Лекция 7

ОСНОВНЫЕ ПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ, ПРИМЕНЯЕМЫЕ ПРИ АНАЛИЗЕ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЯ ДИАГНОСТИЧЕСКОЙ ИНФОРМАЦИИ

Параметрические методы и связанные с ними статистические критерии предполагают известным вид функции распределения генеральной совокупности, и проверка гипотез сводится определению неизвестных значений параметров распределений.

Как уже указывалось, сравнение эффективности критериев может быть проведено на основе применения функции качества (мощности) критерия, под которой понимают вероятность π отклонения нулевой гипотезы в случае ее ложности. В соответствии со сказанным, при равных значениях α функция качества для непараметрического критерия меньше, чем для параметрического. Это значит, что различение выборок с помощью непараметрических критериев осуществляется хуже, чем с помощью параметрических. Непараметрический критерий может не выявить различия, которое обнаруживается параметрическим. Но если различие выявлено непараметрическим критерием, оно будет подтверждено параметрическим.

Для двух сравниваемых критериев существует также понятие эффективности критерия, под которым понимают соотношение объемов выборок, необходимых для одинакового качества контроля. Отсюда следует возможность обеспечения одинакового качества контроля при помощи параметрического и непараметрического критериев за счет увеличения объема выборки в последнем случае. Соответственно, параметрический критерий позволяет получить одинаковую с непараметрическим критерием надежность результатов при меньшем объеме выборки.

Для быстрого *предварительного анализа* (просмотра) данных может быть полезным применение ряда простых критериев, которые широко применяются на практике.

1. Критерий для определения соответствия данных нормальному закону распределения

Остановимся на некоторых часто применяемых параметрических критериях, позволяющих получить результаты с помощью простейших вычислений. При этом предполагается соответствие данных нормальному закону распределения. Поэтому, естественно, в первую очередь подлежит проверке именно это предположение. Для такой проверки представляется целесообразной следующая последовательность действий:

— вычисляют выборочные характеристики: среднее $\overline{x}$, среднеквадратическое отклонение s, третий m_3 и четвертый m_4 центральные моменты, коэффициент вариации γ;

— вычисляют выборочные асимметрию $S_{k\ выб}$ и эксцесс $E_{k\ выб}$:

$$S_{k\ выб} = m_3 / s^3 = m_3 / m_2^{3/2}, \qquad E_{k\ выб} = m_4 / s^4 - 3 = m_4 / m_2^2 - 3;$$

— находят несмещенные оценки асимметрии и эксцесса:

$$S_{k0} = S_{k\ выб} \cdot \left[N(N-1) \right]^{1/2} / (N-2),$$

$$E_{k0} = (N-1)\left[(N+1) E_{k\ выб} + 6 \right] / \left[(N-2)(N-3) \right];$$

— определяют теоретические значения среднеквадратического отклонения асимметрии и эксцесса для данного размера выборки:

$$\sigma_S = \left\{ 6N(N-1) / \left[(N-2)(N+1)(N+3) \right] \right\}^{1/2},$$

$$\sigma_E = \left\{ 24N(N-1)^2 / \left[(N-3)(N-2)(N+3)(N+5) \right] \right\}^{1/2};$$

— проверяют условия: $|S_{k0}| \le 3\sigma_S$, $|E_{k0}| \le 5\sigma_E$. Если оба условия соблюдены, может быть принята гипотеза о нормальном распределении величин в выборке.

Из свойств нормального распределения следует достаточно простая связь между размахом выборки $R = x_{\max} - x_{\min}$ и величиной среднеквадратического отклонения, характеризуемая при небольших размерах выборки, табл. Л7.1.

Таблица Л7.1. **Связь между оценкой среднеквадратического отклонения *s* и величиной размаха *R* при разных размерах выборки *N***

N	2	3	4	5	6	7	8	9	10	11	12
$K = s/R$	0,88	0,59	0,50	0,43	0,40	0,37	0,35	0,33	0,32	0,31	0,31

2. Критерий для отбрасывания резко выделяющихся результатов измерений

Измерения информативных параметров проводятся, как правило, в условиях воздействия различного рода помех, причем при измерениях не исключены грубые ошибки или сбои, приводящие в ряде случаев к результатам измерений, существенно отличающимся от остальных и поэтому вызывающие сомнение. В этих случаях сомнительные данные исключают путем применения специальных критериев.

Результаты измерений представляют в виде *вариационного ряда* — ряда, в котором результаты измерений расположены в порядке возрастания $x_1; x_2; ....; x_n$, где x_1 и x_n — соответственно наименьшее и наибольшее значения при измерениях.

Нулевой гипотезой H_0 при использовании критериев является предположение о том, что наибольшее значение x_n (или наименьшее значение x_1) принадлежат той же генеральной совокупности, что и все остальные $n-1$ измерений.

2.1. Предположим, что известны закон распределения измеряемого параметра — нормальный и генеральная дисперсия σ^2. Неизвестно генеральное среднее μ. Процедура проверки H_0 включает следующие операции:

— рассчитывают выборочное среднее $\bar{x}$ по известной формуле;

— вычисляют статистику $t_n = (x_n - \bar{x})/\sigma$, если вызывает сомнение значение x_n или $t_1 = \left|(x_1 - \bar{x})/\sigma\right|$, если вызывает сомнение значение x_1;

— для выбранного уровня значимости α и числа степеней свободы k (в нашем случае $k = n-1$, так как был рассчитан один параметр $\bar{x}$) с помощью статистических таблиц находят квантиль распределения $t_\alpha(k)$;

— если $t_n \le t_\alpha(k)$ или $t_1 \le t_\alpha(k)$, то H_0 о принадлежности измерения x_n (или x_1) генеральной совокупности принимается, в ином случае H_0 отвергается и x_n (или x_1) исключается из ряда измерений.

2.2. Предположим, что измеряемый параметр распределен по нормальному закону, но генеральное среднее μ и генеральная дисперсия σ^2 неизвестны.

Процедура проверки H_0 включает следующие операции (*критерий К.В. Смирнова*):

— рассчитывают выборочные среднее $\bar{x}$ и дисперсию s^2 по известным формулам;

— вычисляют статистику $u_n = (x_n - \bar{x})/s$, если вызывает сомнение значение x_n или $u_1 = \left|(x_1 - \bar{x})/s\right|$, если вызывает сомнение значение x_1;

— сравнивают t_n (или t_1) с критическим значением $u_\alpha(k)$ для заданного уровня значимости α и числа степеней свободы k. В этом случае $k = n-2$, так как на основе выборки были рассчитаны два параметра - среднее $\overline{x}$ и дисперсия s^2.

Заметим, если $k > 25$, то можно считать $u_\alpha(k) = t_\alpha(k)$.

— если $u_n \le u_\alpha(k)$ или $u_1 \le u_\alpha(k)$, то H_0 о принадлежности измерения x_n (или x_1) генеральной совокупности принимается, в ином случае H_0 отвергается и x_n (или x_1) исключается из ряда измерений.

При отбрасывании величин x_n (или x_1) оценки выборочных среднего $\overline{x}$ и дисперсии s пересчитываются, и проверяются на соответствие выборки оставшиеся крайние значения вариационного ряда. При этом при определении числа степеней свободы k не следует забывать, что объем анализируемой выборки уменьшился за счет отброшенных резко выделяющихся величин.

Проверим наличие грубых ошибок в обеих выборках из примера Л5.2 (лекция 5). Для первой выборки (с покрытием) имеем $d_{\max} = 62$, $d_{\min} = 27$ и выборочные среднее значение и среднеквадратическое отклонение соответственно равны $\overline{d} = 45,75$ и $s = 9,53$. Следовательно, $u_{\max} = \left(d_{\max} - \overline{d}\right)/s = 16,25/9,53 = 1,71$ и $u_{\min} = \left|\left(d_{\min} - \overline{d}\right)/s\right| = 1,97$. Критическое значение u-статистики выражается через квантили распределения Стьюдента $t_{\alpha,\,k}$ и для $\alpha = 0,1$ и с числом степеней свободы $k = n - 2 = 10$ $\left(t_{0,1;\,10} = 1,81\right)$ $u_{0,1}(10)$ составляет

$$u_\alpha(k) = t_{\alpha,\,k}\sqrt{n-1}\Big/\sqrt{k + \left(t_{\alpha,\,k}\right)^2} = 1,81 \cdot 3,32\Big/\sqrt{10 + 1,81^2} = 1,65 .$$

Как видно, критическое значение превышено, следовательно, результат с максимальным отклонением от среднего является грубой ошибкой и выпадает из общего ряда. Прежде всего, следует исключить из выборки значение, равное 27, так как u-статистика, для этого значения в наибольшей степени отличается от критического значения $u_{0,1}(10)$. Необходимо проанализировать, почему это произошло при проведении измерений. Повторение вычислений без этого значения дает новые значения статистик: $\overline{d} = 47,45$; $s = 7,84$; $u_{\max} = \left(d_{\max} - \overline{d}\right)/s = 14,55/7,84 = 1,85$ и

$$u_{0,1}(9) = t_{0,1;\,9}\sqrt{10}\Big/\sqrt{9 + \left(t_{0,1;\,9}\right)^2} = 1,83 \cdot 3,16\Big/\sqrt{9 + 1,83^2} = 1,65 .$$

Вновь критическое значение превышено, и теперь отсеиванию подлежит результат, равный 62. Без этого результата будем иметь $\overline{d} = 46$; $s = 6,51$. Так как теперь $d_{\max} = 59$, $u_{\max} = \left(d_{\max} - \overline{d}\right) \big/ s = 13/6,51 = 2$ и $t_{0,1;\,8} = 1,52$. В итоге получаем еще один выпадающий результат, равный 59. Отбросив его, приходим, наконец, к выборке, не содержащей грубых промахов: 39; 43; 43; 52; 52; 40; 45; 47; 40, для которой нулевая гипотеза об отсутствии грубых промахов может быть принята.

Задача Л7.1. Провести аналогичные вычисления для второй выборки (с покрытием).

Для вновь образованной выборки найдем вариацию: $v = s/x = 4,93 / 44,55 = 0,11 < 0,33$, то есть целесообразна проверка на соответствие нормальному распределению. Заметим, что, как показывают расчеты, гипотезе о нормальности распределения не противоречит даже исходный набор данных (выборка), что является естественным: при столь малом объеме выборки должны иметься веские основания для отклонения нулевой гипотезы.

3. Критерий равенства дисперсий двух выборок

Существенное значение для статистического анализа данных имеет выборочное распределение отношения двух дисперсий.

Пусть две независимые выборки размером n_1 и n_2 подчинены нормальным распределениям $N\left(\mu_1, \sigma_1^2\right)$ и $N\left(\mu_2, \sigma_2^2\right)$. Обозначим через $k_1 = n_1 - 1$ и $k_2 = n_2 - 1$ число степеней свободы для первой и второй выборок соответственно. Проверяется гипотеза H_0 о том, что указанные выборки принадлежат генеральным совокупностям с равными дисперсиями $\sigma_1^2 = \sigma_2^2 = \sigma^2$, при альтернативной гипотезе H_1 о том, что $\sigma_1^2 \neq \sigma_2^2$.

Для проверки гипотезы H_0 используется двусторонний F-критерий (критерий Фишера). Методика проверки сводится к следующему:

— рассчитываются оценочные дисперсии s_1^2 и s_2^2 выборок по известным формулам. Для определенности будем считать $s_1^2 > s_2^2$, тогда статистика $k_1 s_1^2 \big/ \sigma_1^2$ распределена по закону χ_{k1}^2, статистика $k_2 s_2^2 \big/ \sigma_2^2$ — по закону χ_{k2}^2, а отношение

$$F\left(k_1, k_2\right) = s_1^2 \sigma_2^2 \big/ s_2^2 \sigma_1^2 = \left(k_2 / k_1\right)\left(\chi_{k1}^2 \big/ \chi_{k2}^2\right),$$

называемое *отношением Фишера*, распределено по закону Фишера-Снедекора со степенями свободы $\left(k_1; k_2\right)$, плотность вероятности которого имеет вид

$$w_{k1,k2}(x)=\left(\frac{k_1}{k_2}\right)^{k_1/2}\frac{\Gamma\left[\left(k_1+k_2\right)/2\right]}{\Gamma\left(k_1/2\right)\Gamma\left(k_2/2\right)}\times\frac{x^{\left(k_1/2-1\right)}}{\left(1+k_1x/k_2\right)^{\left(k_1+k_2\right)/2}}\ ,\quad x>0,$$

причем, что в случае справедливости H_0 статистика $F\left(k_1,k_2\right)=s_1^2/s_2^2$;

— для заданного уровня значимости α, k_1 и k_2 по таблицам находят критическое значение $F_{1-\alpha/2}\left(k_1,k_2\right)$;

— если $F\left(k_1,k_2\right)=s_1^2/s_2^2\le F_{1-\alpha/2}\left(k_1,k_2\right)$, то гипотеза H_0 о равенстве $\sigma_1^2=\sigma_2^2=\sigma^2$ принимается.

В этом случае $\left(\sigma_1^2=\sigma_2^2=\sigma^2\right)$ по двум выборочным дисперсиям уточняют оценку генеральной дисперсии σ^2 по формуле

$$s^2=\frac{\left(n_1-1\right)s_1^2+\left(n_2-1\right)s_2^2}{n_1+n_2-2}.$$

Отношение Фишера эффективно используют для оценки качества аппроксимации опытных данных той или иной математической моделью. Если согласно отношению Фишера рассеяние экспериментальных данных относительно одной аппроксимирующей кривой меньше, чем относительно другой, то первой следует отдать предпочтение.

Если различие незначимо, предпочтение не может быть отдано той или другой формуле. В частности, степень аппроксимирующего полинома целесообразно повышать только до тех пор, пока дисперсия значимо убывает. Следует иметь в виду, что наилучший аппроксимирующий полином может не содержать некоторых степеней, поэтому необходимо продолжить анализ еще на несколько шагов после достижения ситуации, когда повышение степени полинома не приводит к значимому уменьшению дисперсии. Необходимо помнить, что по мере повышения степени полинома, число степеней свободы убывает: для вычисления коэффициента полинома нулевой степени, то есть среднего значения, использовано одно уравнение, и число степеней свободы уменьшилось на единицу. После вычисления коэффициентов полинома первой степени число степеней свободы уменьшается на два и т. д.

4. Критерий равенства средних двух выборок

Пусть из двух нормально распределенных генеральных совокупностей с неизвестными параметрами $\left(\mu_1, \sigma_1^2\right)$ и $\left(\mu_2, \sigma_2^2\right)$ получены выборки объемом соответственно n_1 и n_2. По результатам измерений оценены параметры распределений $\left(\overline{x}_1, s_1^2\right)$ и $\left(\overline{x}_2, s_2^2\right)$. Необходимо проверить гипотезу H_0 о равенстве средних значений $\mu = \mu_1 = \mu_2$ при альтернативной гипотезе $\mu_1 \neq \mu_2$.

4.1. Рассмотрим случай $\sigma_1^2 = \sigma_2^2 = \sigma^2$, последнее можно проверить по методике, приведенной в п.2.

— по двум выборочным дисперсиям уточняют оценку генеральной дисперсии σ^2 по формуле

$$s^2 = \frac{\left(n_1 - 1\right)s_1^2 + \left(n_2 - 1\right)s_2^2}{n_1 + n_2 - 2};$$

— вычисляется статистика

$$t = \frac{\overline{x}_1 - \overline{x}_2}{s\sqrt{1/n_1 + 1/n_2}};$$

— для заданного уровня значимости α и числа степеней свободы $k = n_1 + n_2 - 2$, по статистическим таблицам находят критическое значение $t_{\alpha,k}$;

— если $|t| \leq t_{\alpha,k}$, то гипотеза H_0 о равенстве генеральных средних значений $\mu_1 = \mu_2$ принимается, в ином случае принимается альтернативная гипотеза.

В случае принятия гипотезы H_0 по двум выборочным средним уточняют оценку генерального среднего значения

$$\overline{x} = \frac{n_1\overline{x}_1 + n_2\overline{x}_2}{n_1 + n_2},$$

которую используют для построения доверительных интервалов средних значений и среднеквадратических отклонений.

Для разности средних значений двух выборок с размерами n_1 и n_2, если их дисперсии равны, доверительный интервал имеет границы

$$\left(x_1 - x_2 \pm t_{\alpha,k} \sqrt{\frac{n_1 + n_2}{n_1 n_2 \left(n_1 + n_2 - 2\right)} \left(n_2 s_1^2 + n_1 s_2^2\right)} \right), \tag{Л9.1}$$

где $k = n_1 + n_2 - 2$, как и выше, число степеней свободы.

4.2. Рассмотрим случай различия генеральных дисперсий $\sigma_1^2 \neq \sigma_2^2$. В этом случае

— вычисляется статистика

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_1^2 / n_1 + s_2^2 / n_2}};$$

— определяют число степеней свободы k из соотношения

$$\frac{1}{k} = \frac{c^2}{n_1 - 1} + \frac{1 - c^2}{n_2 - 1}, \text{ где } \qquad c = \frac{s_1^2 / n_1}{s_1^2 / n_1 + s_2^2 / n_2};$$

— для заданного уровня значимости α и числа степеней свободы k, по статистическим таблицам находят критическое значение $t_{\alpha,k}$;

— если $|t| \leq t_{\alpha,k}$, то принимается гипотеза H_0 о равенстве генеральных средних значений $\mu_1 = \mu_2$, в ином случае принимается альтернативная гипотеза.

В случае принятия гипотезы H_0 уточняют оценку генерального математического ожидания

$$\overline{x} = \frac{n_1 \overline{x}_1 + n_2 \overline{x}_2}{n_1 + n_2}.$$

5. Критерий проверки гипотетической плотности распределения

Пусть дана выборка из n независимых наблюдений случайной величины x, о которой делается предположение (выдвигается гипотеза H_0) о том, что ее плотность вероятности описывается функцией $w(x)$.

Для проверки гипотезы выполняют следующие операции:

— группируют n наблюдений по N интервалам — *интервалам группировки*. Число значений n_i случайной величины x, попавших в i-ый интервал $(x_i; x_i + \Delta x)$, отнесенное к объему выборки n, называют *наблюдаемой частотой* $f_i = n_i / n$, и используют ее для построения гистограммы;

— рассчитывается *ожидаемая частота* i-го интервала $F_i = w(x_i)\Delta x$ — относительная частота попаданий значений x в этот интервал, если бы плотность вероятности случайной величины x описывалась функцией $w(x)$;

— определяется расхождение между наблюдаемой и ожидаемой частотами $f_i - F_i$;

— вычисляется статистика

$$\chi^2 = \sum_{i=1}^{N} \frac{(f_i - F_i)^2}{F_i}.$$

Распределение случайной величины χ^2 приближенно совпадает с распределением χ_k^2, где k — число степеней свободы. В данном случае k равно объему выборки минус число различных независимых линейных ограничений, наложенных на данные наблюдений.

Первое ограничение: наблюдаемые частоты f_i взаимосвязаны между собой, действительно,

$$\sum_{i=1}^{N} f_i = 1,$$

и, следовательно, любую из них можно выразить через остальные, например последнюю f_N

$$f_N = 1 - f_1 - f_2 - \ldots - f_{N-1},$$

таким образом, $k = N - 1$.

Второе ограничение: если проверяется соответствие функции $w(x)$ закону распределения с неизвестными параметрами, например, нормальному распределению с неизвестными средним значением и дисперсией, то возникают еще ограничения, связанные с нахождением оценок этих параметров. Для нормального имеем еще два ограничения, связанных с нахождением оценок $\overline{x}$ и s^2. В результате $k = N - 3$.

— задавшись уровнем значимости, из статистических таблиц определяют критическое значение $\chi_{\alpha,k}^2$. Используем односторонний критерий, по верхней границе, так как любое отклонение от $w(x)$ приводит к возрастанию χ^2;

— если $\chi^2 \le \chi_{\alpha,k}^2$, то принимается гипотеза H_0 о законе распределения $w(x)$ с уровнем значимости α, если неравенство не выполняется, то H_0 отвергается, и подыскивается другая функция.

В заключение заметим, что интервал разбиения выбирается равным $\Delta x \simeq 0,4s$, но такой, чтобы в интервал попало не менее 3-х значений случайной величины x.