Критерии для проверки гипотез о параметрах многомерной нормальной случайной величины на основе отношения правдоподобия

Пусть гипотеза H_0 отличается от гипотезы H_1 тем, что она накладывает дополнительно r ограничений на параметры распределения. Пусть $\Psi_0(X,\theta_0)$ - функция правдоподобия при истинности H_0, а $\Psi_1(X,\theta_1)$ - функция правдоподобия при истинности H_1, где θ_0, θ_1 - векторы неизвестных параметров распределения при истинности H_0 и H_1 соответственно. Найдем оценку максимального правдоподобия $\hat{\theta}_0$ для θ_0 в предположении истинности H_0, а также оценку максимального правдоподобия $\hat{\theta}_1$ для θ_1 в предположении истинности H_1. Заменив неизвестные параметры в функциях правдоподобия их оценками максимального правдоподобия, получим соответственно функции $\hat{\Psi}_0 = \Psi_0(X,\hat{\theta}_0) = \max \Psi_0(X,\theta_0)$ и $\hat{\Psi}_1 = \Psi_1(X,\hat{\theta}_1) = \max \Psi_1(X,\theta_1)$. Запишем отношение правдоподобия:

$$\lambda = \frac{\hat{\Psi}_0}{\hat{\Psi}_1} = \frac{\Psi_0(X,\hat{\theta}_0)}{\Psi_1(X,\hat{\theta}_1)}.$$

При условиях регулярности исходного распределения, статистика $\eta = -2\ln\lambda$ при истинности H_0 асимптотически имеет распределение χ^2 с r степенями свободы. При малых n отклонение распределения статистики η от распределения χ^2 может быть существенным, однако можно улучшить аппроксимацию, если рассматривать статистику $\eta = -2\rho\ln\lambda$, где константу ρ подобрать так, чтобы $M(\eta) = r$ с точностью до бесконечно малых заданного порядка малости относительно величины $1/n$. Как правило, при определении поправочного множителя ρ, в отношении правдоподобия λ выборочные матрицы ковариаций заменяют несмещенными выборочными матрицами, а вместо объемов выборок используют числа степеней свободы.

Гипотеза о равенстве матриц ковариаций

Пусть имеются выборки $X_1, X_2, \ldots, X_q$ объемов n_1, n_2, $\ldots, n_q$ соответственно из k-мерных нормальных совокупностей $N(\vec{m}_1, A_1)$, $N(\vec{m}_2, A_2)$, …, $N(\vec{m}_q, A_q)$. Проверяется гипотеза $H_0 : A_1 = A_2 = \ldots = A_q$.

Отношение правдоподобия в этом случае будет равно:

$$\lambda = \frac{\prod_{i=1}^{q} |S_i|^{\frac{n_i-1}{2}}}{|S|^{\frac{n-q}{2}}},$$

где $S=\frac{1}{n-q}\sum_{i=1}^{q}(n_i-1)S_i=\frac{1}{n-q}\sum_{i=1}^{q}\widehat{X}_i^T\widehat{X}_i$, $n=\sum_{i=1}^{q}n_i$, $S_i=\frac{1}{n_i-1}\widehat{X}_i^T\widehat{X}_i$, $\widehat{X}_i$ - центрированная матрица i-ой выборки.

Если положить: $\rho=1-\left(\sum_{i=1}^{q}\left(\frac{1}{n_i-1}\right)-\frac{1}{n-q}\right)\frac{2k^2+3k-1}{6(k+1)(q-1)}$, то статистика:

$$\eta=-2\rho\ln\lambda=-\rho\left(\sum_{i=1}^{q}\left((n_i-1)\ln|S_i|\right)-(n-q)\ln|S|\right)$$

будет асимптотически иметь распределение χ^2 с $\nu=\frac{1}{2}(q-1)k(k+1)$ степенями свободы.

Таким образом, критерий будет иметь вид:

$$\delta(X,Y)=\begin{cases}H_0, & -2\rho\ln\lambda<\tau_{1-\alpha}\\ H_1, & -2\rho\ln\lambda\geq\tau_{1-\alpha}\end{cases},$$

где $\tau_{1-\alpha}$ - квантиль распределения χ^2 с $\nu=\frac{1}{2}k(k+1)$ степенями свободы.

Гипотеза о равенстве векторов средних

Пусть имеются выборки $X_1,X_2,\ldots,X_q$ объемов n_1, n_2, $\ldots,n_q$ соответственно из k-мерных нормальных совокупностей $N(\vec{m}_1,A)$, $N(\vec{m}_2,A)$, $\ldots$, $N(\vec{m}_q,A)$. Проверяется гипотеза $H_0: m_1=m_2=\ldots=m_q$.

Отношение правдоподобия:

$$\lambda=\frac{|S|^{\frac{n-q}{2}}}{|S_0|^{\frac{n-q}{2}}},$$

где: $S=\frac{1}{n-q}\sum_{i=1}^{q}(n_i-1)S_i$ и $S_0=\frac{1}{n-q}\widehat{X}^T\widehat{X}=S+\frac{1}{n-q}\sum_{i=1}^{q}n_i(\overline{X}_i-\overline{X})(\overline{X}_i-\overline{X})^T$, $n=\sum_{i=1}^{q}n_i$, $\widehat{X}$ - центрированная матрица объединенной выборки, $\overline{X}$ - вектор выборочных средних объединенной выборки, $\overline{X}_i$ - вектор выборочных средних i-ой выборки.

Если положить $\rho=1-\frac{k-q+2}{2(n-q)}$, то статистика $\eta=-2\rho\ln\lambda$ будет приближенно иметь распределение χ^2 с $\nu=qk-k=k(q-1)$ степенями свободы.

Гипотеза об однородности многомерных нормальных совокупностей.

Пусть имеются выборки $X_1,X_2,\ldots,X_q$ объемов n_1, n_2, $\ldots,n_q$ соответственно из k-мерных нормальных совокупностей $N(\vec{m}_1,A_1)$, $N(\vec{m}_2,A_2)$, $\ldots$, $N(\vec{m}_q,A_q)$. Проверяется гипотеза $H_0:\vec{m}_1=\vec{m}_2=\ldots\vec{m}_q,\quad A_1=A_2=\ldots=A_q$.

Отношение правдоподобия для данной гипотезы:

$$\lambda = \frac{\prod_{i=1}^{q} |S_i|^{\frac{n_i - 1}{2}}}{|S_0|^{\frac{n-q}{2}}},$$

где $S_0 = \frac{1}{n-q}\widehat{X}^T \widehat{X} = S + \frac{1}{n-q}\sum_{i=1}^{q} n_i (\overline{X}_i - \overline{X})(\overline{X}_i - \overline{X})^T$, $S = \frac{1}{n-q}\sum_{i=1}^{q}(n_i - 1)S_i = \frac{1}{n-q}\sum_{i=1}^{q}\widehat{X}_i^T \widehat{X}_i$, $n = \sum_{i=1}^{q} n_i$, $\widehat{X}$ - центрированная матрица объединенной выборки, $\overline{X}_i$ - вектор выборочных средних i-ой выборки, $\overline{X}$ - вектор выборочных средних объединенной выборки.

Если положить:

$\rho = 1 - \left(\sum_{i=1}^{q}\left(\frac{1}{n_i - 1}\right) - \frac{1}{n-q}\right)\frac{2k^2 + 3k - 1}{6(k+3)(q-1)} - \frac{1}{n-q}\frac{k-q+2}{k+3}$, то статистика $\eta = -2\rho \ln \lambda$ будет приближенно иметь распределение χ^2 с числом степеней свободы $\nu = \frac{1}{2}(q-1)k(k+3)$.

Гипотеза о независимости множеств случайных величин

Пусть k-мерный нормальный вектор $\vec{\xi}$ разбит на q подвекторов $\vec{\xi}_1, \vec{\xi}_2, \ldots, \vec{\xi}_q$ размерности $k_1, k_2, \ldots, k_q$ соответственно. Требуется по выборке X из n значений вектора $\vec{\xi}$ проверить гипотезу H_0: $\vec{\xi}_1, \vec{\xi}_2, \ldots, \vec{\xi}_q$ взаимно независимы, то есть, $f_{\vec{\xi}} = f_{\vec{\xi}_1} \cdot f_{\vec{\xi}_2} \cdot \ldots \cdot f_{\vec{\xi}_q}$. Гипотезу H_0 можно также сформулировать как гипотезу, что матрица ковариаций имеет блочно диагональный вид:

$$A = \begin{pmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & A_{22} & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & A_{qq} \end{pmatrix},$$

где A_{ii} матрица ковариаций подвектора $\vec{\xi}_i$.

Для данной гипотезы отношение правдоподобия:

$$\lambda = \frac{|\overline{A}|^{\frac{n}{2}}}{\prod_{i=1}^{q}|\overline{A}_{ii}|^{\frac{n}{2}}} = \frac{|S|^{\frac{n}{2}}}{\prod_{i=1}^{q}|S_{ii}|^{\frac{n}{2}}},$$

где $\overline{A}_{ii}, S_{ii}$ - выборочные матрицы ковариаций i-го подвектора.

Если положить:

$$\rho = 1 - \frac{2(k^3 - \sum_{i=1}^{q} k_i^3) + 9(k^2 - \sum_{i=1}^{q} k_i^2)}{6n(k^2 - \sum_{i=1}^{q} k_i^2)},$$

то статистика $\eta = -2\rho \ln \lambda$ будет асимптотически иметь распределение χ^2 с $\nu = \frac{1}{2}(k^2 - \sum_{i=1}^{q} k_i^2)$ степенями свободы.

В случае $q = k$ получаем критерий для проверки гипотезы о независимости компонент вектора $\vec{\xi}$. В этом случае:

$$\lambda = \frac{|\overline{A}|^{\frac{n}{2}}}{\prod_{i=1}^{q} (\overline{D}_i)^{\frac{n}{2}}} = \frac{|S|^{\frac{n}{2}}}{\prod_{i=1}^{k} (s_i^2)^{\frac{n}{2}}}, \qquad \rho = 1 - \frac{2k+11}{6n}, \qquad \nu = \frac{1}{2}k(k-1).$$

Гипотеза о сферичности распределения

Пусть требуется по выборке X из n значений вектора $\vec{\xi}$ проверить гипотезу H_0: $A = \sigma^2 E$, где величина σ^2 не задана. Гипотеза подобного рода называется гипотезой о сферичности распределения.

Для данной гипотезы отношение правдоподобия:

$$\lambda = \frac{|S|^{\frac{n-1}{2}}}{\left(Sp(S/k)\right)^{\frac{k(n-1)}{2}}}.$$

Если: $\rho = 1 - \frac{2k^2 + k + 2}{6k(n-1)}$,

то статистика $\eta = -2\rho \ln \lambda$ будет асимптотически иметь распределение χ^2 с $\nu = \frac{1}{2}k(k+1) - 1$ степенями свободы.

Гипотеза о равенстве матрицы ковариаций заданной матрице.

Пусть $X = \{(X_1^{(1)}, X_2^{(1)}, \ldots, X_k^{(1)}), (X_1^{(2)}, X_2^{(2)}, \ldots, X_k^{(2)}), \ldots, (X_1^{(n)}, X_2^{(n)}, \ldots, X_k^{(n)})\}$ - выборка объема n из k-мерной нормальной совокупности $N(\vec{m}, A)$ случайной величины $\vec{\xi}$. Проверяется гипотеза $H_0 : A = A_0$, против альтернативы $H_1 : A \neq A_0$.

Отношение правдоподобия для данной гипотезы:

$$\lambda = \left|\overline{A} \cdot A_0^{-1}\right|^{\frac{1}{2}n} \exp\left[\frac{1}{2}kn - \frac{1}{2}nSp(\overline{A} \cdot A_0^{-1})\right].$$

Положим: $\eta = -2\ln \lambda$, тогда при истинности H_0, статистика η будет асимптотически иметь распределение χ^2 с $\nu = \frac{1}{2}k(k+1)$ степенями свободы.

Гипотеза о равенстве вектора средних заданному вектору и матрицы ковариаций заданной матрице.

Пусть $X=\{(X_1^{(1)},X_2^{(1)},\ldots,X_k^{(1)}),(X_1^{(2)},X_2^{(2)},\ldots,X_k^{(2)}),\ldots,(X_1^{(n)},X_2^{(n)},\ldots,X_k^{(n)})\}$ - выборка объема n из k-мерной нормальной совокупности $N(\vec{m},A)$ случайной величины $\vec{\xi}$. Проверяется гипотеза $H_0:\vec{m}=\vec{m}_0,\, A=A_0$.

Отношение правдоподобия для H_0:

$$\lambda=\left|\overline{A}\cdot A_0^{-1}\right|^{\frac{1}{2}n}\exp\left[\frac{1}{2}n(k-Sp(\overline{A}\cdot A_0^{-1})-(\overline{X}-\vec{m}_0)^T A_0^{-1}(\overline{X}-\vec{m}_0))\right].$$

Положим: $\eta=-2\ln\lambda$, тогда при истинности H_0, статистика η будет асимптотически иметь распределение χ^2 с $\nu=\dfrac{1}{2}k(k+1)+k$ степенями свободы.