Глава 12

КРИТЕРИИ СОГЛАСИЯ

§ 1. СТАТИСТИЧЕСКИЙ КРИТЕРИЙ

Эксперимент. Предположим, что кто-то подбросил 10 раз монетку, и в 8 случаях она упала гербом вверх. Можно ли считать эту монетку симметричной?

Статистическая модель. Используем для описания эксперимента схему Бернулли, т. е. будем считать данные эксперимента реализацией выборки $\boldsymbol{X} = (X_1, \ldots, X_{10})$, где $X_i = 1$ (выпадает герб) с вероятностью θ и $X_i = 0$ (выпадает решка) с вероятностью $1 - \theta$. Как проверить гипотезу H о том, что $\theta = 1/2$?

Правило, позволяющее принять или отвергнуть гипотезу H на основе реализации выборки $x_1, \ldots, x_n$, называется *статистическим критерием*. Обычно критерий задается при помощи *статистики критерия* $T(x_1, \ldots, x_n)$ такой, что для нее типично принимать умеренные значения в случае, когда гипотеза H верна, и большие (малые) значения, когда H не выполняется.

Для приведенного выше эксперимента в качестве статистики T можно взять сумму $x_1 + \ldots + x_n$. Тогда гипотезе $H\colon \theta = 1/2$ противоречат значения, которые близки к 0 или n.

При проверке гипотез с помощью критериев всегда присутствует возможность ошибочно отвергнуть гипотезу H, когда на самом деле она верна. Например, симметричная монета может случайно упасть 10 раз подряд гербом вверх. Но вероятность наблюдать такое событие равна всего лишь $2^{-10} = 1/1024$. Если мы готовы пренебречь возможностью осуществления столь маловероятного события, то появление 10 гербов подряд следует считать основанием для отклонения гипотезы $H\colon \theta = 1/2$.

В общем случае задается малое число α — вероятность, с которой мы можем позволить себе отвергнуть верную гипотезу H (скажем, $\alpha = 0{,}05$). Это число называют *уровнем значимости*. Исходя из предположения, что гипотеза H верна, определяется наименьшее значение $x_{1-\alpha}$, удовлетворяющее условию

$$\mathsf{P}(T(X_1, \ldots, X_n) \geqslant x_{1-\alpha}) \leqslant \alpha. \qquad (1)$$

Если функция распределения статистики T непрерывна, то $x_{1-\alpha}$ является, очевидно, ее $(1-\alpha)$-квантилью (см. § 3 гл. 7). Такое $x_{1-\alpha}$ называют *критическим значением*: гипотеза H отвергается, если $t_0 = T(x_1,\ldots,x_n) \geqslant x_{1-\alpha}$ (произошло маловероятное событие), и принимается — в противном случае.

При этом величина $\alpha_0 = \mathbf{P}(T(X_1,\ldots,X_n) \geqslant t_0)$ задает *фактический уровень значимости*. Он равен вероятности того, что статистика T (измеряющая степень отклонения полученной реализации от наиболее типичной) за счет случайности примет значение t_0 или даже больше. Фактический уровень значимости — наименьший уровень, на котором проверяемая гипотеза H принимается (рис. 1).

Если значение T попало в область, имеющую при выполнении гипотезы H высокую вероятность, то можно заключить, что данные *согласуются* с гипотезой H. Отсюда происходит термин «*критерии согласия*».

Проверим для данных эксперимента гипотезу $H\colon \theta = 1/2$ на уровне значимости $\alpha = 0{,}05$ и вычислим α_0. Известно, что сумма $T = x_1 + \ldots + x_n$ имеет биномиальное распределение:

$$\mathbf{P}(T \geqslant k) = \sum_{i=k}^{n} C_n^i\, \theta^i (1-\theta)^{n-i}.$$

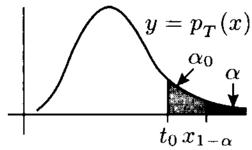

Рис. 1

Для $\theta = 1/2$ правая часть этого выражения при $k = 8$ равна $(45+10+1)/1024 \approx 0{,}055$ и при $k = 9$ равна $(10+1)/1024 \approx 0{,}011$. Поэтому для $\alpha = 0{,}05$ наименьшим $x_{1-\alpha}$, удовлетворяющим условию (1), будет 9. Поскольку полученное в эксперименте значение $t_0 = T(x_1,\ldots,x_n) = 8 < 9$, на заданном уровне значимости гипотеза $H\colon \theta = 1/2$ принимается.

С другой стороны, фактический уровень значимости $\alpha_0 = \mathbf{P}(T \geqslant 8) \approx 0{,}055$, что всего на 0,005 превосходит заданный уровень: уже при $\alpha = 0{,}06$ гипотезу H следует отклонить.

На основе данных эксперимента нельзя уверенно принять или отвергнуть гипотезу H (хотя последнее представляется более правдоподобным). Следовало бы еще несколько раз подбросить монетку, чтобы прийти к более взвешенному заключению.

Вопрос 1.
Чему приближенно равна вероятность наблюдать не менее 60 падений гербом вверх при 100 бросаниях симметричной монеты?
(Воспользуйтесь табл. Т2.)

Вычисление фактического уровня значимости нередко позволяет избегать категоричных (и при этом — ошибочных) выводов, сделанных лишь на основе сравнения t_0 с критическим значением $x_{1-\alpha}$, найденным для формально заданного α.

Не все стриги, что растет.

Козьма Прутков

§ 2. ПРОВЕРКА РАВНОМЕРНОСТИ

Пример 1. ***Орбиты планет и комет*** [72, с. 113]. В 1734 г. Французская академия присудила Даниилу Бернулли премию за исследование по орбитам планет, в котором он пытался показать, что схожесть орбит является неслучайной. Если предположить, что Солнце и планеты образовались в результате концентрации вещества первоначального «волчка» (рис. 2), то согласно *закону* *сохранения момента импульса* J орбиты планет должны лежать примерно в одной плоскости, что и наблюдается в реальности.

Д. Бернулли (1700–1782), швейцарский математик (племянник Якоба Бернулли (1654–1705), установившего в 1713 г. справедливость закона больших чисел для частоты «успехов» в независимых испытаниях). Д. Бернулли известен своими результатами в области механики жидкостей и газов. В 1778 г. им была опубликована в изданиях Петербургской Академии наук работа «Наиболее вероятное определение по нескольким расходящимся между собой наблюдениям и устанавливаемое отсюда наиболее правдоподобное заключение», где впервые был высказан и использован для оценки неизвестного параметра принцип максимального правдоподобия (см. [19, с. 419]).

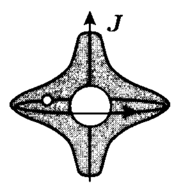

Рис. 2

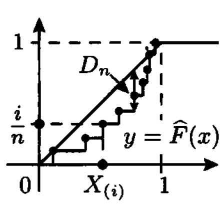

Рис. 3

В 1812 г. Лаплас исследовал схожую проблему: образовались ли и кометы в общем «волчке» или же они — всего лишь «гости», захваченные притяжением Солнца. В последнем случае углы между нормалями к плоскостям орбит комет и вектором J должны не концентрироваться вблизи нуля, а быть равномерно распределенными на отрезке $[0, \pi/2]$. Проведя статистическую обработку известных к тому времени астрономических данных, Лаплас пришел к выводу, что гипотеза о равномерности не отвергается.

Каким же образом можно проверить гипотезу равномерности? Рассмотрим несколько разных методов.

Метод 1. Критерий Колмогорова

Статистикой критерия является величина

$$D_n = \sup_{-\infty<x<\infty} |\widehat{F}_n(x) - F(x)|, \tag{2}$$

где $\widehat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \leqslant x\}}$ — это эмпирическая функция распределения, встречавшаяся ранее в § 1 гл. 9, $F(x)$ — функция распределения элементов выборки (на рис. 3 изображен случай равномерного распределения на $[0,1]$). Для любого фиксированного x_0 согласно усиленному закону больших чисел значение $\widehat{F}_n(x_0)$, равное частоте попаданий X_i левее x_0, при $n \to \infty$ с вероятностью 1 стремится к $F(x_0) = \mathbf{P}(X_1 \leqslant x_0)$. *Теорема Гливенко* утверждает, что для произвольной функции распределения $F(x)$ имеет место сходимость $D_n \xrightarrow{п.\,н.} 0$. (Доказательство этой теоремы приведено в [19, с. 206].) Поэтому в случае, когда гипотеза равномерности верна, значение D_n для выборки достаточно большого размера не должно существенно отклоняться от нуля.

Как количественно характеризуется значимость отклонения от нуля? В силу центральной предельной теоремы (П6)

$$\sqrt{n}\,(\widehat{F}_n(x_0) - F(x_0)) \xrightarrow{d} \xi \sim \mathcal{N}(0, F(x_0)(1 - F(x_0))).$$

Поэтому в фиксированной точке x_0 величина $|\widehat{F}_n(x_0) - F(x_0)|$ имеет порядок малости $1/\sqrt{n}$. Оказывается, что и величина D_n имеет тот же порядок малости, причем справедлив следующий результат.

Теорема Колмогорова. Если функция распределения элементов выборки $F(x)$ непрерывна, то для $x > 0$

$$\lim_{n\to\infty} \mathbf{P}(\sqrt{n}\, D_n \leqslant x) = K(x) = 1 + 2\sum_{k=1}^{\infty} (-1)^k e^{-2k^2x^2}.$$

Быстрая сходимость к предельному закону позволяет пользоваться этим приближением уже при $n \geqslant 20$.

Замечание 1. Особенностью статистики D_n является то, что закон ее распределения оказывается одним и тем же для всех непрерывных функций F. Он зависит только от размера выборки n. Действительно, полагая в формуле (2) $x = F^{-1}(y) = \sup\{x\colon F(x) = y\}$, $0 \leqslant y \leqslant 1$ (рис. 4), получаем $D_n = \sup\limits_{0\leqslant y\leqslant 1} |\widehat{F}_n(F^{-1}(y)) - y|$.

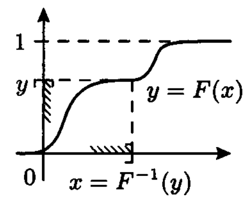

Рис. 4

Согласно методу обратной функции (см. § 1 гл. 4) случайные величины $Y_i = F(X_i)$ образуют выборку из равномерного распределения на отрезке $[0, 1]$. В силу монотонности и непрерывности функции $F(x)$ неравенства $x \leqslant F^{-1}(y)$ и $F(x) \leqslant y$ эквивалентны (см. рис. 4). Поэтому

$$\widehat{F}_n(F^{-1}(y)) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \leqslant F^{-1}(y)\}} = \frac{1}{n}\sum_{i=1}^{n} I_{\{Y_i \leqslant y\}}.$$

Правая часть — эмпирическая функция выборки $Y_1, \ldots, Y_n$.

Приведем таблицу некоторых квантилей функции $K(x)$:

α	0,5	0,15	0,1	0,05	0,025	0,01	0,001
$x_{1-\alpha}$	0,83	1,14	1,23	1,36	1,48	1,63	1,95

Таким образом, для заданного уровня значимости α критерий Колмогорова отвергает *гипотезу равномерности*, если для $F(x) = x$ величина $\sqrt{n}\, D_n(x_1, \ldots, x_n) \geqslant x_{1-\alpha}$.

Так как функция распределения $F(x)$ непрерывна и не убывает, а $\widehat{F}_n(x)$ — кусочно-постоянна, то sup в формуле (2) достигается в одной из точек разрыва функции $\widehat{F}_n$. Отсюда получаем простую формулу для вычисления значения $D_n(x_1, \ldots, x_n)$:

$$D_n(x_1, \ldots, x_n) = \max_{1\leqslant i\leqslant n} \left\{ \frac{i}{n} - F(x_{(i)}), F(x_{(i)}) - \frac{i-1}{n} \right\}.$$

Вопрос 2.
Как может выглядеть эмпирическая функция распределения $\widehat{F}_n$, для которой $D_n(x_1, \ldots, x_n) \neq \max\limits_{1\leqslant i\leqslant n} \left| \frac{i}{n} - F(x_{(i)}) \right|$?

В задаче 1 критерий применяется для проверки качества таблицы случайных чисел T1. Задача 3 показывает, что условие непрерывности функции $F(x)$ в теореме Колмогорова необходимо.

Метод 2. Критерий омега-квадрат

Статистика D_n измеряет отклонение эмпирической функции распределения $\widehat{F}_n$ от теоретической функции распределения F в равномерной метрике. Если воспользоваться (взвешенной) квадратичной метрикой, то получим *статистику критерия омега-квадрат*:

$$\omega_n^2(\psi) = \int_{-\infty}^{\infty} [\widehat{F}_n(x) - F(x)]^2\, \psi[F(x)]\, dF(x),$$

где $\psi(y)$ — заданная на $[0,1]$ весовая функция. Рассмотрим два варианта: $\psi_1 = 1$ (*критерий Крамéра — Мизеса*), $\psi_2(y) = 1/[y(1-y)]$ (*критерий Андерсона — Дарлинга*).

Первый из них хорошо улавливает расхождение между $\widehat{F}_n$ и F в области «типичных значений» случайной величины с функцией распределения F (часто он оказывается более чувствительным, чем критерий Колмогорова). Второй же, благодаря тому, что $\psi_2(y)$ быстро возрастает при $y \to 0$ и $y \to 1$, способен заметить различие «на хвостах» распределения F, которому придается дополнительный вес.

Так же, как и для статистики D_n, закон распределения величины $\omega_n^2(\psi)$ один и тот же для всех непрерывных функций F.

При выполнении ряда условий относительно ψ можно доказать, что существует $\lim\limits_{n\to\infty} \mathsf{P}(n\omega_n^2(\psi) \leqslant x) = A(x)$, зависящий от ψ. Для ψ_1 и ψ_2 известны разложения в ряды соответствующих законов $A_1(x)$ и $A_2(x)$ (см. [10, с. 83]). Приведем таблицу некоторых квантилей y_p и z_p этих законов ($A_1(y_p) = A_2(z_p) = p$):

α	0,5	0,15	0,1	0,05	0,025	0,01	0,001
$y_{1-\alpha}$	0,12	0,28	0,35	0,46	0,58	0,74	1,17
$z_{1-\alpha}$	0,77	1,62	1,94	2,49	3,08	3,88	5,97

Значения $n\omega_n^2(\psi_1)$ и $n\omega_n^2(\psi_2)$ вычисляются по следующим формулам (первая из них выводится в задаче 4):

$$n\omega_n^2(\psi_1) = \frac{1}{12n} + \sum_{i=1}^{n}\left[F(x_{(i)}) - \frac{2i-1}{2n}\right]^2,$$

$$n\omega_n^2(\psi_2) = -n - 2\sum_{i=1}^{n}\left[\frac{2i-1}{2n}\ln F(x_{(i)}) + \left(1 - \frac{2i-1}{2n}\right)\ln(1 - F(x_{(i)}))\right].$$

В гл. 18 появится еще один критерий (так называемый *критерий хи-квадрат*), с помощью которого можно проверять равномерность по сгруппированным данным.

§ 3. ПРОВЕРКА ПОКАЗАТЕЛЬНОСТИ

> Многие вещи нам непонятны не потому, что наши понятия слабы; но потому, что сии вещи не входят в круг наших понятий.
>
> *Козьма Прутков*

Прежде чем познакомиться с методом проверки показательности, введем формально понятие статистической гипотезы.

Напомним, что под статистической моделью в § 1 гл. 6 понималось семейство функций распределения $\{F(x,\theta),\ \theta \in \Theta\}$, где Θ — множество возможных значений параметра. При этом данные

$x_1,\ldots,x_n$ рассматривались как реализация выборки $X_1,\ldots,X_n$, элементы которой имеют функцию распределения $F(x,\theta_0)$ с неизвестным значением $\theta_0 \in \Theta$.

Пусть выделено некоторое подмножество $\Theta_0 \subset \Theta$. Под *статистической гипотезой* H понимается предположение о том, что $\theta_0 \in \Theta_0$. Если множество Θ_0 состоит всего из одной точки, то гипотеза H называется *простой*, иначе — *сложной*. В последнем случае задача заключается в проверке принадлежности закона распределения величин X_i целому классу функций распределения $\{F(x,\theta),\, \theta \in \Theta_0\}$.

Под гипотезой показательности понимается сложная гипотеза, в которой этот класс образуют функции распределения вида $F(x,\theta) = \left(1 - e^{-\theta x}\right) I_{\{x>0\}}$, где $\theta > 0$ (рис. 5). Рассмотрим методы проверки такой гипотезы.

Метод 1. Исключение неизвестного параметра

Согласно лемме 3 гл. 4, вектор $(S_1/S_n,\ldots,S_{n-1}/S_n)$, где $S_k = X_1 + \ldots + X_k$, распределен так же, как вектор порядковых статистик $(\eta_{(1)},\ldots,\eta_{(n-1)})$ для выборки размера $(n-1)$ из равномерного распределения на отрезке $[0,1]$.

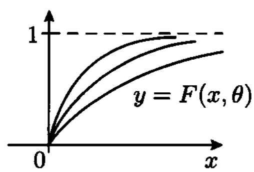

Рис. 5

Так как эмпирическая функция распределения строится по порядковым статистикам, то данное преобразование сводит задачу к проверке равномерности. Однако, за исключение «мешающего» параметра θ приходится платить уменьшением размера выборки на 1.

Пример 2. Следующие данные представляют собой количества летных часов между последовательными отказами установки для кондиционирования воздуха на самолете типа «Боинг-720» [40].

23	261	87	7	120	14	62	47	225	71
246	21	42	20	5	12	120	11	3	14
71	11	14	11	16	90	1	16	52	95

Считая времена между отказами независимыми, проверим гипотезу их показательности. Вычислим $s_k = x_1 + \ldots + x_k$, $k = 1,\ldots,30$:

23	284	371	378	498	512	574	621	846	917
1163	1184	1226	1246	1251	1263	1383	1394	1397	1411
1482	1493	1507	1518	1534	1624	1625	1641	1693	1788

В результате деления на $s_{30} = 1788$, получим ряд значений s_k/s_{30}, $k = 1,\ldots,29$:

0,013	0,159	0,207	0,211	0,279	0,286	0,321	0,347	0,473	0,513
0,650	0,662	0,686	0,697	0,700	0,706	0,773	0,780	0,781	0,789
0,829	0,835	0,843	0,849	0,858	0,908	0,909	0,918	0,947	

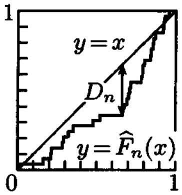

Рис. 6

Построенная по этому ряду эмпирическая функция распределения изображена на рис. 6. Максимальное отклонение $D_n = 0{,}65 - 10/29 \approx 0{,}306$, $\sqrt{29}\,D_n \approx 1{,}65$. Из приведенной выше таблицы квантилей функции Колмогорова находим, что гипотеза показательности отвергается на уровне значимости 1%.

Статистики $n\omega_n^2(\psi_1)$ и $n\omega_n^2(\psi_2)$ критерия омега-квадрат равны, соответственно, 0,627 и 3,036. Из таблицы квантилей $A_1(x)$ и $A_2(x)$ следует, что первая значимо велика на уровне приблизительно 2%, вторая — на уровне 2,5%.

Метод 2. Подстановка оценки параметра

Пусть $\tilde{\theta}_n$ — оценка максимального правдоподобия для параметра θ (см. § 4 гл. 9). Рассмотрим *модифицированные статистики Колмогорова и Краме́ра—Мизеса*:

$$\tilde{D}_n = \sup_x \left|\widehat{F}_n(x) - F(x,\tilde{\theta}_n)\right|,$$

$$\tilde{\omega}_n^2 = \int_{-\infty}^{\infty} \left[\widehat{F}_n(x) - F(x,\tilde{\theta}_n)\right]^2 dF(x,\tilde{\theta}_n). \qquad (3)$$

Замечание 2 [80, с. 317]. Эти статистики, в отличие от их прототипов D_n и $\omega_n^2(\psi_1)$, не обладают свойством «свободы от распределения» элементов выборки, поэтому для каждого параметрического семейства распределений нужны отдельные таблицы. Более того, их распределения могут зависеть и от истинного значения неизвестного параметра (параметров). К счастью, для семейств сдвига-масштаба (к которым относятся, в частности, показательный и нормальный законы) этого последнего осложнения не возникает.

Несложно проверить (см. задачу 3 гл. 9), что оценкой максимального правдоподобия для параметра θ показательного закона является $\tilde{\theta}_n = 1/\overline{X}$, где $\overline{X} = (X_1 + \ldots + X_n)/n$, которая ранее встречалась в замечании к примеру 1 гл. 6.

Вопрос 3.
Почему в случае показательного закона распределение статистики $\tilde{D}_n$ не зависит от θ?

М. Стефенс (см. [35]) предложил вместо статистик $\sqrt{n}\,\tilde{D}_n$ и $n\tilde{\omega}_n^2$ использовать для *показательной модели* их несколько преобразованные варианты $(\sqrt{n}+0{,}26+0{,}5/\sqrt{n})\,(\tilde{D}_n - 0{,}2/n)$ и $(n+0{,}16)\,\tilde{\omega}_n^2$, распределения которых практически не зависят от n, начиная с $n = 5$. Приведем таблицу соответствующих квантилей $\tilde{x}_{1-\alpha}$ и $\tilde{y}_{1-\alpha}$ этих распределений (рассчитанную методом Монте-Карло):

α	0,15	0,1	0,05	0,025	0,01
$\tilde{x}_{1-\alpha}$	0,926	0,990	1,094	1,190	1,308
$\tilde{y}_{1-\alpha}$	0,149	0,177	0,224	0,273	0,337

Еще один критерий для проверки показательности (*«Новое лучше старого»*) рассматривается в § 1 гл. 13.

§ 4. ПРОВЕРКА НОРМАЛЬНОСТИ

Б. Л. Ван дер Варден в [13, с. 84] пишет:

«Я до сих пор живо помню, как однажды, когда я был еще ребенком, мой отец привел меня на край города, где на берегу стояли ивы, и велел мне сорвать наугад сотню ивовых листочков. После отбора листьев с поврежденными кончиками у нас осталось 89 целых листиков. Вернувшись домой, мы расположили их в ряд по росту, как солдат. Затем мой отец через кончики листьев провел кривую и сказал: «Это и есть кривая Кетле. Глядя на нее, ты видишь, что посредственности всегда составляют большинство и лишь немногие поднимаются выше или так и остаются внизу».

Если эту кривую расположить вертикально (рис. 7) и в качестве единицы масштаба на оси ординат выбрать отрезок, длина которого равна высоте всей фигуры, то ордината h, соответствующая абсциссе t, будет, очевидно, представлять собой частоту (или долю) тех ивовых листьев, длина которых меньше t. И так как частота h приближенно равна вероятности, то наша кривая приближенно представляет $p = F(x)$ — функцию распределения длины листьев.»

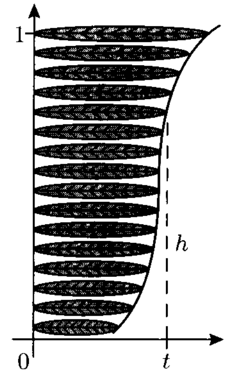

Рис. 7

Л. А. Ж. Кетле (1796–1874), бельгийский социолог.

Как проверить сложную двухпараметрическую *гипотезу нормальности* о том, что выборка была взята из совокупности с функцией распределения $F(x,\mu,\sigma) = \Phi\left(\dfrac{x-\mu}{\sigma}\right)$ с какими-то неизвестными параметрами μ и $\sigma > 0$? (Здесь, как обычно, $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-u^2/2}du$ — функция распределения стандартного нормального закона.)

Прежде чем применять критерии, полезно посмотреть на данные на вероятностной бумаге (см. § 1 гл. 9): если точки $\left(x_{(i)}, \Phi^{-1}\left(\dfrac{i-0{,}5}{n}\right)\right)$ не расположены вблизи некоторой прямой, то гипотеза нормальности скорее всего ошибочна.

Чтобы ее формально отвергнуть, можно использовать

Метод 1. Исключение неизвестных параметров

Пусть m — произвольное, но заранее фиксированное целое число от 1 до n. Положим

$$A_m = \frac{1}{n+\sqrt{n}} \sum_{i=1}^{n} X_i - \frac{1}{1+\sqrt{n}} X_m,$$

$$Y_j = \begin{cases} X_j - A_m, & \text{если } j = 1,\ldots,m-1, \\ X_{j+1} - A_m, & \text{если } j = m,\ldots,n-1. \end{cases}$$

Оказывается, случайные величины $Y_1,\ldots,Y_{n-1}$ независимы и одинаково распределены по закону $\mathcal{N}(0,\sigma^2)$ (ввиду нормальности $Y_1,\ldots,Y_{n-1}$ достаточно (см. П9) проверить, что $\mathbf{M}Y_j = 0$, $\mathbf{D}Y_j = \sigma^2$ и $\mathbf{cov}(Y_i,Y_j) = 0$ при $i \neq j$).

Переход от выборки $X_1,\ldots,X_n$ к набору случайных величин $Y_1,\ldots,Y_{n-1}$ позволяет избавиться от неизвестного параметра сдвига μ, однако при этом размерность данных уменьшается на 1. (Можно было бы попытаться исключить параметр μ за счет следующего простого преобразования: $X_i' = X_i - \overline{X}$, $i = 1,\ldots,n$. Однако X_i' не будут образовывать выборку.)

Вопрос 4. Зависимы ли а) X_i' и X_j' при $i \neq j$, б) X_i' и $\overline{X}$?

Для исключения оставшегося параметра σ совершим еще одно преобразование:

$$Z_k = Y_k \Big/ \sqrt{B_k}\,, \quad \text{где} \quad B_k = \frac{1}{n-k-1}\sum_{j=k+1}^{n-1} Y_j^2, \quad k = 1,\ldots,n-2.$$

К. Саркади показал (см. [10, с. 57]), что случайные величины $Z_1,\ldots,Z_{n-2}$ также независимы, причем Z_k, очевидно, подчиняется закону Стьюдента t_{n-k-1} (см. § 2 гл. 11).

Обозначим через $F_{n-k-1}(x)$ функцию распределения закона t_{n-k-1}. Если гипотеза нормальности верна, то (согласно методу обратной функции из § 1 гл. 4) случайные величины $F_{n-k-1}(Z_k)$ должны быть независимыми и равномерно распределенными на $[0,1]$. Это проверяется с помощью одного из критериев, рассмотренных в § 2.

Вопрос 5. Применим ли к ним критерий Колмогорова?

Таблицы значений функций F_{n-k-1} для разных степеней свободы приведены в [10, с.174]. Для вычисления ее на компьютере можно численно проинтегрировать плотность, задаваемую формулой (3) гл. 11 (множитель c_{n-k-1} в которой легко определяется из свойства гамма-функции $\Gamma(x+1) = x\Gamma(x)$ и тождества $\Gamma(1/2) = \sqrt{\pi}$), или воспользоваться рекуррентными формулами из [20, с. 22, 51], позволяющими выразить F_{n-k-1} через элементарные функции.

Метод 2. Подстановка оценок параметров

Согласно задаче 2 гл. 9, оценками максимального правдоподобия параметров μ и σ являются соответственно $\overline{X}$ и S, где $\overline{X} = \frac{1}{n}\sum X_i$, $S^2 = \frac{1}{n}\sum (X_i - \overline{X})^2$.

Статистики $\sqrt{n}\,\tilde{D}_n$ и $n\tilde{\omega}_n^2$, определяемые формулой (3) при $F(x,\tilde{\theta}_n) = \Phi((x-\overline{X})/S)$, сходятся при $n \to \infty$ к некоторым предельным законам. М. Стефенс (см. [35]) установил, что для *нормальной модели* распределения (сходящихся к тем же законам) модифицированных статистик $(\sqrt{n}-0{,}01+0{,}85/\sqrt{n})\,\tilde{D}_n$ и $(n+0{,}5)\,\tilde{\omega}_n^2$ практически не зависят от n при $n \geqslant 5$. Приведем таблицу соответствующих квантилей $\tilde{x}_{1-\alpha}$ и $\tilde{y}_{1-\alpha}$ этих распределений:

α	0,15	0,1	0,05	0,025	0,01
$\tilde{x}_{1-\alpha}$	0,775	0,819	0,895	0,955	1,035
$\tilde{y}_{1-\alpha}$	0,091	0,104	0,126	0,148	0,178

Замечание 3. Важно отметить, что предельные распределения статистик $\sqrt{n}\,\tilde{D}_n$ и $n\tilde{\omega}_n^2$ отличаются от $K(x)$ и $A_1(x)$. Дело в том, что при вычислении значения $\tilde{\theta}_n$ используются те же самые $x_1,\ldots,x_n$, что и при построении эмпирической функции распределения. Поэтому $\widehat{F}_n(x)$ и $F(x,\tilde{\theta}_n)$ (в случае, если проверяемая гипотеза верна) оказываются ближе друг к другу, чем $\widehat{F}_n(x)$ и $F(x,\theta_0)$. При этом критическими становятся *существенно меньшие* значения статистик, чем в случае простой гипотезы. Например, сравнение $\tilde{x}_{0,95} = 0{,}895$ с медианой $x_{1/2} = 0{,}83$ и квантилью $x_{0,95} = 1{,}36$ функции распределения $K(x)$ показывает, что значения, типичные для $\sqrt{n}\,D_n$, оказываются *критическими* для статистики $\sqrt{n}\,\tilde{D}_n$.

Метод 3. Центральные выборочные моменты

Простые критерии (см. [13, с. 281]), которые несколько больше, чем критерий Колмогорова, учитывают поведение «хвостов» распределения, основаны на *центральных выборочных моментах* $M_k = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{X})^k$, $k = 1,2,\ldots$.

При помощи величин M_2, M_3 и M_4 вычисляются *выборочные коэффициенты асимметрии G_1 и эксцесса G_2*:

$$G_1 = M_3/M_2^{3/2}, \quad G_2 = M_4/M_2^2 - 3.$$

Эти случайные величины можно использовать в качестве оценок для (независящих от сдвига и масштаба) *теоретических коэффициентов асимметрии $\gamma_1 = \mu_3/\mu_2^{3/2}$ и эксцесса $\gamma_2 = \mu_4/\mu_2^2 - 3$*, где $\mu_k = \mathbf{M}(X_1 - \mathbf{M}X_1)^k$ — *центральные теоретические моменты*. (Для нормального закона $\gamma_1 = \gamma_2 = 0$.)

При конечных n целесообразно заменить G_1 и G_2 на

$$G_1' = \frac{\sqrt{n(n-1)}}{n-2}\,G_1 \quad \text{и} \quad G_2' = \frac{n-1}{(n-2)(n-3)}\,[(n+1)\,G_2 + 6].$$

Если истинное распределение является *нормальным*, то математические ожидания величин G_1' и G_2' в точности равны нулю, а дисперсии задаются формулами

$$\sigma_1^2 = \frac{6\,n(n-1)}{(n-2)(n+1)(n+3)}, \quad \sigma_2^2 = \frac{24\,n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}.$$

При этом статистики G_1'/σ_1 и G_2'/σ_2 асимптотически нормальны (см. § 4 гл. 7): распределение каждой из них сходится к закону $\mathcal{N}(0,1)$ при $n \to \infty$. Значимость их отклонения от нуля можно определить по таблице Т2.

Замечание 4 [10, с. 56]. Как показал Э. Пирсон (1930 г.), распределение статистики G_1'/σ_1 довольно быстро приближается к $\mathcal{N}(0,1)$, тогда как распределение величины G_2'/σ_2 даже при больших n

оказывается далеким от нормального. Р. Гири (1935 г.) предложил заменить ее на статистику $G_3 = \frac{1}{n}\sum |X_i - \overline{X}|/S$, у которой

$$\mathbf{M}G_3 = \sqrt{\frac{2}{\pi}}\left[1 + \frac{2}{8n-9} + O\left(\frac{1}{n^3}\right)\right],$$

$$\mathbf{D}G_3 = \frac{1}{n}\left[\left(1 - \frac{3}{\pi}\right) - \frac{1}{4\pi n} + O\left(\frac{1}{n^2}\right)\right].$$

Распределение величины $(G_3 - \mathbf{M}G_3)/\sqrt{\mathbf{D}G_3}$ удовлетворительно аппроксимируется стандартным нормальным законом при $n \geqslant 50$.

Задача 2 показывает, что, как правило, *нельзя надежно проверить сложную гипотезу по небольшой выборке* (состоящей из нескольких десятков наблюдений): критерии улавливают только очень крупные отклонения, так как за счет варьирования параметра (параметров) обычно удается достаточно хорошо подогнать $F(x,\theta)$ к эмпирической функции распределения $\widehat{F}_n(x)$.

С другой стороны, для выборок большого размера (порядка нескольких сотен наблюдений) трудно гарантировать одинаковость условий при сборе данных (однородность наблюдений).

По-видимому, хороший способ разрешения этой проблемы — использование статистических методов, не предполагающих строгую нормальность наблюдений, для которых требуется лишь непрерывность функции распределения элементов выборки. Именно такие критерии рассматриваются в гл. 14–17.

Проверить нормальность по сгруппированным данным можно также при помощи критерия хи-квадрат (см. гл. 18).

§ 5. ЭНТРОПИЯ

Обозначим через ξ некоторый эксперимент с исходами $A_1,\ldots,A_N$, которые осуществляются с вероятностями $p_1,\ldots,p_N$ соответственно. *Энтропией* этого эксперимента называется величина

$$H = H(\xi) = -\sum_{i=1}^{N} p_i \log_2 p_i, \tag{4}$$

где по непрерывности полагаем $0 \cdot \log_2 0 = 0$. Ясно, что $H \geqslant 0$, причем $H = 0$ тогда и только тогда, когда все вероятности p_i, кроме одной, равны нулю.

Утверждение. Максимум энтропии H, равный $\log_2 N$, достигается при $p_1 = \ldots = p_N = 1/N$.

Доказательство. Поскольку вторая производная функции $\varphi(x) = x\log_2 x$ положительна при $x > 0$, то эта функция выпукла (П4). Записывая неравенство Иенсена (П4) для случайной

Задачи

1. Вычислить теоретический коэффициент эксцесса γ_2 для равномерного распределения на отрезке [0, 1].

2. Вычислить теоретический коэффициент асимметрии γ_1 для показательного (экспоненциального) распределения.

3. Вычислить центральный теоретический момент 4-го порядка μ_4 для стандартного нормального распределения $N(0, 1)$.

4. Для выборки из распределения Бернулли построить график: а) теоретической функции распределения, б) эмпирической функции распределения.

5. Найти предельное распределение статистики критерия Колмогорова $\sqrt{n}D_n$ для выборки из распределения Бернулли. (Указание. Используйте центральную предельную теорему.)

6*. Для статистики критерия Крамера — Мизеса $\omega_n^2(\psi_1)$ вывести следующую формулу:

$$\omega_n^2(\psi_1) = \frac{1}{12n} + \sum_{i=1}^{n} \left(F(X_{(i)}) - \frac{2i-1}{2n} \right)^2.$$