СТАТИСТИЧЕСКАЯ ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Статистическая гипотеза. Нулевая и конкурирующая, простая и сложная гипотезы

Часто необходимо знать закон распределения генеральной совокупности. Если закон распределения неизвестен, но имеются основания предположить, что он имеет определенный вид (назовем его A), выдвигают гипотезу: генеральная совокупность распределена по закону A. Таким образом, в этой гипотезе речь идет о *виде предполагаемого распределения*.

Возможен случай, когда закои распределения известен, а его параметры неизвестны. Если есть основания предположить, что неизвестный параметр Θ равен определенному значенню Θ_0, выдвигают гипотезу: $\Theta=\Theta_0$. Таким образом, в этой гипотезе речь идет о *предполагаемой величине параметра* одного известного распределения.

Возможны и другие гипотезы: о равенстве параметров двух или нескольких распределений, о независимости выборок и многие другие.

Возможны и другие гипотезы: о равенстве параметров двух или нескольких распределений, о независимости выборок и многие другие.

Статистической называют гипотезу о виде неизвестного распределения, или о параметрах известных распределений.

Например, статистическими являются гипотезы:

1) генеральная совокупность распределена по закону Пуассона;

2) дисперсии двух нормальных совокупностей равны между собой.

В первой гипотезе сделано предположение о виде неизвестного распределения, во второй—о параметрах двух известных распределений.

Гипотеза «на Марсе есть жизнь» не является статистической, поскольку в ней не идет речь ни о виде, ни о параметрах распределения.

Наряду с выдвинутой гипотезой рассматривают и противоречащую ей гипотезу. Если выдвинутая гипотеза будет отвергнута, то имеет место противоречащая гипотеза. По этой причине эти гипотезы целесообразно различать.

Нулевой (основной) называют выдвинутую гипотезу H_0.

Конкурирующей (альтернативной) называют гипотезу H_1, которая противоречит нулевой.

Например, если нулевая гипотеза состоит в предположении, что математическое ожидание a нормального распределения равно 10, то конкурирующая гипотеза, в частности, может состоять в предположении, что $a \neq 10$. Коротко это записывают так: $H_0: a = 10$; $H_1: a \neq 10$.

Различают гипотезы, которые содержат только одно и более одного предположений.

Простой называют гипотезу, содержащую только одно предположение. Например, если λ — параметр показательного распределения, то гипотеза $H_0: \lambda = 5$ — простая. Гипотеза H_0: математическое ожидание нормального распределения равно 3 (σ известно) — простая.

Сложной называют гипотезу, которая состоит из конечного или бесконечного числа простых гипотез. Например, сложная гипотеза $H: \lambda > 5$ состоит из бесчисленного множества простых вида $H_i: \lambda = b_i$, где b_i — любое число, большее 5. Гипотеза H_0: математическое ожидание нормального распределения равно 3 (σ неизвестно) — сложная.

Ошибки первого и второго рода

Выдвинутая гипотеза может быть правильной или неправильной, поэтому возникает необходимость ее проверки. Поскольку проверку производят статистическими методами, ее называют *статистической*. В итоге статистической проверки гипотезы в двух случаях может быть принято неправильное решение, т. е. могут быть допущены ошибки двух родов.

Ошибка первого рода состоит в том, что будет отвергнута правильная гипотеза.

Ошибка второго рода состоит в том, что будет принята неправильная гипотеза.

Подчеркнем, что последствия этих ошибок могут оказаться весьма различными. Например, если отвергнуто правильное решение «продолжать строительство жилого дома», то эта ошибка первого рода повлечет материальный ущерб; если же принято неправильное решение «продолжать строительство», несмотря на опасность обвала стройки, то эта ошибка второго рода может повлечь гибель людей. Можно привести примеры, когда ошибка первого рода влечет более тяжелые последствия, чем ошибка второго рода.

З а м е ч а н и е 1. Правильное решение может быть принято также в двух случаях:

1) гипотеза принимается, причем и в действительности она правильная;

2) гипотеза отвергается, причем и в действительности она неверна.

З а м е ч а н и е 2. Вероятность совершить ошибку первого рода принято обозначать через α; ее называют *уровнем значимости*. Наиболее часто уровень значимости принимают равным 0,05 или 0,01. Если, например, принят уровень значимости, равный 0,05, то это означает, что в пяти случаях из ста имеется риск допустить ошибку первого рода (отвергнуть правильную гипотезу).

Статистический критерий проверки нулевой гипотезы. Наблюдаемое значение критерия

Для проверки нулевой гипотезы используют специально подобранную случайную величину, точное или приближенное распределение которой известно. Эту величину обозначают через U или Z, если она распределена нормально, F или v^2 — по закону Фишера—Снедекора, T — по закону Стьюдента, χ^2 — по закону «хи квадрат» и т. д. Поскольку в этом параграфе вид распределения во внимание приниматься не будет, обозначим эту величину в целях общности через K.

Статистическим критерием (или просто *критерием*) называют случайную величину K, которая служит для проверки нулевой гипотезы.

Например, если проверяют гипотезу о равенстве дисперсий двух нормальных генеральных совокупностей, то в качестве критерия K принимают отношение исправленных выборочных дисперсий:

$$F = s_1^2/s_2^2.$$

$$F = s_1^2/s_2^2.$$

Эта величина случайная, потому что в различных опытах дисперсии принимают различные, наперед неизвестные значения, и распределена по закону Фишера—Снедекора.

Для проверки гипотезы по данным выборок вычисляют частные значения входящих в критерий величин и таким образом получают частное (наблюдаемое) значение критерия.

Наблюдаемым значением $K_{\text{набл}}$ называют значение критерия, вычисленное по выборкам. Например, если по двум выборкам найдены исправленные выборочные дисперсии $s_1^2 = 20$ и $s_2^2 = 5$, то наблюдаемое значение критерия F

$$F_{\text{набл}} = s_1^2/s_2^2 = 20/5 = 4.$$

Критическая область. Область принятия гипотезы. Критические точки

После выбора определенного критерия множество всех его возможных значений разбивают на два непересекающихся подмножества: одно из них содержит значения критерия, при которых нулевая гипотеза отвергается, а другая — при которых она принимается.

Критической областью называют совокупность значений критерия, при которых нулевую гипотезу отвергают.

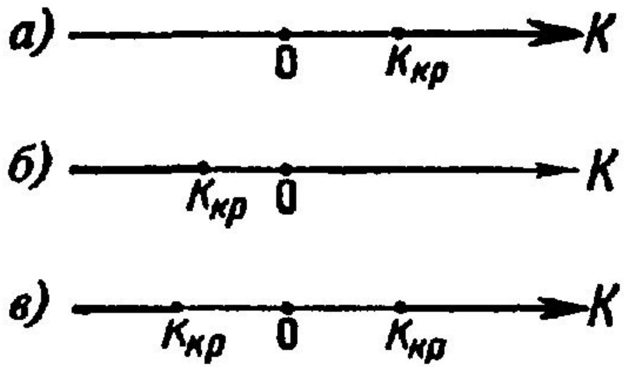

Рис. 23

Областью принятия гипотезы (областью допустимых значений) называют совокупность значений критерия, при которых гипотезу принимают.

Основной принцип проверки статистических гипотез можно сформулировать так: если наблюдаемое значение критерия принадлежит критической области — гипотезу отвергают, если наблюдаемое значение критерия принадлежит области принятия гипотезы — гипотезу принимают.

Поскольку критерий K — одномерная случайная величина, все ее возможные значения принадлежат некоторому интервалу. Поэтому критическая область и область принятия гипотезы также являются интервалами и, следовательно, существуют точки, которые их разделяют.

Критическими точками (границами) $k_{кр}$ называют точки, отделяющие критическую область от области принятия гипотезы.

Различают одностороннюю (правостороннюю или левостороннюю) и двустороннюю критические области.

Правосторонней называют критическую область, определяемую неравенством $K > k_{кр}$, где $k_{кр}$ — положительное число (рис. 23, *а*).

Левосторонней называют критическую область, определяемую неравенством $K < k_{кр}$, где $k_{кр}$ — отрицательное число (рис. 23, *б*).

Односторонней называют правостороннюю или левостороннюю критическую область.

Двусторонней называют критическую область, определяемую неравенствами $K < k_1$, $K > k_2$, где $k_2 > k_1$.

В частности, если критические точки симметричны относительно нуля, двусторонняя критическая область определяется неравенствами (в предположении, что $k_{кр} > 0$): $K < -k_{кр}$, $K > k_{кр}$, или равносильным неравенством $|K| > k_{кр}$ (рис. 23, *в*).

Отыскание правосторонней критической области

Как найти критическую область? Обоснованный ответ на этот вопрос требует привлечения довольно сложной теории. Ограничимся ее элементами. Для определенности начнем с нахождения правосторонней критической области, которая определяется неравенством $K > k_{кр}$, где $k_{кр} > 0$. Видим, что для отыскания правосторонней критической области достаточно найти критическую точку. Следовательно, возникает новый вопрос: как ее найти?

Для ее нахождения задаются достаточной малой вероятностью — уровнем значимости α. Затем ищут критическую точку $k_{кр}$, исходя из требования, чтобы при условии справедливости нулевой гипотезы вероятность того, что критерий K примет значение, большее $k_{кр}$, была равна принятому уровню значимости:

$$P(K > k_{кр}) = \alpha.$$

Для каждого критерия имеются соответствующие таблицы, по которым и находят критическую точку, удовлетворяющую этому требованию.

Отыскание левосторонней и двусторонней критических областей

Отыскание левосторонней и двусторонней критических областей сводится (так же, как и для правосторонней) к нахождению соответствующих критических точек.

Левосторонняя критическая область определяется (см. § 4) неравенством $K < k_{кр}$ $(k_{кр} < 0)$. Критическую точку находят исходя из требования, чтобы при справедливости нулевой гипотезы вероятность того, что критерий примет значение, меньшее $k_{кр}$, была равна принятому уровню значимости:

$$P(K < k_{кр}) = \alpha.$$

Двусторонняя критическая область определяется (см. § 4) неравенствами $K < k_1$, $K > k_2$. Критические точки находят исходя из требования, чтобы при спра-

Отыскание левосторонней и двусторонней критических областей сводится (так же, как и для правосторонней) к нахождению соответствующих критических точек.

Левосторонняя критическая область определяется неравенством $K < k_{кр}$ ($k_{кр} < 0$). Критическую точку находят исходя из требования, чтобы при справедливости нулевой гипотезы вероятность того, что критерий примет значение, меньшее $k_{кр}$, была равна принятому уровню значимости:

$$P(K < k_{кр}) = \alpha.$$

Двусторонняя критическая область определяется (см. § 4) неравенствами $K < k_1$, $K > k_2$. Критические точки находят исходя из требования, чтобы при справедливости нулевой гипотезы сумма вероятностей того, что критерий примет значение, меньшее k_1 или большее k_2, была равна принятому уровню значимости:

$$P(K < k_1) + P(K > k_2) = \alpha. \qquad (*)$$

Ясно, что критические точки могут быть выбраны бесчисленным множеством способов. Если же распределение критерия симметрично относительно нуля и имеются основания (например, для увеличения мощности *)) выбрать симметричные относительно нуля точки $-k_{кр}$ и $k_{кр}$ ($k_{кр} > 0$), то

$$P(K < -k_{кр}) = P(K > k_{кр}).$$

Учитывая (*), получим

$$P(K > k_{кр}) = \alpha/2.$$

Это соотношение и служит для отыскания критических точек двусторонней критической области.

Как уже было указано критические точки находят по соответствующим таблицам.

Дополнительные сведения о выборе критической области. Мощность критерия

Мы строили критическую область, исходя из требования, чтобы вероятность попадания в нее критерия была равна α при условии, что нулевая гипотеза справедлива. Оказывается целесообразным ввести в рассмотрение вероятность попадания критерия в критическую область при условии, что нулевая гипотеза неверна и, следовательно, справедлива конкурирующая.

Мощностью критерия называют вероятность попадания критерия в критическую область при условии, что справедлива конкурирующая гипотеза. Другими словами, мощность критерия есть вероятность того, что нулевая гипотеза будет отвергнута, если верна конкурирующая гипотеза.

Пусть для проверки гипотезы принят определенный уровень значимости и выборка имеет фиксированный объем. Остается произвол в выборе критической области. Покажем, что ее целесообразно построить так, чтобы мощность критерия была максимальной. Предварительно убедимся, что если вероятность ошибки второго рода (принять неправильную гипотезу) равна β, то мощность равна $1-\beta$. Действительно, если β — вероятность ошибки второго рода, т. е. события «принята нулевая гипотеза, причем справедлива конкурирующая», то мощность критерия равна $1-\beta$.

Пусть мощность $1-\beta$ возрастает; следовательно, уменьшается вероятность β совершить ошибку второго рода. Таким образом, чем мощность больше, тем вероятность ошибки второго рода меньше.

Итак, если уровень значимости уже выбран, то *критическую область следует строить так, чтобы мощность критерия была максимальной*. Выполнение этого требования должно обеспечить минимальную ошибку второго рода, что, конечно, желательно.