

## СТАТИСТИЧЕСКИЕ ГИПОТЕЗЫ И ИХ ПРОВЕРКА

1. Понятие статистической гипотезы
2. Проверка статистических гипотез

1. **Понятие статистической гипотезы.** *Статистической гипотезой* называют научное предположение о тех или иных статистических законах распределения рассматриваемых случайных величин, которое может быть проверено на основе выборки.

Гипотезы в свою очередь классифицируются на:

- простые и сложные;
- параметрические и непараметрические;
- основные (высказанные) и альтернативные (конкурирующие).

Если выдвигаемая гипотеза сводится к утверждению о том, что значение некоторого неизвестного параметра генеральной совокупности в точности равно заданной величине, то эта гипотеза называется **простой**.

**Сложной** называют гипотезу, которая состоит из конечного или бесконечного множества простых гипотез, при этом указывается некоторая область вероятных значений параметра.

Гипотезы о параметрах генеральной совокупности называются **параметрическими**, о распределениях — **непараметрическими**.

Выдвинутая гипотеза называется **нулевой (основной)**. Ее обозначают  $H_0$ . При этом предполагается, что действительное различие сравниваемых величин равно нулю, а выявленное по данным отличие от нуля носит случайный характер. Нулевая гипотеза отвергается тогда, когда по выборке получается результат, который при истинности выдвинутой нулевой гипотезы маловероятен.

По отношению к высказанной (основной) гипотезе всегда можно сформулировать **альтернативную (конкурирующую)**, противоречащую ей. Альтернативную (конкурирующую) гипотезу принято обозначать  $H_1$ .

По своему содержанию статистические гипотезы подразделяются на несколько основных типов:

- гипотезы о виде закона распределения исследуемой случайной величины;
- гипотезы о числовых значениях параметров исследуемой генеральной совокупности;
- гипотезы об однородности двух или нескольких выборок или некоторых характеристик анализируемых совокупностей;

- гипотезы об общем виде модели, описывающей статистическую зависимость между признаками.

Так как проверка статистических гипотез осуществляется на основании выборочных данных, решения относительно нулевой гипотезы  $H_0$  имеют вероятностный характер. Следовательно, такое решение сопровождается некоторой, хотя возможно и очень малой, вероятностью ошибочного заключения как в ту, так и в другую сторону. То есть в небольшой доле случаев а нулевая гипотеза  $H_0$  может оказаться отвергнутой, в то время как в действительности в генеральной совокупности она является справедливой. *Такую ошибку называют ошибкой 1-го рода, а ее вероятность – уровнем значимости и обозначают  $\alpha$ .*

В какой-то небольшой доле случаев нулевая гипотеза  $H_0$  принимается, в то время как на самом деле в генеральной совокупности она ошибочна, а справедлива альтернативная гипотеза  $H_x$ . *Такую ошибку называют ошибкой 2-го рода. Вероятность ошибки 2-го рода обозначается как  $\beta$ . Вероятность 1 -  $\beta$  называют мощностью критерия.*

При фиксированном объеме выборки можно выбрать по своему усмотрению величину вероятности только одной из ошибок  $\alpha$  или  $\beta$ . Увеличение вероятности одной из них приводит к снижению другой.

Принято задавать вероятность ошибки 1-го рода  $\alpha$  – уровень значимости. Как правило, пользуются некоторыми стандартными значениями уровня значимости  $\alpha$ : 0,1; 0,05; 0,025; 0,01; 0,005; 0,001. В агрономических исследованиях обычно используется 0,05 или 5% уровень значимости.

В отношении свойств генеральной совокупности могут выдвигаться некоторые гипотезы о величине средней, дисперсии, характере распределения, форме и тесноте связи между переменными.

Проверка гипотезы осуществляется на основе выявления согласованности эмпирических данных с гипотетическими (теоретическими). Если расхождение между сравниваемыми величинами не выходит за пределы случайных ошибок, гипотезу принимают. При этом не делается никаких заключений о правильности самой гипотезы, речь идет лишь о согласованности сравниваемых данных.

**2. Проверка статистических гипотез.** Проверка статистических гипотез осуществляется с помощью статистического критерия. *Статистический критерий* – это правило (формула), по которому определяется мера расхождения результатов выборочного наблюдения с высказанной гипотезой  $H_0$ .

Как уже отмечалось выше, следует иметь в виду, что статистическая проверка гипотез имеет вероятностный характер, так как принимаемые выводы основываются на изучении свойств распределения случайной переменной по данным выборки, а потому всегда существует риск допустить ошибку. Однако с помощью статистической проверки гипотез можно определить вероятность принятия ложного решения. Если вероятность последнего невелика, то можно считать, что применяемый критерий обеспечивает малый риск ошибки.

При проведении проверки статистических гипотез в первую очередь приходится решать задачи статистической проверки гипотез о:

- 1) принадлежности «выделяющихся» единиц исследуемой выборочной совокупности генеральной совокупности;
- 2) виде распределения изучаемых признаков;
- 3) величине средней арифметической и доли;
- 4) наличии и тесноте связи между изучаемыми признаками;
- 5) о форме корреляционной связи.

В статистике в настоящее время имеется большое число критериев для проверки практически любых гипотез. При этом основные принципы их построения и применения являются общими. Для построения статистического критерия, позволяющего проверить некоторую гипотезу, необходимо следующее:

- 1) сформулировать проверяемую гипотезу  $H_0$ . Наряду с проверяемой гипотезой формулируется также конкурирующая (альтернативная) гипотеза;
- 2) выбрать уровень значимости  $\alpha$ , контролирующий допустимую вероятность ошибки первого рода;
- 3) определить область допустимых значений и так называемую критическую область;
- 4) принять то или иное решение на основе сравнения фактического и критического значений критерия.

Проверка статистических гипотез складывается из следующих этапов:

- формулируется в виде статистической гипотезы задача исследования;
- выбирается статистическая характеристика гипотезы;
- выбираются испытываемая и альтернативная гипотезы на основе анализа возможных ошибочных решений и их последствий;
- определяются область допустимых значений, критическая область, а также критическое значение статистического критерия ( $t$ ,  $F$ ,  $\chi^2$ ) по соответствующей таблице;

- вычисляется фактическое значение статистического критерия;
- проверяется испытуемая гипотеза на основе сравнения фактического и критического значений критерия, и в зависимости от результатов проверки гипотеза либо отклоняется, либо не отклоняется.

Проверка каждого типа статистических гипотез осуществляется с помощью соответствующего критерия.