

Лекция 3. Проверка статистических гипотез и оценка статистической взаимосвязи генеральных совокупностей

Вопросы лекции:

1. Проверка статистических гипотез
2. Проверка гипотез о законе распределения. Критерий согласия Пирсона
3. Оценка статистической взаимосвязи генеральных совокупностей

Вопрос 1. Проверка статистических гипотез

Под **статистической гипотезой** понимают всякое высказывание о генеральной совокупности, которое можно проверить по выборке. Как правило, статистические гипотезы делят на гипотезы о законах распределения и гипотезы о параметрах распределения.

Основные гипотезы:

- поведение случайной величины с течением времени не изменяется;
- случайная величина (генеральная совокупность) обладает некоторой функцией плотности вероятности;
- существуют моменты: математическое ожидание, дисперсия

Задача: оценить эти характеристики по набору наблюдаемых значений, которые называются выборкой.

Предполагается, что для оценивания характеристик используется выборка конечного объема $X_1 \dots X_n$, где X_i – выборочное значение случайной величины при i - том наблюдении. (например, выборка значений рентабельности деятельности 100 строительных предприятий Москвы, при этом все строительные предприятия Москвы – генеральная совокупность) (Приведите примеры выборок и соответствующих генеральных совокупностей).

Общая схема проверки гипотез

1) Выдвигается гипотеза о некотором свойстве генеральной совокупности, эта гипотеза называется основной или нулевой, обозначается H_0 . Формулируется также альтернативные гипотезы, которые в какой-либо мере являются отрицанием основной гипотезы. Если альтернативная гипотеза единственная, то она обозначается H_1 .

2) Выбирается критерий, который позволяет сравнить выдвинутые гипотезы на предмет их противоречивости выборочным данным.

Пусть $z=g(x)$ – критерий сравнения выборочных совокупностей, для которого известна функция плотности вероятности при нулевой гипотезе, то есть: $W_g(Z/H_0)$. Предположим, что наша гипотеза состоит в утверждении о некотором свойстве генеральной совокупности. Назовём эту гипотезу нулевой и обозначим её H_0 . (например, на основе анализа выборки строительных предприятий сделан вывод о снижении эффективности инвестиционно-строительной деятельности. Верна ли эта гипотеза для всей совокупности строительных предприятий России? Все предприятия в

данном случае – генеральная совокупность. Нулевая гипотеза – снижение эффективности строительной деятельности в России. Для разных предприятий вероятность утверждения различна, но у большинства наблюдается снижение эффективности, то есть вероятность и плотность вероятности известны. Приведите альтернативную гипотезу)

Альтернативной, или конкурирующей гипотезой, которую обозначим H_1 , будет $H = H_1$. Перед нами стоит задача проверки гипотезы H_0 относительно конкурирующей гипотезы H_1 на основании выборки, состоящей из p независимых наблюдений $X_1, X_2, \dots, X_p$. (наблюдение эффективности деятельности другой выборки предприятий в нашем примере, тех, представителей которых не опрашивали в предыдущий раз) Следовательно, всё возможное множество выборок объёма p можно разделить на два непересекающихся подмножества (O и W) таких, что проверяемая гипотеза H_0 должна быть отвергнута, если наблюдаемая выборка попадает в подмножество W (множество предприятий, у которых эффективность не снизилась) и принята, если выборка принадлежит подмножеству O (множество предприятий, у которых интенсивность снизилась). Границы подмножества O составляют $[a_\alpha; b_\alpha]$.

Осуществляется регистрация выборочных значений и вычисляется значение критерия $g(\bar{x})$. Принимаются решения:

Если $a_\alpha < g(\bar{x}) < b_\alpha$, то H_0 – не противоречит данным; если $g \notin [a_\alpha; b_\alpha]$, то H_0 противоречит данным и поэтому отвергается на уровне значимости α .

Подмножество O называют **областью допустимых значений**, а подмножество W – **критической областью**. При формировании критической области возможны ошибки.

Ошибка первого рода состоит в том, нулевая гипотеза отвергается, то есть принимается гипотеза H_1 , в то время как в действительности верна гипотеза H_0 .

Ошибка второго рода состоит в том, что принимается гипотеза H_0 , а в действительности верна гипотеза H_1 .

В качестве критерия обычно выступают некоторые статистики, то есть функции от выборочных значений.

Любое оценивание предполагает осуществление некоторых вычислений с использованием выборочных значений, то есть, если оценивается некоторая характеристика «а», то предполагается получение её оценки как некоторой функции от наблюдений.

$$a \rightarrow \widehat{a}_m = q(x_1, \dots, x_m) \quad (3.1)$$

Эти функции называются статистиками. При оценивании желательно достичь наилучшего соответствия с оцениваемой величиной. Понятие соответствие по существу означает качество оценки.

Иногда в литературе понятие «оценки» и «оценивающей функции» отождествляются и поэтому говорят о требованиях к оценкам. Вместе с тем, следует иметь в виду, что оценивающие функции получаются (выбираются) априорно, то есть, ещё до проведения опыта, а оценка – некоторое число, которое получается с использованием оценивающей функции от выборочных значений.

Обычно, оценивающим функциям предъявляются следующие требования:

1) Требование несмещённости. Так как оценка вычисляется как функция случайной величины, то она до опыта должна рассматриваться как случайная величина с некоторой функцией плотности вероятности. Обычно, оценивающие функции обладают математическим ожиданием и дисперсией. Требование означает выполнение условия:

$$E(g(x_1, \dots, x_n)) = E(\widehat{a}_n) = a \quad (3.2)$$

то есть математическое ожидание оценки должно быть равно оцениваемой величине.

2) Требование состоятельности – означает, что при

$$n \rightarrow \infty \Rightarrow \widehat{a}_n \rightarrow a \quad (3.3)$$

то есть оценка сходится к оцениваемой величине при неограниченном увеличении длины выборки. Если дисперсия оценки стремится к нулю, то это называется **среднеквадратической сходимостью**.

3) Требование эффективности. Оценка называется эффективной, если оценочная функция обладает наименьшей дисперсией. Таким образом, среди всех оценочных функций следует выбирать такую, которая удовлетворяет требованию:

$$E(g(x_1, \dots, x_n) - a)^2 = \min \quad (3.4)$$

Основные методы построения оценочных функций

1. Метод аналогии. Аналогия заключается в выполнении процедуры усреднения, причем усреднение с функцией плотности вероятности заменяется на усреднение по количеству наблюдений.

2. Метод моментов. Идея этого метода заключается в приравнивании теоретических и эмпирических моментов.

3. Метод максимума правдоподобия (термин был впервые использован в работе Фишера, 1922) - это общий метод оценивания параметров генеральной совокупности с помощью максимизации правдоподобия (L) выборки. Метод предназначен для оценивания численных значений параметров, от которых зависит функция плотности вероятности.

Следует иметь в виду, что не для всех видов функций плотности вероятности метод максимального правдоподобия применим. Одно из главных условий заключается в том, что производные функции плотности вероятности по каждому из оцениваемых параметров должны существовать во всём диапазоне изменений значений случайной величины. Например, в равномерной функции плотности вероятности производные не существуют на концах интервалов, поэтому этот метод здесь не применяется.

Сравнение центров распределения нормальных генеральных совокупностей

Имеем две случайные величины X и Y . Обе подчиняются нормальному закону распределения. Допустим, что мы располагаем двумя независимыми выборками объёмами m и n соответственно. Нулевая гипотеза : $M(X)=M(Y)$. (математические ожидания величин равны) За альтернативную гипотезу примем $|M(X) - M(Y) > 0|$

Для проверки исходного предположения о равенстве математического ожидания используется статистика вида:

$$t_v = \frac{|\bar{x} - \bar{y}|}{\sqrt{(m-1)\tilde{\delta} + (n-1)\delta}} \cdot \sqrt{\frac{(m+n-2)}{m+n}} \sqrt{mn} \quad (3.5)$$

Эта статистика называется статистикой Стьюдента с $v = m + n - 2$ - степенями свободы. Для неё составлены таблицы соответствующих им уровней значимости. Если вычисленная статистика Стьюдента превышает значение, соответствующее выбранному уровню значимости, то делается вывод о том, что исходная гипотеза о равенстве математических ожиданий двух генеральных совокупностей отвергается на уровне значимости - α . Обычно α выбирается из диапазона $0,001 \leq \alpha \leq 0,1$.

Например, таким образом можно проверить равенство средней ожидаемой прибыли по нескольким проектам, либо до и после осуществления каких-то мероприятий.

Вопрос 2. Проверка гипотез о законе распределения. Критерий согласия Пирсона.

Существует несколько критериев согласия для проверки законов распределения случайной величины. Это критерии Колмогорова, Смирнова, χ^2 Пирсона и др. Критерий Пирсона – это наиболее часто употребляемый критерий для проверки закона распределения случайной величины. Нулевая гипотеза заключается в том, что делается предположение о конкретном типе функции плотности вероятности.

Схема проверки:

1. Выдвигается гипотеза и виде закона распределения.
2. Проводится n – измерений значения случайной величины.
3. Строится гистограмма.
4. Оцениваются необходимые параметры закона распределения.
5. Вычисляются значения P_i , согласно формуле:

$$P_i = P\{\varepsilon \in D_i\} = \int_{a_i}^{b_i} W_\varepsilon(x) dx \quad (3.6)$$

$$D_i = [a_i; b_i]$$

в которой подставляются значения границ интервалов гистограммы и оцененных параметров.

6. Проводится вычисление по формуле:

$$\chi_{L-1}^2 = \sum_{i=1}^L P_i \left(\frac{P_i - v_i/n}{P_i} \right)^2 \quad (3.7)$$

Эту статистику называют критерием согласия Пирсона, то есть критерием того, что исходное предположение о виде закона распределения согласуется с экспериментальными данными. Для распределения χ^2 также составлены таблица квантилей, которые соответствуют выбранным уровням значимости.

7. Выбирается уровень значимости α и определяется по таблицам h_α .

8. Вычисленные значения сравниваются с h_α и выносится решение о том, что значение выборки противоречит исходной гипотезе, если выполняется неравенство:

$$\chi_{L-1}^2 > h_\alpha \quad (3.8)$$

иначе говорят, что исходная гипотеза о виде закона распределения противоречит выборочным данным на уровне значимости α .

Вопрос 3. Оценка статистической взаимосвязи генеральных совокупностей

Достаточно часто, особенно в экономике, приходится рассматривать несколько генеральных совокупностей, которые могут оказывать взаимное влияние. Возникает проблема описания этого влияния с помощью математических моделей. В противном случае речь идет о проверке гипотезы об отсутствии взаимосвязи между двумя генеральными совокупностями на основе статистических данных. Если нулевая гипотеза отвергается, то принимается гипотеза о наличии влияния одной генеральной совокупности на другую (о их взаимосвязи).

Итак, гипотеза имеет вид: $\{H_0 : \varepsilon_2 \text{ не зависит от } \varepsilon_1 \text{ и/или наоборот } \varepsilon_1 \text{ не зависит от } \varepsilon_2, \text{ но } \varepsilon_1: x_1, \dots, x_n; \varepsilon_2: y_1, \dots, y_n\}$.

Для проверки ее справедливости предполагается использовать выборочные значения. Так как по предположению генеральные совокупности – являются случайными величинами, то можно говорить о вероятностных моделях, описывающих влияние генеральной совокупности. То есть следует взаимозависимость понимать в вероятностном смысле.

Моделирование взаимосвязей случайной величины требует рассмотрения многомерных распределений вероятностей.

По аналогии с одномерным случаем можно ввести функцию плотности вероятности совокупности двух случайных величин (генеральных совокупностей) $(\varepsilon_1, \varepsilon_2)$:

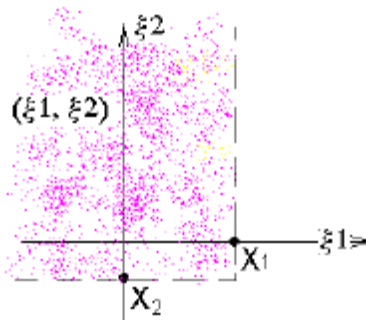


Рис. 3.1. Графическое изображение функции плотности вероятностей

$$F_2(x_1, x_2) = P\{\varepsilon_1 < x_1 \cap \varepsilon_2 < x_2\} \quad (3.9)$$

Кроме понятия независимости, которое сводится к произведению функции плотности вероятности, имеется несколько (очень мало) конкретных моделей двумерной функции плотности вероятности, которые позволяют описать в том числе зависимость между случайными величинами. (корреляция или ковариация).

Исследование взаимосвязи случайной величины с прикладной точки зрения целесообразно потому, что может позволить предсказывать значение одних случайных величин по значению других. Особенно это полезно тогда, когда некоторые величины могут быть ненаблюдаемыми.

Контрольные вопросы по теме 3:

1. Что такое оценочные функции?
2. Какие существуют методы построения оценочных функций?
3. Охарактеризуйте процесс проверки гипотез.
4. Какие функции распределения Вы знаете?
5. Что такое критерии согласия?