение при уровне значимости $\alpha = 0.01$.

Решение:

1) Применяем левостороннюю гипотезу о среднем $H_0 : m = 29.4$ $H_1 : m < 29.4$.

2) Для проверки гипотезы о среднем применяется z - критерий, для левосторонней гипотезы его критическое значение составляет $z_\alpha = -2.33$.

3) Находим расчетное значение критерия z по выборке

$$z = \frac{\overline{x} - m_0}{\sigma / \sqrt{n}};\ m_0 = 29.4;\ \sigma = 2;\ n = 25;\ \overline{x} = 27;\ z = \frac{27 - 29.4}{2 / \sqrt{25}} = -6.$$

4) Принятие решения. Так как -6 < -2.33, то значение критерия находится в критической области и гипотеза H_0 отклоняется (принимается H_1).

5) **Вывод.** Студенты колледжа в среднем смотрят телевизор меньше, чем население в страны в целом.

Комментарий. Могло так оказаться, что только выбранные студенты смотрят телевизор меньше, чем другие люди. Но если выборка сделана корректно (например, по схеме случайного отбора), то вероятность, что среднее для 25 человек окажется 27 при истинном генеральном среднем 29.4, очень мала — менее 1%. Поэтому мы пренебрегаем этой возможностью и утверждаем - это отличие не случайно, студенты действительно смотрят телевизор меньше.

Пример. Директор магазина молодежной одежды полагает, что средний возраст покупателей одной из моделей джинсов составляет 15 лет. Для проверки утверждения опрошены 36 покупателей, средний возраст которых составил 15.6 года. Проверить гипотезу директора при уровне значимости 1%, если среднее квадратическое отклонение возраста покупателей известно и равно 1 году.

Решение:

1) Применяем двустороннюю гипотезу о среднем $H_0 : m = 15$; $H_1 : m \neq 15$.

2) Для проверки гипотезы о среднем применяется z - критерий, для двусторонней гипотезы его критическое значение составляет $z_{\alpha/2} = 2.58$.

3) Находим значение критерия для данной выборки

$$z = \frac{\overline{x} - m_0}{\sigma / \sqrt{n}};\ m_0 = 15;\ \sigma = 1;\ n = 36;\ \overline{x} = 15.6;\ z = \frac{15.6 - 15}{1 / \sqrt{36}} = 3.6$$

4) Принятие решения. Условие $z > z_{\alpha/2}, 3.6 > 2.58$ показывает, что значение критерия попало в критическую область и гипотеза H_0 должна быть отклонена.

5) **Вывод.** Средний возраст покупателей магазина не равен 15 годам.

Следующая лекция:

Гипотезы о значении показателя. Гипотеза о среднем