Проверка статистической гипотезы о математическом ожидании нормального распределения при известной дисперсии.

Пусть имеется нормально распределенная случайная величина ξ,, определенная на множестве объектов некоторой генеральной совокупности. Известно, что $D\xi = \sigma^2$. Математическое ожидание $M\xi$ неизвестно. Допустим, что имеются основания предполагать, что $M\xi = a$, где a – некоторое число (такими основаниями могут быть ограниченные сведения об объектах генеральной совокупности, опыт исследования подобных совокупностей и т. д.). Будем считать также, что имеется другая информация, указывающая на то, что $M\xi = a_1$, где $a_1 > a$.

I. Выдвигаем нулевую гипотезу H_0:$M\xi$=a; при конкурирующей гипотезе H_1:$M\xi$=a_1

Делаем выборку объема n: $x_1, x_2,..., x_n$. В основе проверки лежит тот факт, что случайная величина $\bar{x}$ (выборочная средняя) распределена по нормальному закону с дисперсией σ^2/n и математическим ожиданием, равным a в случае справедливости H_0, и равным a_1 в случае справедливости H_1.

Очевидно, что если величина $\bar{x}$ оказывается достаточно малой, то это дает основание предпочесть гипотезу H_0 гипотезе H_1. При достаточно большом значении $\bar{x}$ более вероятна справедливость гипотезы H_1. Задачу можно было бы поставить так: требуется найти некоторое критическое число, которое разбивало бы все возможные значения выборочной средней (в условиях данной задачи это все действительные числа) на два полубесконечных промежутка. При попадании $\bar{x}$ в левый промежуток следовало бы принимать гипотезу H_0, а при попадании $\bar{x}$ в правый промежуток предпочтение следовало бы оказать гипотезе H_1. Однако на самом деле поступают несколько иначе.

В качестве статистического критерия выбирается случайная величина $z = (\bar{x} - a)\sqrt{n}/\sigma$, распределенная по нормальному закону , причем $Mz = 0$ и $Dz = 1$ (это следует из свойств математического ожидания и дисперсии) в случае справедливости гипотезы H_0. Если справедлива гипотеза H_1, то $Mz = a^* = (a_1 - a)\sqrt{n}/\sigma$, $Dz = 1$.

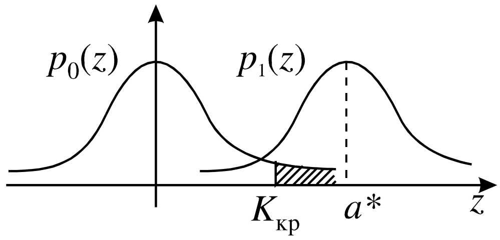

Рис.1.

На рисунке 1. изображены графики $p_0(z)$ и $p_1(z)$ – функций плотности распределения случайной величины z при справедливости гипотез H_0 и H_1, соответственно.

Если величина $\bar{x}$, полученная из выборочных данных, относительно велика, то и величина z велика, что является свидетельством в пользу гипотезы H_1. Относительно малые значения $\bar{x}$ приводят к малым значениям z, что свидетельствует в пользу гипотезы H_0. Отсюда следует, что должна быть выбрана правосторонняя критическая область. По принятому уровню значимости α (например α = 0,05), используя то, что случайная величина z распределена по нормальному закону, определим значение $K_{кр}$ из формулы

$$\alpha = P(K_{кр} < z < \infty) = \Phi(\infty) - \Phi(K_{кр}) = 0{,}5 - \Phi(K_{кр}).$$

Отсюда, $\Phi(K_{кр})=(1-2\alpha)/\alpha$ и осталось воспользоваться таблицей функции Лапласа для нахождения числа $K_{кр}$.

Если величина z, полученная при выборочном значении $\bar{x}$, попадает в область принятия гипотезы ($z < K_{кр}$), то гипотеза H_0 принимается (делается вывод, что выборочные данные не противоречат гипотезе H_0). Если величина z попадает в критическую область, то гипотеза H_0 отвергается.

В данной задаче может быть подсчитана мощность критерия:

$$1-\beta = \Phi(\infty) - \Phi(K_{кр} - (a_1 - a)\sqrt{n}/\sigma)$$

Мощность критерия тем больше, чем больше разность $a_1 - a$.

II. Если в задаче поставить другое условие: H_0: $M\xi = a$; H_1: $M\xi = a_1$, $a_1 < a$,

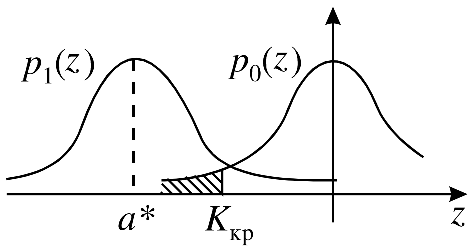

Рис. 2.

то сохранив смысл всех рассуждений, здесь придется рассматривать левостороннюю критическую область, как изображено на рисунке 2. Здесь, как и в предыдущем случае, $a^*=(a_1-a)\sqrt{n}/\sigma$, а величина $K_{кр}$ определяется из формулы

$$\alpha = P(-\infty < z < K_{кр}) = \Phi(K_{кр}) - \Phi(-\infty) = \Phi(K_{кр}) + 1/2.$$

Используя формулу $-\Phi(K_{кр}) = \Phi(-K_{кр})$, получаем: $\Phi(-K_{кр}) = (1-2\alpha)/2$.

Отметим, что по смыслу задачи здесь $K_{кр}$ – отрицательное число.

Значения z, вычисленные по выборочным данным, превышающие $K_{кр}$, согласуются с гипотезой H_0. Если величина z попадает в критическую область ($z<K_{кр}$), то гипотезу H_0 следует отвергнуть, считая предпочтительной гипотезу H_1.

III. Рассмотрим теперь такую задачу: H_0: $M\xi = a$; H_1: $M\xi \neq a$.

В данном случае большие отклонения величины z от нуля в положительную или отрицательную сторону должны приводить к заключению о ложности гипотезы H_0, то есть здесь следует рассматривать двустороннюю

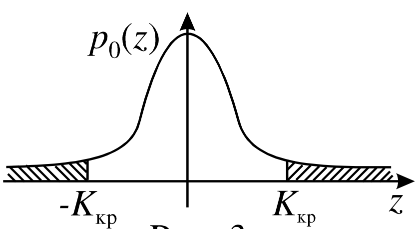

Рис. 3.

критическую область, как изображено на рисунке 3. Критическое значение $K_{кр}$

определяется с помощью соотношения $P(-K_{кр}<z<K_{кр}) = 1-\alpha = \Phi(K_{кр})-\Phi(-K_{кр}) = 2\Phi(K_{кр})$. Из этого соотношения следует: $\Phi(K_{кр}) = (1-\alpha)/2$.

Проверка гипотезы о равенстве дисперсий.

Гипотезы о дисперсии играют очень важную роль в экономико–математическом моделировании, так как величина рассеяния экспериментальных выборочных данных относительно рассчитанных теоретических значений соответствующих параметров, характеризующаяся дисперсией, дает возможность судить о пригодности (адекватности) теории или модели, на основании которой строится теория.

Пусть нормально распределенная случайная величина ξ определена на некотором множестве, образующем генеральную совокупность, а нормально распределенная случайная величина η определена на другом множестве, которое тоже составляет генеральную совокупность. Из обеих совокупностей делаются выборки: из первой – объема n_1, а из второй – объема n_2 (отметим, что объем выборки не всегда можно определить заранее, как например в случае, если он равен количеству рыб, попавших в сеть). По каждой выборке рассчитывается исправленная выборочная дисперсия: s_1^2 для выборки из первой совокупности и s_2^2 для выборки из второй совокупности.

Поставим задачу: с помощью выборочных данных проверить статистическую гипотезу H_0: $D\xi = D\eta$. В качестве конкурирующей гипотезы будем рассматривать идею, заключающуюся в том, что дисперсия той совокупности, для которой исправленная выборочная дисперсия оказалась наибольшей, больше дисперсии другой совокупности. Критерий берется в следующем виде: $F = S^{**}/S^{*}$. Здесь S^{**}– наибольшая из двух оценок s_1^2 и s_2^2, а S^{*}– наименьшая из тех же двух оценок. Критерий F распределен по закону Фишера с k_1 и k_2 степенями свободы. Здесь $k_1 = n_1-1$, $k_2 = n_2-1$, если $S^{**}= s_1^2$; $k_1 = n_2-1$, $k_2 = n_1-1$, если $S^{**}= s_2^2$.

В этой задаче естественно рассматривать правостороннюю критическую область, так как достаточно большие выборочные значения критерия F свидетельствуют в пользу конкурирующей гипотезы.

При заданном уровне значимости q (обычно q =0,05 или q =0,01) критическое значение $F_{кр}$ определяется из таблицы распределения Фишера. В случае $F > F_{кр}$ гипотеза H_0 отвергается, а в случае $F < F_{кр}$ – принимается.

Пусть два множества некоторых объектов, обладающих количественным признаком, подвергнуты выборочному контролю. Значения количественного признака есть распределенные по нормальному закону случайные величины, которые мы обозначим ξ_1 и ξ_2, соответственно, для первого и для второго множеств.

Из первого множества сделана выборка объема n_1=21 и подсчитана исправленная выборочная дисперсия, оказавшаяся равной 0,75. Из второго множества сделана выборка объема n_2=11. Эта выборка дала значение исправленной выборочной дисперсии, равное 0,25. Выдвигаем гипотезу H_0: $D\xi_1=D\xi_2$. Конкурирующая гипотеза H_1 заключается в том, что $D\xi_1>D\xi_2$. В данном случае выборочное значение $F_в$ критерия Фишера равно 3. При выбранном уровне значимости q = 0,05 по числам степеней свободы k_1=20, k_2=10 находим по таблице распределения Фишера $F_{кр}$=2,77. Так как $F_в > F_{кр}$, гипотеза о равенстве дисперсий должна быть отвергнута.

Проверка статистической значимости выборочного коэффициента корреляции.

Проверкой статистической значимости выборочной оценки δ **параметра** Δ **генеральной совокупности называется проверка статистической гипотезы** H_0: $\Delta = 0$, **при конкурирующей гипотезе** H_1: $\Delta \neq 0$. **Если гипотеза** H_0 **отвергается, то оценка** δ **считается статистически значимой**.

Пусть имеются две случайные величины ξ и η, определенные на множестве объектов одной и той же генеральной совокупности, причем обе имеют нормальное распределение. Задача заключается в проверке статистической гипотезы об отсутствии корреляционной зависимости между случайными величинами ξ и η. H_0: $\rho_{\xi\eta} = 0$; H_1: $\rho_{\xi\eta} \neq 0$. Здесь $\rho_{\xi\eta}$ – коэффициент линейной корреляции.

Производится выборка объема n и вычисляется выборочный коэффициент корреляции r. За статистический критерий принимается случайная величина

$$t = r\sqrt{n-2}\,/\sqrt{1-r^2}\ ,$$

которая распределена по закону Стьюдента с n – 2 степенями свободы.

Отметим сначала, что все возможные значения выборочного коэффициента корреляции r лежат в промежутке [–1;1]. Очевидно, что относительно большие отклонения в любую сторону значений t от нуля получаются при относительно больших, то есть близких к 1, значениях модуля r. Близкие к 1 значения модуля r противоречат гипотезе H_0, поэтому здесь естественно рассматривать двустороннюю критическую область для критерия t.

По уровню значимости α и по числу степеней свободы n – 2 находим из таблицы распределения Стьюдента значение $t_{кр}$. Если модуль выборочного значения критерия $t_в$ превосходит $t_{кр}$, то гипотеза H_0 отвергается и выборочный коэффициент корреляции считается статистически значимым. В противном случае, то есть если $|t_в| < t_{кр}$ и принимается гипотеза H_0, выборочный коэффициент корреляции считается статистически незначимым.