С. М. БОРОДАЧЁВ

МЕТОДЫ
МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Министерство образования и науки Российской Федерации
Уральский федеральный университет
имени первого Президента России Б. Н. Ельцина

С. М. БОРОДАЧЁВ

МЕТОДЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Учебное пособие

Екатеринбург
УрФУ
2012

УДК 519.22(075.8)

ББК 22.172я73

Б83

Рецензенты:

УрГУПС кафедра «Высшая и прикладная математика» (зав кафедрой д-р физ.-мат. наук, проф. Г. А. Тимофеева);

В. И. Зенков д-р физ.-мат. наук директор негосударственного образовательного учреждения «Новые технологии и программы».

Научный редактор – д-р экон. наук проф. С. В. Кортов

Бородачёв, С. М.

Б83 Методы математической статистики : учебное пособие /

С. М. Бородачёв. – Екатеринбург : УрФУ, 2012. – 129 с.

ISBN 978–5–321–02218–4

Математическая статистика составляет методологическую основу всех областей знаний, работающих с реальными данными. Пособие содержит теоретический материал, упражнения, лабораторный практикум и задания для самостоятельной работы. Предназначено для студентов естественнонаучных, экономических и др. направлений всех форм обучения.

Библиогр. : 12 назв. Прил. 1.

УДК 519.22(075.8)

ББК 22.172я73

ISBN 978–5–321–02218–4

© Уральский федеральный университет, 2012

© Бородачёв С. М., 2012

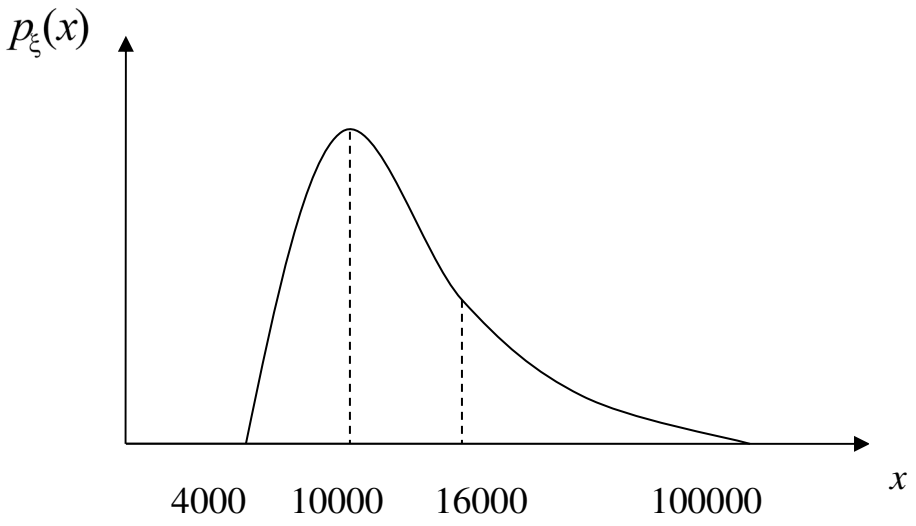
Введение

Задача статистики – изучение реального мира путём анализа числовых данных, полученных в результате наблюдений (экспериментов) над ним.

Пример

Пусть хотим изучить положение с зарплатой сотрудников некоторого предприятия (какая бывает, средняя величина и т. п.). Обозначим ξ - зарплату случайно встреченного сотрудника этого предприятия. В том, какая она может оказаться, какая чаще и т. п., т. е. в распределении случайной величины ξ, и находится информация о положении с зарплатой.

Плотность распределения случайной величины ξ может иметь следующий вид (но нам неизвестна!):



Средняя зарплата на предприятии есть математическое ожидание ξ:

$a = M\xi = \int\limits_{-\infty}^{+\infty} x p_\xi(x) dx$. Пусть она равна 16000 р., нам это не узнать, т. к. плотность неизвестна.

Для хотя бы приближённого определения средней зарплаты проведём опрос случайно встреченных сотрудников. Опросили $N = 10$ чел.:

$x^1 = 10\,000,\ x^2 = 21\,000, \ldots, x^{10} = 8000$. Подсчитали среднее арифметическое:

$\overline{x} = \dfrac{x^1 + x^2 + \ldots + x^{10}}{10} = 12\,000$. Можно ли считать, что a = 12 000 р.?

Чтобы подстраховаться, решили повторить опрос во второй день: $x^1 = 20\,000,\ x^2 = 9000, \ldots, x^{10} = 18\,000$. Подсчитали среднее арифметическое:

$\overline{x} = \dfrac{x^1 + x^2 + \ldots + x^{10}}{10} = 14\,000$. Можно ли считать, что a = 14 000 р.? Но математическое ожидание – величина не случайная (постоянная), так что a, скорее всего, и не 12 000 и не 14 000. Почему же $\overline{x}$ получились разными? Если рассмотреть данные различных дней, то видно, что результат первого наблюдения X^1* был разным, т. е. это величина случайная. Аналогично ей случайны остальные X^n, поэтому и $\overline{X} = \dfrac{X^1 + X^2 + \ldots + X^N}{N} = \dfrac{1}{N}\sum_{n=1}^{N} X^n$ случайна. Зачем тогда её подсчитывали? Найдём её математическое ожидание

$$M\overline{X} = M\frac{1}{N}\sum_{n=1}^{N} X^n = \frac{1}{N}M\sum_{n=1}^{N} X^n = \frac{1}{N}\sum_{n=1}^{N} MX^n = \frac{1}{N}a\sum_{n=1}^{N} 1 = \frac{1}{N}aN = a. \qquad (1)$$

Здесь использовано равенство $MX^n = M\xi = a$, т. к. результат каждого n-го наблюдения есть у нас «экземпляр» ξ. То, что математическое ожидание среднего арифметического результатов наблюдений равно искомой величине a, оправдывает его (среднего арифметического) применение.

---

* Будем обозначать случайные величины греческими буквами ξ, η, … или прописными латинскими X, Y, …, а их наблюдённые значения, как и прочие не случайные величины, соответствующими строчными буквами: x, y, …

Однако если ограничиться одним измерением, то будет $MX^1 = a$. Зачем тогда измерять N раз? Найдём дисперсию $D\overline{X}$. Поскольку X^n – независимые величины, $DX^n = D\xi = \sigma^2$, постольку

$$D\overline{X} = D\frac{1}{N}\sum_{n=1}^{N} X^n = \frac{1}{N^2} D\sum_{n=1}^{N} X^n = \frac{1}{N^2}\sum_{n=1}^{N} DX^n = \\ = \frac{1}{N^2}\sigma^2 \sum_{n=1}^{N} 1 = \frac{1}{N^2}\sigma^2 N = \frac{\sigma^2}{N}. \tag{2}$$

Поэтому чем больше опрошенных, тем скачки результата (вокруг искомого a) будут меньше, а значит, вероятнее, что $\overline{x}$ находится вблизи a.

1. Точечные оценки

Выборка объёма N – совокупность N независимых и одинаково распределённых случайных величин

$$X^1,\ X^2, ..., X^n, ..., X^N, \tag{3}$$

представляющих собой результаты N измерений (наблюдений) случайной величины ξ. Функция распределения: $F_{X^n}(x) = F_\xi(x)$ для всех n.

Множество возможных значений случайной величины ξ с распределением $F_\xi(x)$ называется *генеральной совокупностью (популяцией)* с распределением $F_\xi(x)$, а наблюдённые значения $x^1, x^2, ..., x^n, ..., x^N$ – *выборка* из генеральной совокупности.

Задача статистики – на основе выборочных данных сделать выводы обо всей популяции. При этом случайная выборка должна представлять всю популяцию (быть *репрезентативной*). Например, если проводить опросы о зарплате (см. пример из «Введения») в утренние часы, то вы узнаете кое-что о зарплате только той категории сотрудников, которые приходят на работу в ранние часы. Или ещё

пример нерепрезентативной выборки: попытка выяснить отношение к Интернету всего населения, путём публикации вопроса на каком-либо сайте.

Любая функция выборки $Y = f(X^1, X^2, ..., X^n, ..., X^N)$ называется *статистикой*.

1.1. Примеры статистик и их распределений

Пусть (3) – выборка, имеющая нормальное стандартное распределение (т. е. $\xi \sim N(0,1)$. Знак «тильда» (~) означает «имеет распределение»).

1. Статистика $\chi_N^2 = (X^1)^2 + (X^2)^2 + ... + (X^N)^2$ имеет распределение, называемое *распределением «хи-квадрат» с N степенями свободы*.

Можно показать, что

$$M\chi_N^2 = N,\ D\chi_N^2 = 2N\ . \tag{4}$$

2. Статистика

$$T_{N-1} = \frac{X^N}{\sqrt{\chi_{N-1}^2 / (N-1)}} \tag{5}$$

имеет распределение, называемое *распределением Стьюдента с* (N-1) *степенями свободы.*

3. Обозначим

$$F_{N_1,N_2} = \frac{\chi_{N_1}^2 / N_1}{\chi_{N_2}^2 / N_2}. \tag{6}$$

Если $\chi_{N_1}^2$ и $\chi_{N_2}^2$ независимы, то F_{N_1,N_2} имеет распределение, называемое *распределением Фишера с* N_1, N_2 *степенями свободы.*

1.2. Оценка

Пусть по выборке объёма N требуется оценить неизвестный параметр θ распределения $P\{\xi \le x\} = F_\xi(\theta, x)$.

Оценкой неизвестного параметра θ называется статистика $\widehat{\Theta}_N = \hat{\theta}_N(X^1, X^2, ..., X^N)$, реализация которой, полученная в результате опытов, $\hat{\theta}_N = \hat{\theta}_N(x^1, x^2, ..., x^N)$ – принимается за неизвестное истинное значение параметра θ.

Произвольно построенная оценка не обязательно будет близкой к оцениваемому параметру θ. Введём два свойства оценок, которые описывают их близость к соответствующим параметрам.

Оценка $\widehat{\Theta}_N$ параметра θ называется *несмещённой*, если её математическое ожидание $M\widehat{\Theta}_N = \theta$. В противном случае $M\widehat{\Theta}_N = \theta + b(\theta)$, $b(\theta)$ – смещение (*bias*) оценки.

Оценка $\widehat{\Theta}_N$ параметра θ называется *состоятельной*, если она сходится по вероятности к оцениваемому параметру θ: $\widehat{\Theta}_N \xrightarrow[N\to\infty]{P} \theta$. Это означает, что для любого $\varepsilon > 0$, $P\left\{\left|\widehat{\Theta}_N - \theta\right| \le \varepsilon\right\} \xrightarrow[N\to\infty]{} 1$.

Свойство несмещённости означает, что оценка не имеет систематической ошибки. Свойство состоятельности обеспечивает сближение оценки с измеряемым параметром при увеличении числа измерений.

Ошибка оценки, т. е. отклонение оценки от истинного значения параметра ($\widehat{\Theta}_N$ – θ), есть величина случайная. *Среднеквадратическая ошибка* (или *погрешность* как нечто среднее, характеризующее многократные действия) *оценки* $\widehat{\Theta}_N$ параметра θ – это корень квадратный из математического ожидания квадрата ошибки оценки.

$$RMSE\widehat{\Theta}_N = \sqrt{M(\widehat{\Theta}_N - \theta)^2}\,, \qquad (7)$$

где *RMS – root mean square*.

Иногда говорят о *точности оценки* параметра (в процентах):

$$\delta = \frac{RMSE\widehat{\Theta}_N}{\theta} \cdot 100\ \%\,.$$

Стандартной ошибкой оценки $\widehat{\Theta}_N$ параметра θ называется квадратный корень из какой-либо подходящей оценки её дисперсии

$$S.E.\widehat{\Theta}_N = \sqrt{\widehat{D\widehat{\Theta}_N}}\,. \qquad (8)$$

В ГОСТ Р 50779.10–2000 «Статистические методы. Вероятность и основы статистики. Термины и определения» в п. 2.56 содержится неверное определение этого понятия: стандартное отклонение оценки!

1.3. Оценка математического ожидания и дисперсии

Пусть нас интересует параметр $a = M\xi$. Предположим, что существует $D\xi = \sigma^2$.

Несмещённой и состоятельной оценкой математического ожидания случайной величины ξ является статистика $\hat{A}_N = \overline{X} = \frac{1}{N}\sum_{n=1}^{N} X^n$, называемая *выборочное среднее*.

Доказательство. Несмещённость – см. формулу (1), состоятельность – следует доказать в упражнении* 1.5.

Пусть теперь по выборке (3) нужно оценить два параметра: математическое ожидание и дисперсию.

---

* Далее – *упр.*

В качестве оценки первого выберем выборочное среднее. Предположим, что существует $M\xi^4$.

Несмещённой и состоятельной оценкой дисперсии случайной величины ξ является статистика

$$\hat{\Sigma}_N^2 = S^2 = \frac{1}{N-1}\sum_{n=1}^{N}(X^n - \overline{X})^2 . \qquad (9)$$

Эту формулу можно представить в виде

$$S^2 = \frac{1}{N-1}\left[\sum_{n=1}^{N} X^n - N\overline{X}^2\right]. \qquad (10)$$

Пример 1.1

Оценить математическое ожидание и дисперсию случайной величины по выборке: 3, 7, 7, 0, 2, 1, 2, 3, 5, 4. $N = 10$.

$$\hat{a} = \overline{x} = \frac{1}{10}\cdot(3+7+7+0+2+1+2+3+5+4) = 3.4;$$

$$\widehat{\sigma^2} = s^2 = \frac{1}{9}\cdot[3^2 + 7^2 + ... + 4^2 - 10\cdot(3.4)^2] = 5.6.$$

Стандартное отклонение можно оценить так: $\hat{\sigma} = \sqrt{\widehat{\sigma^2}} = s = \sqrt{5.6} = 2.37.$ Но это будет смещённая оценка стандартного отклонения (в частности, для нормального распределения всегда заниженная).

Стандартная ошибка оценки математического ожидания:

$$S.E.\hat{A}_N = \sqrt{\widehat{D\hat{A}_N}} = \sqrt{\widehat{D\overline{X}}} = \sqrt{\frac{\widehat{\sigma^2}}{N}} = \sqrt{\frac{5.6}{10}} = 0.748.$$

1.4. Теорема Фишера

Теорема

Пусть выборка (3) взята из нормального распределения: $\xi \sim N(a, \sigma)$. Тогда статистики $\overline{X}$ и S^2 независимы. Статистика $\dfrac{(N-1)S^2}{\sigma^2}$ имеет распределение χ^2_{N-1}.

Как известно из линейной алгебры, если матрица A [N x N] симметрична и rang $A = r <= N$, то имеет место сингулярное разложение $A = S\Lambda S^T$, где $S = (\vec{e}^1, \vec{e}^2, ..., \vec{e}^N)$ – матрица из собственных векторов матрицы A; $\lambda_1 > 0, \lambda_2 > 0, ..., \lambda_r > 0, \lambda_{r+1} = 0, ..., \lambda_N = 0$ – соответствующие собственные числа. $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_N)$. Собственные векторы ортонормированы, матрица S ортогональна. Сингулярное разложение можно переписать как $A = \sum_{k=1}^{r} \lambda_k \vec{e}^k \vec{e}^{kT}$. Легко доказать, что $Sp(A) = \sum_{k=1}^{r} \lambda_k$. Если матрица A ещё и идемпотентна, то у идемпотентной матрицы $\lambda_k = 1$ или 0, постольку $Sp(A) = r$.

Лемма 1

Обозначим $\vec{X} = (X^1, X^2, ..., X^n, ..., X^N)^T$ – вектор-выборка из $N(0,\ 1)$. Пусть $Q = \vec{X}^T A \vec{X}$, где матрица A симметрична. $\vec{t} = B\vec{X}$, где B – матрица [m x N]. Если $BA = 0$, то функции Q и $\vec{t}$ независимы.

Доказательство леммы 1. По условию леммы $0 = BA = \sum_{k=1}^{r} \lambda_k (B\vec{e}^k)\vec{e}^{kT}$.

Умножим это равенство на вектор $\vec{e}^s$ справа. В силу ортогональности собственных векторов получим

$$B\vec{e}^s = \vec{0} \text{ для любого } s. \tag{11}$$

Подставляя в Q сингулярное разложение, получим $Q = \sum_{k=1}^{r} \lambda_k (\vec{e}^{kT} \vec{X})^2$. $\vec{t}$ состоит из t_i, Q – из $\vec{e}^{kT} \vec{X}$, поэтому утверждение будет доказано, если показать, что $\operatorname{cov}(t_i, \vec{e}^{kT} \vec{X}) = 0$ (т. к. величины нормальны, это будет означать независимость). В последнем равенстве легко убедиться, используя (11).

Лемма 2

Обозначим $\vec{X} = (X^1, X^2, ..., X^n, ..., X^N)^T$ – вектор-выборка из $N(0,\ 1)$. Пусть $Q = \vec{X}^T A \vec{X}$. Матрица A симметрична и идемпотентна, rang $A = r$, тогда $Q \sim \chi^2_r$.

Доказательство леммы 2. Имеем $Q = \sum_{k=1}^{r} (\vec{e}^{kT} \vec{X})^2$. Так как собственные векторы ортонормированы, то случайные величины $\vec{e}^{kT} \vec{X}$, $k = 1, \ldots, r$, некоррелированны, независимы (поскольку нормальны), с распределением $N(0,1)$. Значит $Q \sim \chi^2_r$.

Доказательство теоремы Фишера. Ясно, что после стандартизации достаточно рассмотреть случай $N(0,\ 1)$.

Введём N-мерный вектор $\vec{b} = (1/N, 1/N, ..., 1/N)^T$ и матрицу $B = (\vec{b},\ \vec{b}, ..., \vec{b})$. Тогда $\overline{X} = \vec{b}^T \vec{X}$, а $(N-1)S^2 = (\vec{X} - B\vec{X})^T (\vec{X} - B\vec{X})$, или $(N-1)S^2 = \vec{X}^T A \vec{X}$, где $A = I - B$ идемпотентна. Вследствие того, что $\vec{b}^T A = \vec{b}^T - \vec{b}^T B = \vec{b}^T - \vec{b}^T = 0$, по лемме 1 $\overline{X}$ и S^2 независимы.

Поскольку $Sp(A) = Sp(I) - Sp(B) = N - 1$, постольку по лемме 2 $(N-1)S^2 \sim \chi^2_{N-1}$.

1.5. Оценка математического ожидания и ковариационной матрицы случайного вектора

Пусть есть N независимых наблюдений $\vec{X}^n = \begin{pmatrix} X_1^n \\ X_2^n \\ \dots \\ X_r^n \end{pmatrix}$, $n = 1, \dots, N$, случайного вектора $\vec{\xi} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \dots \\ \xi_r \end{pmatrix}$.

Математическое ожидание случайного вектора. $M\vec{\xi}$ есть $\begin{pmatrix} M\xi_1 \\ M\xi_2 \\ \dots \\ M\xi_r \end{pmatrix}$, поэтому оценка:

$$\widehat{M\vec{\xi}} = \begin{pmatrix} \frac{1}{N}\sum_{n=1}^{N} X_1^n \\ \frac{1}{N}\sum_{n=1}^{N} X_2^n \\ \dots \\ \frac{1}{N}\sum_{n=1}^{N} X_r^n \end{pmatrix} = \frac{1}{N}\sum_{n=1}^{N} \begin{pmatrix} X_1^n \\ X_2^n \\ \dots \\ X_r^n \end{pmatrix} = \frac{1}{N}\sum_{n=1}^{N} \vec{X}^n = \overline{\vec{X}}.$$

Ковариационная матрица.

Теорема

Несмещённой и состоятельной оценкой ковариации случайных величин ξ_i, ξ_j является

$$\widehat{\operatorname{cov}(\xi_i,\xi_j)} = \frac{1}{N-1}\sum_{n=1}^{N}(X_i^n - \overline{X_i})(X_j^n - \overline{X_j}), \tag{12}$$

где $\overline{X_i} = \frac{1}{N}\sum_{n=1}^{N} X_i^n$.

Оценка ковариационной матрицы – матрица из оценок ковариаций:

$$\widehat{K_{\vec{\xi}}} = \begin{pmatrix} \widehat{D\xi_1} & \widehat{\operatorname{cov}(\xi_1,\xi_2)} & \cdots & \widehat{\operatorname{cov}(\xi_1,\xi_r)} \\ \widehat{\operatorname{cov}(\xi_2,\xi_1)} & \widehat{D\xi_2} & \dots & \widehat{\operatorname{cov}(\xi_2,\xi_r)} \\ \dots & \dots & \ddots & \dots \\ \widehat{\operatorname{cov}(\xi_r,\xi_1)} & \widehat{\operatorname{cov}(\xi_r,\xi_2)} & \dots & \widehat{D\xi_r} \end{pmatrix}.$$

Это можно представить в виде

$$\widehat{K_{\vec{\xi}}} = \frac{1}{N-1}\sum_{n=1}^{N}(\overrightarrow{X^n} - \overline{\vec{X}})(\overrightarrow{X^n} - \overline{\vec{X}})^T, \tag{13}$$

или если ввести матрицу «объект–свойство», содержащую все наблюдения

$$z = \begin{pmatrix} \vec{x}^{1T} \\ \vec{x}^{2T} \\ \dots \\ \vec{x}^{NT} \end{pmatrix} = \begin{pmatrix} x_1^1 & \dots & x_r^1 \\ \dots & \dots & \dots \\ x_1^N & \dots & x_r^N \end{pmatrix}, \tag{14}$$

то

$$\widehat{K_{\vec{\xi}}} = \frac{1}{N-1}[z^T z - \frac{1}{N} z^T \vec{1}\vec{1}^T z] \tag{15}$$

для наблюдённой оценки.

Оценка коэффициента корреляции:

$$\widehat{\rho(\xi_i,\xi_j)} = \frac{\widehat{\mathrm{cov}(\xi_i,\xi_j)}}{\sqrt{\widehat{D\xi_i}\,\widehat{D\xi_j}}}; \qquad (16)$$

оценка корреляционной матрицы: $\widehat{R_\xi} = (\widehat{\rho[\xi_i,\xi_j]})$.

1.6. Оценка стандартного отклонения через размах выборки

Размахом выборки (3) называется случайная величина R_N:

$$R_N = \max(X^1, X^2, \ldots, X^N) - \min(X^1, X^2, \ldots, X^N).$$

Теорема

Функция распределения размаха $F_{R_N}(x)$ имеет вид:

$$F_{R_N}(x) = N\int_{-\infty}^{+\infty} p_\xi(y)[F_\xi(y+x) - F_\xi(y)]^{N-1}dy,$$

где $p_\xi(x)$, $F_\xi(x)$ – плотность и функция распределения случайной величины ξ. Если величина ξ имеет нормальное распределение [ξ ~ $N(a, \sigma)$], то $MR_N = \sigma d_N$ и $DR_N = \sigma^2 D_N$, коэффициенты пропорциональности d_N и D_N зависят только от N.

N	2	3	4	5	6	7	8	9
d_N	1,128	1,693	2,059	2,326	2,534	2,704	2,847	2,970
D_N	0,727	0,789	0,774	0,747	0,719	0,694	0,672	0,653

$M(\frac{R_N}{d_N}) = \sigma$, следовательно,

$$\hat{\sigma} = \frac{r_N}{d_N} \qquad (17)$$

– несмещённая оценка стандартного отклонения σ. Напомним, что часто применяемая оценка s – смещённая оценка стандартного отклонения.

Пример 1.2

По выборке из нормального распределения 40, 38, 37, 52, 42 найти несмещённую оценку стандартного отклонения.

Размах выборки $r_5 = 52 - 37 = 15$, несмещённая оценка стандартного отклонения $\hat{\sigma} = \frac{r_5}{d_5} = \frac{15}{2.326} = 6.45$. Сравним: $\hat{\sigma} = \sqrt{s^2} = 6.02$ – смещённая оценка.

Пусть есть несколько (I) выборок объёмом N каждая. Найдём несмещённую оценку стандартного отклонения по этим выборкам. Введём среднее арифметическое размахов R_N^i $\overline{R}_N = \frac{1}{I}\sum_{i=1}^{I} R_N^i$. Ясно, что

$$\hat{\sigma} = \frac{\overline{r}_N}{d_N} \tag{18}$$

– несмещённая оценка стандартного отклонения с дисперсией в I раз меньше, чем у (17).

1.7. Оценка параметра *p* гипергеометрического распределения

Пусть есть партия M лотерейных билетов, из которых Q выигрышные. Обозначим долю выигрышных билетов $p = \frac{Q}{M}$. Вы купили N билетов. M_N – число выигрышных билетов из купленных, случайная дискретная величина. Её распределение называется гипергеометрическим: $M_N \sim \mathrm{Hypgeo}(M, N, p)$. Известно, что

$$MM_N = Np, \tag{19}$$

$$DM_N = Npq\frac{M-N}{M-1}, \qquad (20)$$

где $q = 1 - p$.

Из (19) следует, что $M(\frac{M_N}{N}) = p$, которое означает, что доля выигрышных билетов $\hat{P}_N = \frac{M_N}{N}$ в выборке есть несмещённая оценка доли p в партии.

Рассмотрим дисперсию этой оценки:

$$\sigma^2_{\hat{P}_N} = D\hat{P}_N = D\frac{M_N}{N} = \frac{1}{N^2}DM_N = \frac{1}{N^2}Npq\frac{M-N}{M-1} = \frac{pq(M-N)}{N(M-1)}.$$

Заменив p и q их подходящими оценками, получим оценку дисперсии оценки: $\hat{\sigma}^2_{\hat{P}_N} = \frac{\hat{p}\hat{q}(M-N)}{N(M-1)}$. Поэтому стандартная ошибка оценки доли:

$$S.E.\hat{P}_N = \hat{\sigma}_{\hat{P}_N} = \sqrt{\frac{\hat{p}\hat{q}(M-N)}{N(M-1)}}. \qquad (21)$$

Из формулы (21) можно найти объем бесповторной выборки N, обеспечивающей заданную стандартную ошибку:

$$N = \frac{\hat{p}\hat{q}M}{\hat{p}\hat{q} + \hat{\sigma}^2_{\hat{P}_N}(M-1)}. \qquad (22)$$

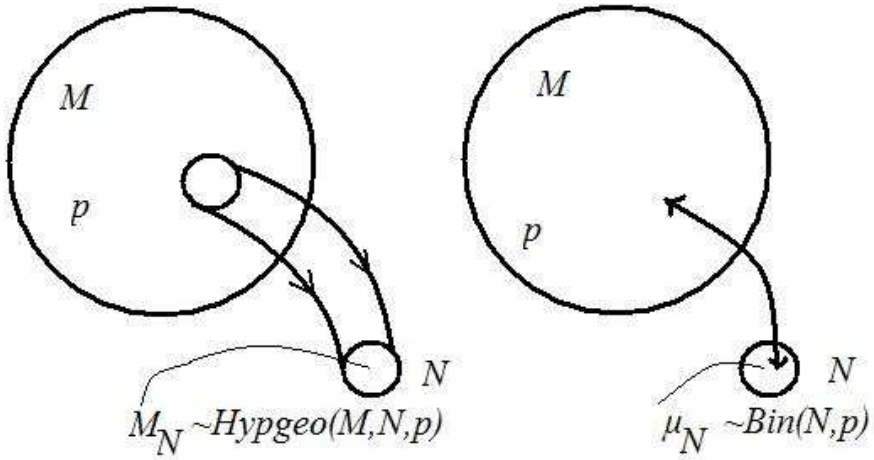
Иногда в формуле (22) заменяют $(M - 1)$ на M и вводят максимальную погрешность, которая равна стандартной ошибке, умноженной на γ – коэффициент запаса (2÷3): $\Delta_{\hat{p}} = \gamma\hat{\sigma}_{\hat{P}_N}$, тогда

$$N = \frac{\hat{p}\hat{q}\gamma^2 M}{\hat{p}\hat{q}\gamma^2 + M(\Delta_{\hat{p}})^2}. \qquad (23)$$

1.8. Оценка параметра *p* биномиального распределения

Число успехов в N испытаниях схемы Бернулли имеет биномиальное распределение $\mu_N \sim \mathrm{Bin}(N, p)$, где p – вероятность успеха в отдельном испытании.

Рассмотрим связь гипергеометрического распределения с биномиальным. На рисунке изображены выборки: бесповторная и независимых испытаний.



Ясно, что различие в распределении числа «выигрышей, успехов» в выборке N уменьшается при увеличении объёма партии M:

$$\mathrm{Hypgeo}(M, N, p) \xrightarrow[M \to \infty]{} \mathrm{Bin}(N, p).$$

Поэтому $\widehat{p_{\mathrm{Bin}}} = \lim\limits_{M \to \infty} \hat{p}_{\mathrm{Hypgeo}} = \lim\limits_{M \to \infty} \dfrac{m_N}{N} = \dfrac{\mu_N}{N}$, т. е. частота успеха $\dfrac{\mu_N}{N}$ есть несмещённая и состоятельная оценка параметра биномиального распределения

1.9. Оценка функции распределения

Определение функции распределения: $F_\xi(x) = P\{\xi \le x\}$. Назовём событие в фигурных скобках успехом. N независимых наблюдения случайной величины

(выборка) можно рассматривать как испытания, в которых может быть успех или неудача. Поэтому оценка функции распределения на аргументе x находится как оценка вероятности p успеха в отдельном испытании. По предыдущему разделу она равна частоте успеха:

$$\widehat{F_\xi(x)} = \frac{N\{\xi \le x\}}{N} = \frac{\textit{число элементов выборки, меньших либо равных } x}{N}.$$

Последнее выражение обозначается $\widehat{F_N}(x)$ и называется эмпирической функцией распределения. Таким образом, эмпирическая функция распределения есть несмещённая и состоятельная оценка функции распределения.

1.10. Метод максимального правдоподобия (МП)

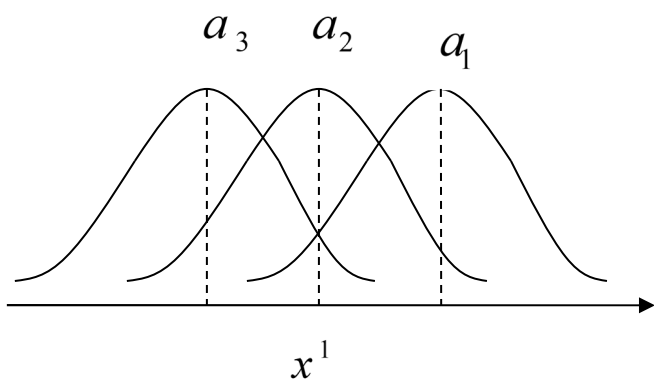
Пусть мы знаем распределение случайной величины ξ с точностью до входящих в него параметров $\xi \sim p_\xi(x, \vec{\theta})$, $\vec{\theta} = (\theta_1, \theta_2, \ldots, \theta_r)^T$. Например, ξ имеет нормальное распределение, а вектор параметров $\vec{\theta} = \begin{pmatrix} a \\ \sigma \end{pmatrix}$ неизвестен.

Метод максимального правдоподобия – это общий метод нахождения (конструирования) оценок, когда за оценку параметра $\vec{\theta}$ берётся такое его значение, при котором вероятность получить наблюдённую выборку $x^1, x^2, \ldots, x^N$ максимальна.

Пример 1.3

Знаем, что $\xi \sim N(a, \sigma)$ – нормальное распределение. Какое значение лучше принять для математического ожидания, зная одно измерение x^1 величины ξ?

Рассмотрим три варианта (см. рисунок).



Вероятнее всего получить x^1 при $a = a_2$, поэтому $\hat{a} = a_2 = x^1$.

Плотность совместного распределения выборки (3): $p_{X^1, X^2, ..., X^N}(x^1, x^2, ..., x^N, \vec{\theta}) = p_{\vec{X}}(\vec{x}, \vec{\theta})$. Вероятность получить выборку вблизи $\vec{x}$ как раз пропорциональна этой плотности, поэтому оценка параметра: $\hat{\vec{\theta}}_{\text{МП}} = \arg\max\limits_{\hat{\vec{\theta}}} p_{\vec{X}}(\vec{x}, \hat{\vec{\theta}})$, где $\vec{x}$ – наблюдённое значение выборки.

Плотность распределения выборки на наблюдённом значении выборки (или вероятность получить наблюдённую выборку) как функция оцениваемого параметра $\vec{\theta}$ называется *функцией правдоподобия* $L(\vec{\theta}) = p_{\vec{X}}(\vec{x}, \vec{\theta})$.

Поскольку $L(\vec{\theta})$ и $\ln L(\vec{\theta})$ одновременно достигают максимума, то иногда функцией правдоподобия называют $\ln L(\vec{\theta})$ и работают с ней.

$$\hat{\vec{\theta}}_{\text{МП}} = \arg\max_{\hat{\vec{\theta}}} \ln L(\hat{\vec{\theta}}). \qquad (24)$$

Если наблюдения в выборке независимы, то функция правдоподобия $L(\vec{\theta})$ распадается на произведение одномерных распределений: $L(\vec{\theta}) = p_{X^1}(x^1, \vec{\theta}) p_{X^2}(x^2, \vec{\theta}) ... p_{X^N}(x^N, \vec{\theta})$, а если они ещё одинаково распределены, то $L(\vec{\theta}) = p_{\xi}(x^1, \vec{\theta}) ... p_{\xi}(x^N, \vec{\theta})$.

Пример 1.4

Найти МП-оценку параметра λ показательного распределения $\xi \sim M(\lambda)$ по выборке $x^1, x^2, ..., x^N$.

$$L(\lambda) = p_\xi(x^1, \lambda)...p_\xi(x^N, \lambda) = \lambda e^{-\lambda x^1}...\lambda e^{-\lambda x^N} = \lambda^N e^{-\lambda(x^1+x^2+...+x^N)},$$

$$\ln L(\lambda) = N \ln \lambda - \lambda(x^1 + x^2 + ... + x^N),$$

$$\frac{\partial \ln L(\lambda)}{\partial \lambda} = \frac{N}{\lambda} - (x^1 + x^2 + ... + x^N) = 0,\ \hat{\lambda}_{\text{МП}} = \frac{N}{x^1 + x^2 + ... + x^N} = \frac{1}{\overline{x}}.$$

Эта оценка согласуется со связью истинного λ и истинного математического ожидания: $\lambda = \dfrac{1}{M\xi}$.

Оценки метода МП являются состоятельными, но могут быть смещёнными.

1.11. Неравенство Рао–Крамера. Эффективные оценки

Теорема

Пусть (3) – независимая выборка, $p_\xi(x, \theta)$ – единая для всех N плотность распределения наблюдений, тогда в некоторых достаточно общих условиях регулярности этой плотности имеет место неравенство для *RMS*-ошибки оценки $\widehat{\Theta}_N$ параметра θ:

$$\sqrt{M(\widehat{\Theta}_N - \theta)^2} \geq \frac{\left|1 + db(\theta)/d\theta\right|}{\sqrt{NI(\theta)}}, \tag{25}$$

где $b(\theta)$ – смещение оценки; $I(\theta) = M\left[\dfrac{\partial \ln p_\xi(X, \theta)}{\partial \theta}\right]^2$ – информация Фишера о параметре θ, содержащаяся в одном наблюдении.

Оценка параметра называется *эффективной,* если её среднеквадратичная ошибка равняется границе Рао–Крамера (т. е. обращает неравенство (25) в равенство, или минимально возможна).

Можно показать, что если эффективная оценка существует, то она единственна, и что любая эффективная оценка является несмещённой.

Пример 1.5

Проверить эффективность оценки $\bar{x}$ в нормальной модели.

Левая часть неравенства Рао–Крамера:

$$\sqrt{M(\bar{X}-a)^2} = \sqrt{D\bar{X}} = \sqrt{\frac{\sigma^2}{N}} = \frac{\sigma}{\sqrt{N}}.$$

Правая часть неравенства Рао–Крамера: $b(a)=0$,

$$I(a) = M\left[\frac{\partial}{\partial a}\ln\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(X-a)^2}{2\sigma^2}}\right]^2 = M\left[\frac{\partial}{\partial a}(\ln\frac{1}{\sqrt{2\pi}\sigma} - \frac{(X-a)^2}{2\sigma^2})\right]^2 =$$

$$= M\left[\frac{1}{2\sigma^2}2(X-a)\right]^2 = \frac{1}{\sigma^4}DX = \frac{1}{\sigma^4}\sigma^2 = \frac{1}{\sigma^2},$$

$$\frac{1}{\sqrt{N/\sigma^2}} = \frac{\sigma}{\sqrt{N}}.$$

Левая часть неравенства равна правой части, следовательно, $\bar{x}$ – эффективная оценка математического ожидания.

Пример 1.6

Проверить эффективность оценки s^2 в нормальной модели $N(a, \sigma)$.

Левая часть (из теоремы Фишера):

$$\sqrt{M(S^2-\sigma^2)^2} = \sqrt{DS^2} = \sqrt{D\frac{\chi^2_{N-1}\sigma^2}{N-1}} = \frac{\sigma^2\sqrt{2}}{\sqrt{N-1}}. \qquad (26)$$

Правая часть: $b(\sigma^2)=0, I(\sigma^2)=\dfrac{1}{2\sigma^4}$ (см. *упр.* 1.22), $\dfrac{1}{\sqrt{N/2\sigma^4}}=\dfrac{\sigma^2\sqrt{2}}{\sqrt{N}}$.

Левая часть больше правой части: $\dfrac{\sigma^2\sqrt{2}}{\sqrt{N-1}}>\dfrac{\sigma^2\sqrt{2}}{\sqrt{N}}$, следовательно, s^2 не является эффективной оценкой.

Можно доказать, что не существует эффективной оценки дисперсии в этом случае.

1.12. Достаточные статистики

Некоторые статистики способны извлекать из данных больше информации, чем другие. На этом основан принцип достаточности. Достаточной оценкой называется статистика, в определённом смысле собирающая всю информацию о параметре, которая содержится в выборке.

Определение: статистика $\widehat{\Theta}_N^*$ называется *достаточной* для параметра θ, если условное распределение выборки при $\widehat{\Theta}_N^*=\hat{\theta}_N^*$ не зависит от θ (т. е. зная $\hat{\theta}_N^*$, мы вычерпали всю информацию о θ из выборки).

Критерий факторизации для достаточности: статистика $\hat{\theta}_N^*$ достаточна тогда и только тогда, когда плотность распределения выборки может быть разложена в произведение вида

$$p_{X^1,X^2,\ldots,X^N}(x^1,x^2,\ldots,x^N,\theta)=g\{\hat{\theta}_N^*(\vec{x}),\theta\}h(\vec{x}),$$

где множитель h не зависит от θ (в частности, может быть постоянным).

Пример 1.7

Нормальная модель с известной дисперсией – $N(\theta, 1)$.

$$p_{X^1,X^2,...,X^N}(x^1, x^2, ..., x^N, \theta) = (2\pi)^{-N/2} \exp\{-\tfrac{1}{2}\sum_{n=1}^{N}(x^n - \theta)^2\} =$$

$$= (2\pi)^{-N/2} \exp\{-\tfrac{1}{2}\sum_{n=1}^{N}(x^n - \overline{x})^2\} \exp\{-\tfrac{1}{2}N(\overline{x} - \theta)^2\} = \qquad (27)$$

$$= h(\vec{x}) g\{\hat{\theta}_N^*(\vec{x}), \theta\},$$

где $\hat{\theta}_N^*(\vec{x}) = \overline{x}$.

Из (27) следует, что статистика $\overline{x}$ достаточна для математического ожидания θ нормальной случайной величины.

Пример 1.8

Нормальная модель с нулевым математическим ожиданием $N(0, \theta)$:

$$p_{X^1,X^2,...,X^N}(x^1, x^2, ..., x^N, \theta) = (2\pi\theta^2)^{-N/2} \exp\{-\tfrac{1}{2}\sum_{n=1}^{N}(x^n)^2 / \theta^2\} = g\{\hat{\theta}_N^*(\vec{x}), \theta\},$$

откуда ясно, что $\hat{\theta}_N^*(\vec{x}) = \sum_{n=1}^{N}(x^n)^2$ – достаточная статистика для стандартного отклонения θ (а значит, и для дисперсии θ^2).

Из последнего примера видно, что достаточная статистика ещё не есть несмещённая оценка параметра.

Теорема Рао–Блекуэлла–Колмогорова.

Оптимальная оценка (несмещённая с минимально возможной дисперсией), если она существует, является функцией от достаточной статистики. Более того, если существует полная достаточная статистика, то всякая функция от неё является оптимальной оценкой математического ожидания этой функции.

Пример 1.9

Оценивание параметра θ равномерного распределения $U(0, \theta)$. Можно показать, что последний элемент $X_{(N)}$ вариационного ряда является достаточной статистикой для параметра θ. Легко показать, что математическое ожидание

$M(X_{(N)}) = \dfrac{N}{N+1}\theta$. Значит, $\dfrac{N+1}{N}X_{(N)}$ является оптимальной оценкой параметра θ. Дисперсия её $\dfrac{\theta^2}{N(N+2)}$.

Всякая эффективная оценка является одновременно достаточной статистикой.

1.13. Теоретико-игровой подход к оцениванию

Необходимые сведения о статистических решающих правилах см. в гл. 6.

Оценивание – частный случай применения решающего правила. Истинное значение параметра θ — это состояние природы θ, значение $\hat{\theta}$ оценки, которое мы принимаем вместо параметра – это действие *a*, которое предписывает решающее правило $a = d(\vec{x})$. Данные эксперимента $\vec{x}$ - теперь это выборка. Решающая функция *d*(.) – теперь оценка (статистика) $\hat{\theta} = d(\vec{x})$. Несоответствие значения оценки и параметра есть недостаток, ведущий к потерям, описанным функцией потерь $L(\hat{\theta}, \theta)$, аналогично потерям из-за несоответствия действия состоянию природы $L(a, \theta)$. Ищется оптимальная оценочная статистика – решающее правило *d*(.).

Распределение выборки считаем известным $p_{\vec{X}}(\vec{x} \mid \theta)$, функция риска (средние потери от применения оценки *d*(.) при истинном параметре θ)

$$R(d, \theta) = M\{L[d(\vec{X}), \theta] \mid \theta\}.$$

Минимаксная оценка

Так как мы не знаем точного θ, то для каждой оценки предположим наихудшею ситуацию для риска, а затем найдем, у которой оценки это наихудшее получше $d_{mm} = \arg\min_{d} \max_{\theta} R(d, \theta)$.

Если есть набор конкурирующих оценок, перебором можно найти $d_{mm}(.)$. Неудобно, т.к. оценка сама не конструируется.

Байесовская оценка

Будем предполагать, что параметр θ = Θ – случайная величина с известным априорным распределением $p(\theta)$, тогда любую оценку $d(.)$ можно охарактеризовать средним по значениям параметра риском (*байесовский риск оценки d*) $r_B(d) = MR(d, \Theta)$ и затем выбрать $d_B = \arg\min_d r_B(d)$. Но здесь у нас есть теорема (байесовский принцип), автоматически конструирующая формулу байесовской оценки:

$$d_B(\vec{x}) = \arg\min_{\hat{\theta}} M[L(\hat{\theta},\Theta) \mid \vec{x}] = \arg\min_{\hat{\theta}} \int L(\hat{\theta},\theta)\, p(\theta \mid \vec{x}) d\theta .$$

Например, если функция потерь квадратична $L(\hat{\theta},\theta) = (\hat{\theta} - \theta)^2$,

$$d_B(\vec{x}) = \arg\min_{\hat{\theta}}[\hat{\theta}^2 1 - 2\hat{\theta}\int \theta p(\theta \mid \vec{x}) d\theta + \int \theta^2 p(\theta \mid \vec{x}) d\theta] =$$

$$= \arg\min_{\hat{\theta}}[\hat{\theta}(\hat{\theta} - 2\overline{\theta}_{\vec{x}})] = \overline{\theta}_{\vec{x}}.$$

Во втором выражении 3-е слагаемое опущено, т.к. не зависит от $\hat{\theta}$. <u>Т.о. байесовская оценка параметра при квадратичной функции потерь есть среднее апостериорного распределения параметра.</u>

$$d_B(\vec{x}) = M(\Theta \mid \vec{x}) = \int \theta p(\theta \mid \vec{x}) d\theta .$$

Другие естественные функции потерь ведут к медиане и моде (см. [10]).

Пример 1.10

Модель регрессии [6]. $\vec{Y} = z\vec{\beta} + \vec{E}$ или $\vec{\beta} = (z^T z)^{-1}[z^T \vec{Y} - z^T \vec{E}]$. Считаем параметр $\vec{\beta} = \vec{\mathrm{B}}$ случайным. Апостериорное его распределение $p_{\vec{\mathrm{B}}}(\vec{\beta} \mid \vec{y}, z) =$ распределение$[(z^T z)^{-1} z^T \vec{y} - (z^T z)^{-1} z^T \vec{E}]$ = |если $\vec{E} \sim N(\vec{0}, \sigma^2 I)$, то линейная функция нормальной величины вновь распределена нормально [9, стр. 204] | =

$$= N[(z^T z)^{-1} z^T \vec{y} - (z^T z)^{-1} z^T \vec{0}, -(z^T z)^{-1} z^T \sigma^2 I(-(z^T z)^{-1} z^T)^T] =$$

$$= N[(z^T z)^{-1} z^T \vec{y}, \sigma^2 (z^T z)^{-1}].$$

Значит, среднее этого апостериорного распределения $(z^T z)^{-1} z^T \vec{y} = \hat{\vec{\beta}}_B$ и есть байесовская оценка вектора регрессии. Формула совпадает с оценкой по МНК. Сразу видим ковариационную матрицу этой оценки $\sigma^2 (z^T z)^{-1}$.

В более широком плане байесовский подход в статистике состоит в получении и изучении апостериорного распределения параметра $p(\theta \mid \vec{x}) \propto p(\theta) p(\vec{x} \mid \theta)$, где функция правдоподобия $p(\vec{x} \mid \theta)$ считается известной, а априорное распределение $p(\theta)$ выражает представление (догадки) исследователя о значении параметра. Если полученное распределение выражается через стандартные, то можно провести генерацию выборки параметра и методом Монте-Карло получить точечную оценку параметра, правдоподобные интервалы и т.п. Если возникают трудности с получением выборки, то просто рассчитывают апостериорное распределение на сетке точек параметра.

Естественно, результаты при концентрированном априорном распределении будут тяготеть к нему, а при неопределённом – в основном определяться данными $\vec{x}$, зачастую совпадая с традиционными («небаевскими») оценками.

Упражнения

Упражнение 1.1

Доказать свойство симметрии распределения Фишера

$$P\{F_{N_1,N_2} \le x\} = P\{F_{N_2,N_1} \ge \frac{1}{x}\}.$$

Упражнение 1.2

По выборке 8, 5, 7, 8, 3, 7, 5, 3, 2, 6 значений случайной величины ξ найти несмещённые оценки математического ожидания $M\xi$ и дисперсии $D\xi$, оценку

стандартного отклонения σ. Какова стандартная ошибка оценки математического ожидания?

Упражнение 1.3

Пусть известно $M\xi = a$. Какую можно предложить оценку дисперсии по выборке? Доказать её несмещённость.

Упражнение 1.4

Доказать несмещённость оценки дисперсии по S^2.

Упражнение 1.5

С помощью неравенства Чебышёва $P\{|\xi - M\xi| < \varepsilon\} \geq 1 - \frac{D\xi}{\varepsilon^2}$ доказать, что для состоятельности несмещённой (асимптотически несмещённой) оценки достаточно стремление её дисперсии к 0 при N, стремящемся к бесконечности. Доказать состоятельность оценки математического ожидания по выборочному среднему.

Упражнение 1.6

Найти среднеквадратические ошибки оценки дисперсии по S^2 и по $S^{*2} = \frac{1}{N}\sum_{n=1}^{N}(X^n - \overline{X})^2$ в нормальной модели. Сделать выводы.

Упражнение 1.7

Что больше: среднеквадратическая ошибка оценки или её стандартное отклонение? Найти связь между ними через смещение оценки $b(\theta)$.

Упражнение 1.8

В 5 магазинах одинакового профиля зафиксированы за месяц следующие показатели, тыс. р.:

Затраты на рекламу	8	10	5	8	9
Объём продаж	100	300	100	200	300

Найти оценку коэффициента корреляции между затратами на рекламу и объёмом продаж у произвольного магазина этого профиля.

Упражнение 1.9

Доказать несмещённость и состоятельность оценки ковариации (выразить ковариацию из формулы дисперсии суммы и оценить правую часть).

Упражнение 1.10

Доказать формулу (13).

Упражнение 1.11

Доказать теорему о распределении размаха выборки.

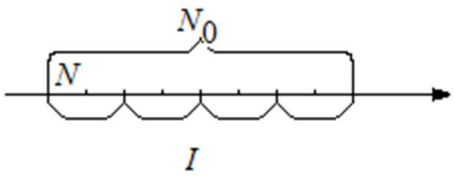
Упражнение 1.12

Показать, что для нормального распределения математическое ожидание размаха пропорционально стандартному отклонению. Рассчитать в *MathCAD* коэффициенты пропорциональности для объёмов выборки 2÷10.

Упражнение 1.13

При каком объёме N выборок, при общем фиксированном числе наблюдений N_0 (см. рисунок), оценка стандартного отклонения через размахи (18) будет иметь наименьшую погрешность?

Какова наибольшая точность оценки (в процентах)?



Упражнение 1.14

По данным опроса 400 респондентов, перед выборами мэра Екатеринбурга за кандидата Г. высказались 21 %. Какова оценка процента избирателей города, поддерживающих Г., и её стандартная ошибка?

Упражнение 1.15

Вывести формулу для необходимого объёма бесповторной выборки (22).

Упражнение 1.16

Какое число студентов в группе из 25 чел. следует опросить («за», «против»), чтобы выяснить долю «за» со стандартной ошибкой 0.1? Предварительный опрос 5 чел. дал 3 «за». Сколько нужно было бы провести опросов «с возвращениями»?

Упражнение 1.17

Доказать состоятельность оценки параметра $\mathrm{Bin}(N, p)$ (рассмотреть дисперсию оценки параметра, используя связь с гипергеометрическим распределением).

Упражнение 1.18

N независимых измерений размера a проведены различными приборами с погрешностями σ_n. Считая $X^n \sim N(a, \sigma_n)$, найти МП-оценку a. Доказать несмещённость этой оценки и найти её погрешность.

Упражнение 1.19

Найти МП-оценку σ для $N(a, \sigma)$ при известном a.

Упражнение 1.20

Баскетболист впервые промахнулся в k-м броске. Найти МП-оценку вероятности его промаха в отдельном броске. Будет ли она несмещённой?

Упражнение 1.21

Рассмотреть эффективность оценки $\overline{x}$ параметра $\theta = 1/\lambda$ в $M(\lambda)$.

Упражнение 1.22

Найти в нормальной модели информацию Фишера о дисперсии, содержащуюся в одном наблюдении.

Упражнение 1.23

Доказать, что если математическое ожидание a известно, то $S_0^2 = \frac{1}{N}\sum_{n=1}^{N}(X^n - a)^2$ является эффективной оценкой дисперсии в случае нормального распределения.

Упражнение 1.24

Доказать критерий факторизации для достаточности статистики.

2. Интервальные оценки (доверительные интервалы)

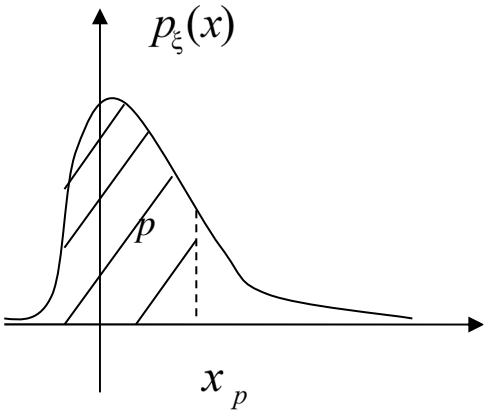
В примере 1.1 мы получили оценку математического ожидания по выборке $\hat{a} = \overline{x} = 3.4.$ Это точечная оценка. Однако не ясно, насколько она близка к истинному значению *a*. Те же выборочные данные могут дать информацию: с вероятностью γ = 0.95 интервал (1.7, 5.1) накрывает *a*. Это интервальная оценка параметра *a*.

Доверительным интервалом параметра θ с доверительной вероятностью γ, порождённым выборкой $x^1, x^2, ..., x^N$, называется интервал с границами $w_1(x^1, x^2, ..., x^N)$ и $w_2(x^1, x^2, ..., x^N)$, которые являются реализациями случайных величин $W_1 = w_1(X^1, X^2, ..., X^N)$ и $W_2 = w_2(X^1, X^2, ..., X^N)$, таких, что $P\{W_1 \le \theta \le W_2\} = \gamma$.

2.1. Доверительный интервал для математического ожидания нормальной случайной величины

Случай, когда дисперсия известна. Пусть $\xi \sim N(a, \sigma)$, где a неизвестно; σ известна. $\overline{X} \sim N(a, \frac{\sigma}{\sqrt{N}})$ имеет нормальное распределение (как линейная комбинация нормальных случайных величин). Стандартизируем её: $\frac{\overline{X} - a}{\sigma / \sqrt{N}} \sim N(0, 1)$. Будем строить интервал: $P\left\{ ? \le \frac{\overline{X} - a}{\sigma / \sqrt{N}} \le ? \right\} = \gamma$.

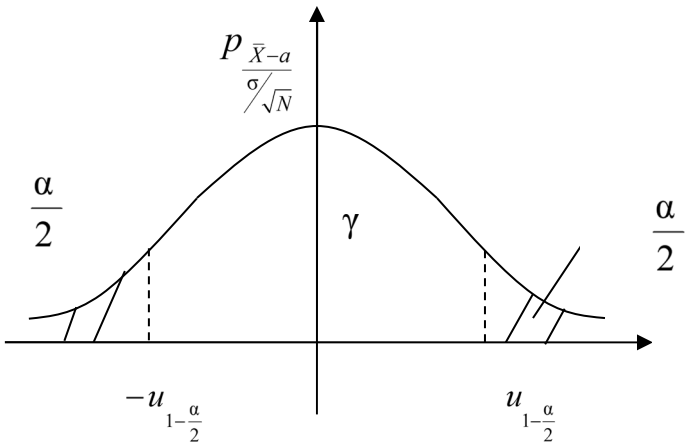
Квантилью порядка *p* распределения случайной величины ξ называется число x_p такое, что $P\{\xi \le x_p\} = p$.



Ещё x_p называют *нижней критической точкой* порядка p (или $p \cdot 100$ процентной точкой). *Верхняя критическая точка* порядка α – это квантиль $x_{1-\alpha}$.

Квантиль нормального стандартного распределения обозначается u_p. Очевидно, что $u_{0.5} = 0, u_0 = -\infty, u_1 = +\infty$.

Обозначим $\alpha = 1 - \gamma$, тогда из рисунка ясно положение границ интервала.



Таким образом, неизвестные границы найдены: $P\left\{-u_{1-\frac{\alpha}{2}} < \frac{\bar{X}-a}{\sigma/\sqrt{N}} < u_{1-\frac{\alpha}{2}}\right\} = \gamma.$

Преобразовав данное неравенство, получим

$$P\left\{\bar{X} - u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{N}} \le a \le \bar{X} + u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{N}}\right\} = \gamma.$$

Случай, когда дисперсия неизвестна. Заменим дисперсию её оценкой – статистикой s^2 и построим случайную величину:

$$\frac{\bar{X}-a}{S/\sqrt{N}}=\frac{\bar{X}-a}{\sigma/\sqrt{N}}\frac{1}{\sqrt{S^2/\sigma^2}}=\frac{\dfrac{\bar{X}-a}{\sigma/\sqrt{N}}}{\sqrt{[(N-1)S^2/\sigma^2]\cdot[1/(N-1)]}}.$$

По теореме Фишера последнее отношение равно $\dfrac{N(0,1)}{\sqrt{\chi^2_{N-1}/(N-1)}}\sim \mathrm{T}_{N-1}$,

т. е. имеет распределение Стьюдента с (N–1) степенями свободы. Далее аналогично случаю известной дисперсии получим:

$$P\left\{\bar{X}-t_{1-\frac{\alpha}{2},N-1}\frac{S}{\sqrt{N}}\le a\le \bar{X}+t_{1-\frac{\alpha}{2},N-1}\frac{S}{\sqrt{N}}\right\}=\gamma, \qquad (28)$$

где $t_{p,n}$ – квантиль порядка p распределения Стьюдента с n степенями свободы.

Пример 2.1 (продолжение примера 1.1)

По выборке 3, 7, 7, 0, 2, 1, 2, 3, 5, 4 найти доверительный интервал для математического ожидания нормальной случайной величины с вероятностью 0.95.

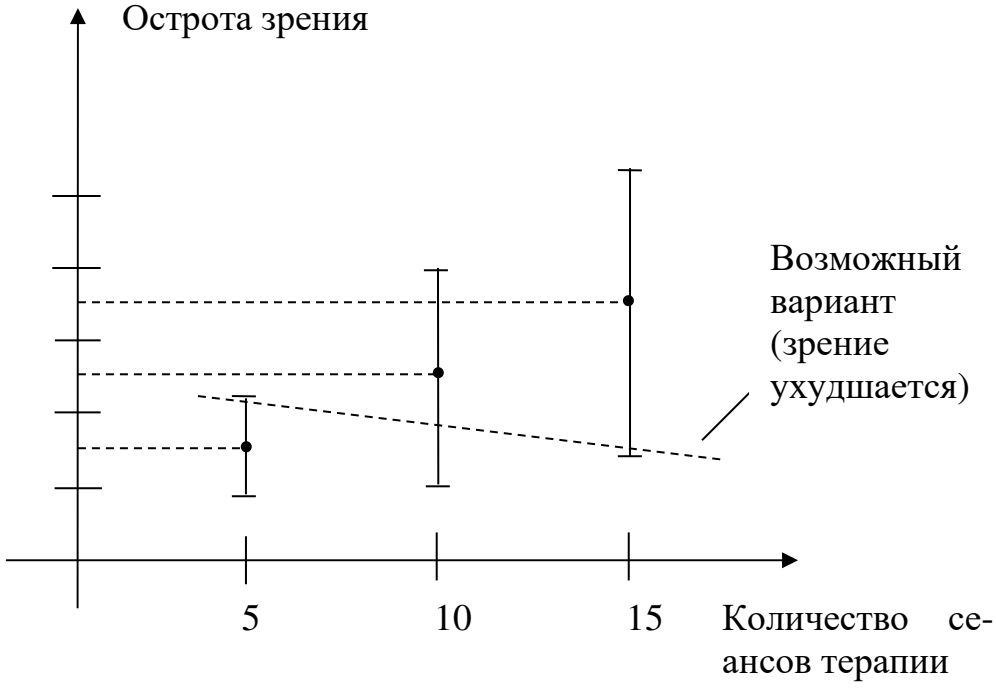
Из таблицы квантилей распределения Стьюдента находим: $t_{1-\frac{\alpha}{2},N-1}=t_{0.975,9}=2.26$.

$$\left(3.4-2.26\cdot\frac{2.37}{\sqrt{10}}\le a\le 3.4+2.26\cdot\frac{2.37}{\sqrt{10}}\right),\ \left(1.7\le a\le 5.1\right).$$

Истинное математическое ожидание находится в этих границах с вероятностью $\gamma=0.95$.

Пример 2.2

Исследуется острота зрения в зависимости от числа сеансов терапии. Важно при нанесении на график экспериментальных точек, полученных по нескольким наблюдениям, указывать для каждой точки доверительный интервал.



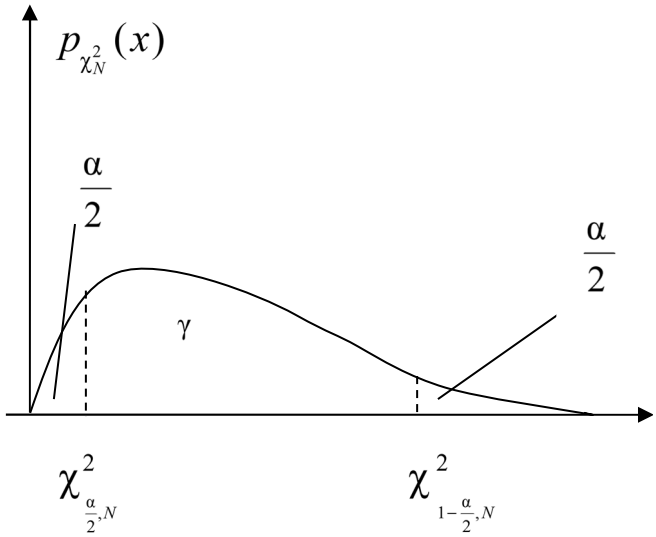
При построении доверительных интервалов мы предполагали нормальную распределённость наблюдений. Однако при больших N это ограничение можно снять, т. к. по центральной предельной теореме (ЦПТ) величина $\overline{X}$ всегда будет иметь примерно нормальное распределение (независимо от вида распределения наблюдений).

2.2. Доверительный интервал для дисперсии нормальной случайной величины

Случай, когда математическое ожидание известно. Имеем выборку (3). Определим

$$S_0^2 = \frac{1}{N}\sum_{n=1}^{N}\left(X^n - a\right)^2, \qquad (29)$$

тогда т. к. $\frac{X^n - a}{\sigma} \sim N(0,1)$, то $\frac{NS_0^2}{\sigma^2} = \sum_{n=1}^{N}\left(\frac{X^n - a}{\sigma}\right)^2 \sim \chi_N^2$. Поместим вероятность γ в середине (по площади) плотности распределения.



Таким образом, определились границы интервала

$P\left\{\chi^2_{\frac{\alpha}{2},N} \le \frac{NS_0^2}{\sigma^2} \le \chi^2_{1-\frac{\alpha}{2},N}\right\} = \gamma$, после преобразований

$$P\left\{\frac{NS_0^2}{\chi^2_{1-\frac{\alpha}{2},N}} \le \sigma^2 \le \frac{NS_0^2}{\chi^2_{\frac{\alpha}{2},N}}\right\} = \gamma.$$

Случай, когда математическое ожидание неизвестно. По теореме Фишера $\frac{(N-1)S^2}{\sigma^2} \sim \chi^2_{N-1}$, поэтому аналогично случаю известного математического ожидания

$$P\left\{\frac{(N-1)S^2}{\chi^2_{1-\frac{\alpha}{2},N-1}} \le \sigma^2 \le \frac{(N-1)S^2}{\chi^2_{\frac{\alpha}{2},N-1}}\right\} = \gamma. \qquad (30)$$

Пример 2.3 (продолжение примера 2.1)

По выборке 3, 7, 7, 0, 2, 1, 2, 3, 5, 4 найти доверительный интервал для дисперсии с доверительной вероятностью 0.9.

По таблице квантилей распределения «хи-квадрат»: $\chi^2_{0.05,\,9} = 3.32, \chi^2_{0.95,\,9} = 16.9$. Имеем по формуле (30) $\left(\frac{9\cdot 5.6}{16.9} \le \sigma^2 \le \frac{9\cdot 5.6}{3.32}\right)$, или $\left(2.9 \le \sigma^2 \le 15.2\right)$.

Поскольку $P\left\{W_1 \le \sigma^2 \le W_2\right\} = \gamma$ эквивалентно $P\left\{\sqrt{W_1} \le \sigma \le \sqrt{W_2}\right\} = \gamma$, последнее даёт доверительный интервал для стандартного отклонения. Для сравнения отметим, что $\sqrt{S^2}$ – корень квадратный из несмещённой оценки дисперсии σ^2, не является несмещённой оценкой σ.

2.3. Односторонние доверительные интервалы

Иногда бывает желательно сосредоточить доверительную вероятность в полубесконечном интервале, скажем, построить для дисперсии правосторонний интервал вида: $P\left\{W_1 \le \sigma^2 \le +\infty\right\}$. Анализируя вывод п. 2.2, легко догадаться, что сначала γ должна быть сосредоточена под левым крылом графика плотности распределения. После преобразования получим интервал

$$P\left\{\frac{(N-1)S^2}{\chi^2_{1-\alpha,N-1}} \le \sigma^2 \le +\infty\right\} = \gamma.$$

Пример 2.4 (продолжение примера 2.3)

Найти правосторонний доверительный интервал для дисперсии с вероятностью 0.9.

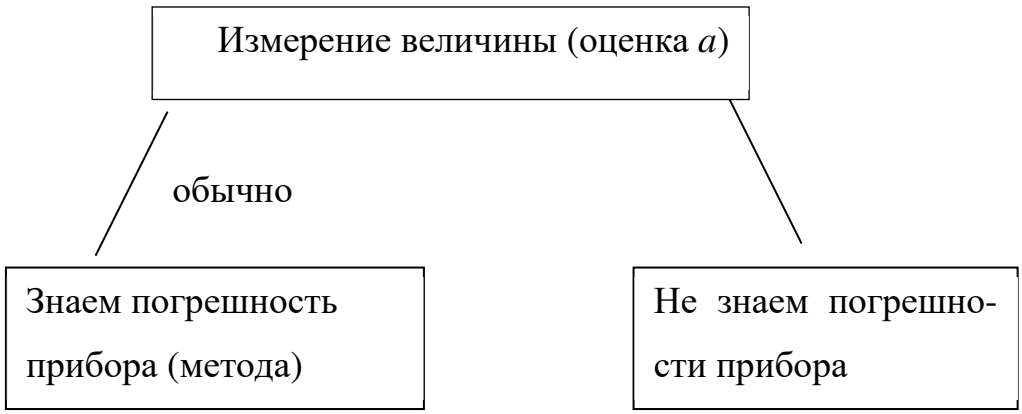
$\chi^2_{0.9,\,9} = 14.7$, и наша выборка даёт интервал: $(3.42 \le \sigma^2 \le +\infty)$.

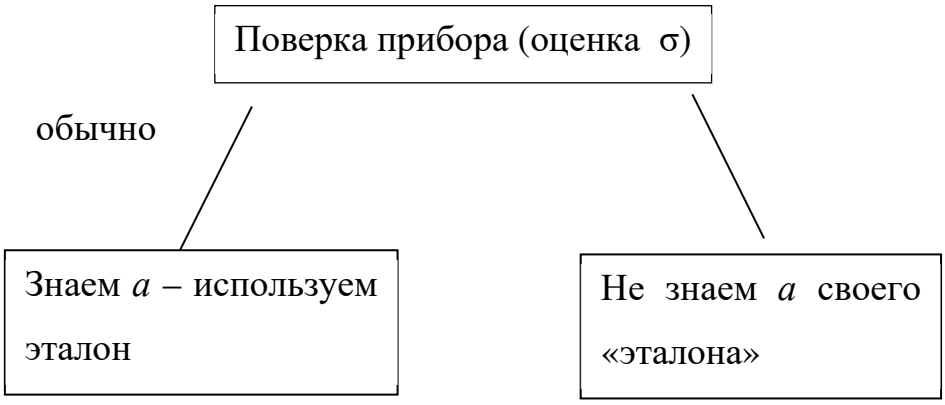
2.4. Решение двух основных задач в теории ошибок измерений

Пусть выборка (3) – результаты *N* независимых измерений постоянной величины *a*. В теории ошибок измерений считают, что результат *n*-го измерения есть измеряемая (истинная) величина, на которую накладывается E^n – случайная ошибка в *n*-м измерении (*Error*). $X^n = a + E^n$. Чаще всего предполагают свойства ошибок: $ME^n = 0$ (нет систематической ошибки), $DE^n = \sigma^2$. Так как $MX^n = M(a + E^n) = Ma + ME^n = Ma = a$, то результат измерения X^n есть несмещённая оценка измеряемой величины и среднеквадратическая ошибка этой оценки (результата измерения) $\sqrt{M(X^n - a)^2} = \sqrt{DE^n} = \sigma$ – погрешность прибора.

Измерения проводятся для решения двух типов задач: измерения величины (оценка истинного значения *a*) и поверки прибора (оценка его погрешности σ). Надёжными будут результаты, представленные интервальными оценками с высокой доверительной вероятностью. Возможные варианты этих задач приведены на рисунке.

Эти 4 случая рассмотрены в п. 2.1–2.2.





Упражнения

Упражнение 2.1

Из космоса с помощью прибора с погрешностью 4 м оценивается диаметр ракетной шахты. После 16 измерений получен результат 20 м. Можно ли утверждать с $p = 0.9$, что диаметр шахты не 22 м? А с $p = 0.99$?

Упражнение 2.2

Диаметр вала измерялся 9 раз. $\overline{x} = 30$, погрешность прибора неизвестна, несмещённая оценка дисперсии равна 9. Найти 90 и 95 % доверительные интервалы для диаметра. Найти то же, если погрешность прибора равна 3.

Упражнение 2.3

Производится поверка измерителя ёмкости конденсаторов. Эталона нет. Произвели 16 измерений, несмещённая оценка дисперсии равна 4.5 мкф^2. Какую погрешность присвоить прибору, если требуется 90 % надёжность?

Упражнение 2.4

На президентских выборах для победы в 1-м туре претенденту надо набрать более 50 % голосов избирателей. В 2000 г. после обработки 2000 бюллетеней за В. В. Путина было подано 1060 голосов. Построить 99 % доверительный интервал для доли сторонников В. В. Путина.

Упражнение 2.5

Некоторая величина измерялась двумя методами, для которых погрешности одного измерения равны соответственно 0.2 и 0.1. Первый метод после 5 измерений дал значение 8.74, второй после 3 дал 8.66. Какое значение является более вероятным? Чему равны погрешности результатов? Каков общий результат и какова его погрешность?

3. Статистическая проверка гипотез

Все мы бываем свидетелями явлений, которые, как нам кажется, значимо свидетельствуют о чём-то необычном, или сами ставим эксперимент, чтобы обнаружить это необычное. Пусть есть гипотеза H_0 и свершившееся событие (наблюдение) *A*, которое мы хотим истолковать как опровержение H_0 в пользу альтернативной гипотезы H_1. Насколько значимо (определённо, достоверно, надёжно) *A* свидетельствует против H_0?

Пример 3.1

Вы оказались свидетелем события: молодой человек не уступил женщине место в трамвае – событие *A*.

H_0: основная гипотеза – он воспитанный;

H_1: альтернативная гипотеза – он невоспитанный.

Мерой значимости события (наблюдения, данных) *A* можно принять редкость (малую вероятность) получить *A* при условии справедливости H_0:

$$sl(A) = P(A \mid H_0) \tag{31}$$

– *уровень значимости* (*significance level*) данных *A*. Если вероятность получить данные *A* при условии справедливости H_0 мала, $sl < 0.01$, а мы их получили, то означает ли это нашу необыкновенную везучесть? Нет, надо отказываться от H_0, поскольку при допущении H_0 вероятность рассчиталась малой.

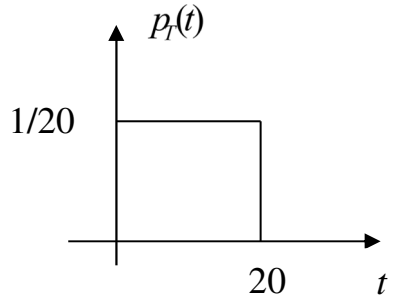
В нашем примере вероятность (31) будет мала и связана с возможной болезнью, невнимательностью (редкие события у воспитанный людей).

Пример 3.2

Вы прождали $t = 19$ мин трамвай с объявленным интервалом движения 20 мин. Гипотеза H_0: $b = 20$ мин (интервал движения трамваев 20 мин). Гипотеза H_1: $b > 20$ мин (сегодня интервал составляет больше 20 мин).

$$sl(t) = P\{T \geq 19 \mid H_0\} = P\{19 \leq T \leq 20 \mid H_0\} = \int_{19}^{20} p_T(t)dt = \frac{1}{20}\int_{19}^{20} dt = \frac{1}{20} = 0.05.$$

В этом расчёте замена события $T = 19$ на $T \geq 19$ может только увеличить вероятность, но если она и после этого мала, тем более следует отказываться от H_0. Случайное время ожидания *T* имеет равномерное распределение (см. рисунок).



Т. о., вероятность прождать 19 мин или больше при должном интервале движения весьма мала.

Традиционно значения уровня значимости интерпретируются следующим образом.

sl Интерпретация

> 0.1….....Данные согласуются с H_0

≈ 0.05……Возможна значимость данных. Есть сомнения в истинности гипотезы H_0

≈ 0.02……Значимость. Довольно сильный довод против H_0

< 0.01……..Высокая значимость. H_0 почти наверняка не подтверждается

Пример 3.3 [10]

В 1980–81 гг. в одной английской деревне большую долю новорождённых составляли девочки (25 из 35). Это сочли необычным явлением, в прессе давались различные объяснения. Одно из них: в окрестностях заработала новая каменоломня и пыль из неё, содержащая кадмий, воздействовала на организмы родите-

лей и повышала вероятность рождения девочек p. Были противники такого объяснения, утверждавшие, что всё в порядке: $H_0 : p = 0.48$ (как и в целом по стране, а 25 из 36 это, мол, дело случая). Первые же утверждали: H_1:

$p > 0.48$.

Вычислим

$$sl(A) = P\{\mu_{35} \geq 25 \mid H_0\} = 1 - P\{\mu_{35} \leq 24 \mid H_0\} = 1 - \Phi^{\text{Bin}}_{35,\,0.48}(24) = 0.004.$$

μ_{35} – число успехов (рождения девочек) из 35 испытаний. sl мал, данные значимо свидетельствуют против H_0, и её надо отвергать, следовательно, в деревне существует причина высокой рождаемости девочек.

Т. о., основная идея проверки гипотез состоит в следующем: гипотеза отвергается, если вероятность при ней получить наблюдённые данные мала.

3.1. Общая схема проверки гипотез

Схема проверки состоит в следующем.

1. Формулируется основная гипотеза H_0 и конкурирующая (альтернативная) H_1. Обычно эксперимент бывает направлен на опровержение H_0 в пользу H_1.

2. Проводятся (или уже имеются) наблюдения – выборка $x^1, x^2, ..., x^N = \{x^n\}$.

3. Обычно выборка используется не непосредственно, а для вычисления некоторой статистики $k = k(\{x^n\})$ – статистики критерия, которая имеет тесную смысловую связь с гипотезами H_0 и H_1.

4. Назначается область значимости $G_{1k} = \{k :$ содержит наблюдённое k и все значения k, ещё более соответствующие гипотезе H_1 и менее – $H_0\}$.

5. Вычисляется уровень значимости данных

$$sl(\{x^n\}) = sl(k) = P\{K \in G_{1k} \mid H_0\},$$

где K – случайное значение статистики критерия. $K = k(\{X^n\})$.

6. Интерпретация полученного sl такова: если sl мало, то H_0 отвергается, в противном случае говорят, что наблюдения согласуются с H_0.

3.2. Проверка гипотезы о математическом ожидании нормальной случайной величины

Пример 3.4

В продаже появилось лекарство, обещающее снижение массы тела в среднем на 6 кг. Дали его $N = 10$ испытуемым и проследили, насколько уменьшилась их масса. Результаты: 3, 7, 7, 0, 2, 1, 2, 3, 5, 4. Вам показалось, что снижение массы мало, и вы проверяете гипотезу $H_0: M\xi = a_0 = 6$ против альтернативы, на которую намекают данные $H_1: M\xi < a_0$. Статистикой, по которой будем судить, выберем $\overline{X}$, в область значимости возьмём значения: наблюдённое $\overline{x}$ и все меньшие. В таком случае $sl = P\{\overline{X} \le \overline{x} \mid H_0\}$. Однако при неизвестной дисперсии распределение $\overline{X}$ неизвестно, поэтому воспользуемся статистикой $\dfrac{\overline{X} - a_0}{S/\sqrt{N}}$.

$$sl = P\left\{\frac{\overline{X} - a_0}{S/\sqrt{N}} \le \frac{\overline{x} - a_0}{s/\sqrt{N}} \mid H_0\right\} = \Phi_{N-1}^{\text{Ст}}\left(\frac{\overline{x} - a_0}{s/\sqrt{N}}\right).$$

$$sl = \Phi_{N-1}^{\text{Ст}}\left(\frac{\overline{x} - a_0}{s/\sqrt{N}}\right). \qquad (32)$$

В нашем примере:

$$sl = \Phi_9^{\text{Ст}}\left(\frac{3.4-6}{2.37/\sqrt{10}}\right) = \Phi_9^{\text{Ст}}(-3.474) = 1 - \Phi_9^{\text{Ст}}(3.474) = 1 - 0.9965 = 0.0035.$$

Вероятность получить наблюдённые результаты при лекарстве, снижающем массу тела в среднем на 6 кг, очень мала, следовательно, гипотезе H_0 говорим «нет» и производитель говорит неправду.

Иногда заранее назначают граничное значение α уровня значимости *sl* для различения гипотез: *критерий уровня* α *проверки гипотезы* H_0 против альтернативы H_1, т. е вычисляется уровень значимости и сравнивается с α:

$$sl \overset{>}{\underset{<}{}} \alpha \Rightarrow \begin{matrix} H_0 \\ H_1 \end{matrix}. \qquad (33)$$

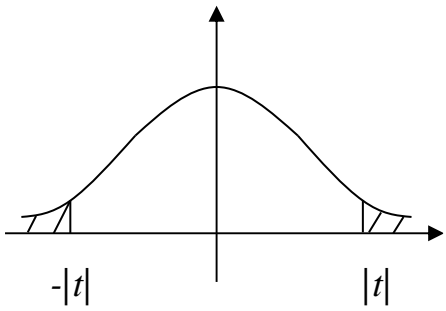
В нашем примере если задать α = 0.05, то получим вывод в пользу гипотезы H_1, если задать α = 0.001, то вывод в пользу гипотезы H_0. О смысле выбора того или иного α см. в п. 3.8.

3.3. Критерий равенства двух математических ожиданий

Пусть есть выборка $x^1, x^2, ..., x^{N_1}$ из нормального распределения $N(a_1, \sigma)$ и выборка $y^1, y^2, ..., y^{N_2}$ из распределения $N(a_2, \sigma)$, т. е. дисперсия одинакова (но неизвестна). Проверяется гипотеза $H_1 : a_1 = a_2$. Альтернативная гипотеза $H_1 : a_1 \neq a_2$. Статистика критерия $\dfrac{\overline{X} - \overline{Y}}{S/\sqrt{\dfrac{N_1 N_2}{N_1 + N_2}}} \sim T_{N_1+N_2-2}$ имеет

(при H_0) распределение Стьюдента с $N_1 + N_2 - 2$ степенями свободы, где $S^2 = \dfrac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{N_1 + N_2 - 2}$.

Уровень значимости данных $t = \dfrac{\overline{x} - \overline{y}}{s \Big/ \sqrt{\dfrac{N_1 N_2}{N_1 + N_2}}}$, т. к. область значимости – большие по модулю значения величины $T_{N_1+N_2-2}$ (см. рисунок),



$sl = P\left\{ |T_{N_1+N_2-2}| \geq |t| \right\} = 2P\left\{ T_{N_1+N_2-2} \geq |t| \right\} = 2\left[1 - P\left\{ T_{N_1+N_2-2} \leq |t| \right\} \right]$. Переходя к функции распределения, получаем

$$sl = 2\left[1 - \Phi^{\text{Ст}}_{N_1+N_2-2}\left(\left| \frac{\overline{x} - \overline{y}}{s \Big/ \sqrt{\dfrac{N_1 N_2}{N_1 + N_2}}} \right| \right) \right]. \qquad (34)$$

Проверка гипотезы H_0: $a_2 - a_1 = c$ при альтернативе H_1: $a_2 - a_1 \neq c$, т. к. это эквивалентно H_0: $a_1 = a_2 - c = a_2'$ при альтернативе H_1: $a_1 \neq a_2'$, проводится по формуле (34) с заменой: $\overline{x} - \overline{y} \to \overline{x} - (\overline{y} - c)$.

3.4. Проверка гипотезы о коэффициенте корреляции

Напомним, что если $\rho(\xi,\eta)\neq 0$, то ξ и η зависимы. Поэтому представляют интерес гипотезы $\begin{matrix} H_0:\rho=0, \\ H_1:\rho\neq 0. \end{matrix}$ Для выбора из них нужны парные наблюдения:

$$\begin{pmatrix} x^1 \\ y^1 \end{pmatrix},\begin{pmatrix} x^2 \\ y^2 \end{pmatrix}\ldots\begin{pmatrix} x^N \\ y^N \end{pmatrix}. \qquad (35)$$

Оказывается, в случае нормального распределения при H_0 величина $\dfrac{\hat{R}\sqrt{N-2}}{\sqrt{1-\hat{R}^2}}\sim T_{N-2}$, т. е. имеет распределение Стьюдента. Здесь $\hat{R}$ – оценка коэффициента корреляции по формуле (16). Область значимости – большие по модулю значения этой величины, поэтому получаем

$$sl=P\left\{|T_{N-2}|\geq\left|\frac{\hat{\rho}\sqrt{N-2}}{\sqrt{1-\hat{\rho}^2}}\right|\right\}=2\left[1-\Phi_{N-2}^{\text{Ст}}\left(\left|\frac{\hat{\rho}\sqrt{N-2}}{\sqrt{1-\hat{\rho}^2}}\right|\right)\right],$$

где $\hat{\rho}$ – наблюдённое значение $\hat{R}$, рассчитанное по данным (35).

3.5. Критерий равенства двух дисперсий

Пусть $x^1, x^2, ..., x^{N_1}$ – выборка из нормального распределения $N(a_1,\sigma_1)$, а $y^1, y^2, ..., y^{N_2}$ – выборка из нормального распределения $N(a_2,\sigma_2)$. Пусть a_1, a_2 известны. Проверяются гипотезы $\begin{matrix} H_0:\sigma_1=\sigma_2, \\ H_1:\sigma_1>\sigma_2. \end{matrix}$

Рассмотрим отношение $\dfrac{S_{01}^2}{S_{02}^2}$, где S_0^2 определено по (29). При справедливости H_1 вероятны большие значения этого отношения, поэтому

$sl = P\left\{ \frac{S_{01}^2}{S_{02}^2} \geq \frac{s_{01}^2}{s_{02}^2} \mid H_0 \right\}$. Поскольку при $H_0 : \sigma_1 = \sigma_2 = \sigma$:

$\frac{N_1 S_{01}^2}{\sigma^2} \sim \chi_{N_1}^2$, $\frac{N_2 S_{02}^2}{\sigma^2} \sim \chi_{N_2}^2$, постольку $\frac{S_{01}^2}{S_{02}^2} = \frac{\chi_{N_1}^2 / N_1}{\chi_{N_2}^2 / N_2} \sim F_{N_1,N_2}$. В результате получаем

$$sl = P\left\{ F_{N_1,N_2} \geq \frac{s_{01}^2}{s_{02}^2} \right\} = 1 - \Phi_{N_1,N_2}^F \left(\frac{s_{01}^2}{s_{02}^2} \right).$$

Если a_1, a_2 неизвестны, то

$$sl = P\left\{ F_{N_1-1,N_2-1} \geq \frac{s_1^2}{s_2^2} \right\} = 1 - \Phi_{N_1-1,N_2-1}^F \left(\frac{s_1^2}{s_2^2} \right). \qquad (36)$$

Пример 3.5

До наладки станка была проверена идентичность изготовления 10 втулок и найдена оценка дисперсии диаметра s^2 = 9.6 мкм^2. После наладки станка подверглись контролю ещё 15 втулок, и нашли s^2 = 5.7 мкм^2. Можно ли считать, что разброс диаметра уменьшился?

$H_0 : \sigma_1 = \sigma_2$, $H_1 : \sigma_1 > \sigma_2$, $sl = 1 - \Phi_{9,14}^F \left(\frac{9.6}{5.7} \right) = 1 - 0.815 = 0.185$. Данные согласуются с гипотезой H_0, т. е. ничего не изменилось – наладка станка не улучшила ситуацию.

Если рассматривать двустороннюю альтернативу $H_1 : \sigma_1 \neq \sigma_2$, то

$$sl = 2 \cdot [1 - \Phi_{N_1-1,N_2-1}^F \left(\frac{s_1^2}{s_2^2} \right)], \qquad (37)$$

где первые степени свободы и значение в числителе всегда относятся к большей оценке дисперсии.

3.6. Связь проверки гипотез с интервальным оцениванием

Задача проверки простой гипотезы о параметре θ: $H_0 : \theta = \theta_0$ – и построение доверительного интервала для этого параметра тесно связаны друг с другом. Если доверительный интервал для θ с доверительной вероятностью $\gamma = 1 - \alpha$ накрывает θ_0, то гипотеза $H_0 : \theta = \theta_0$ принимается на уровне значимости α, и наоборот (не накрывает – не принимается).

Пример 3.6

Дисперсия зарплаты рабочих в цехе была 150 р.^2. После введения КТУ – коэффициента трудового участия – опросили $N = 20$ рабочих и вычислили $\hat{\sigma}^2 = s^2$ = 300 р.^2. Противники КТУ говорят, что полученное увеличение – дело случая, а дисперсия зарплаты не изменилась. Проверим критерием уровня $\alpha = 0.01$ гипотезу: $H_0 : \sigma^2 = \sigma_0^2 = 150$ против альтернативы $H_1 : \sigma^2 \neq \sigma_0^2$. Доверительный интервал $\left(\dfrac{(N-1)s^2}{\chi^2_{1-\frac{\alpha}{2}, N-1}} \leq \sigma^2 \leq \dfrac{(N-1)s^2}{\chi^2_{\frac{\alpha}{2}, N-1}} \right) =$

$= \left(147.6 \leq \sigma^2 \leq 833.3\right)$ накрывает дисперсию в 150 р.^2, значит, нет основания отвергать H_0 на уровне значимости $\alpha = 0.01$. Противники КТУ оказались правы (см. упр. 3.9).

3.7. Мощность критерия

Если параметрическая гипотеза имеет вид $H : \theta = \theta_0$, т. е. вполне определяет распределение, то она называется *простой*. Если гипотеза имеет вид:

$H:\theta\in M_\theta$, где M_θ – некоторое множество значений параметра (в частности, $\theta>\theta_0$, $\theta<\theta_0$, $\theta\neq\theta_0$), то такая гипотеза *сложная*.

Хороший критерий (33) различения двух гипотез должен при фактической ситуации H_0 говорить «H_0», т. е. $P\left\{\overset{\text{сказать}}{H_0}\mid H_0\right\}=P\{SL>\alpha\mid H_0\}$ должна быть велика. Здесь SL – случайная величина, определяемая случайными данными (выборкой). Эту вероятность всегда можно увеличить, понижая α. Однако хороший критерий должен при $(H_1^{'}:\theta=\theta_1)\in H_1$ говорить «H_1», т. е. вероятность $P\left\{\overset{\text{сказать}}{H_1}\mid H_1^{'}\right\}=P\{SL<\alpha\mid H_1^{'}\}$ должна быть велика. Этого всегда можно добиться, повышая α. Но в таком случае при H_0 потеряем способность говорить «H_0». Поэтому, зафиксировав α (скажем, $\alpha=0{,}01$), чтобы обеспечить узнавание H_0, желательно обеспечить большую вероятность $P\{SL<\alpha\mid H_1^{'}\}$ – мощность критерия.

Функция мощности критерия уровня α:

$$w(\alpha,\theta_1)=P\left\{\overset{\text{сказать}}{H_1}\mid H_1^{'}:\theta=\theta_1\right\}=P\{SL<\alpha\mid H_1^{'}\}. \qquad (38)$$

3.8. Мощность критерия проверки гипотезы о математическом ожидании нормальной случайной величины

Случай, когда дисперсия известна. Пусть $\xi\sim N(a,\sigma)$, σ – известно.

Проверяем гипотезу о параметре $\theta=M\xi=a$: $\begin{matrix}H_0:a=a_0,\\ H_1:a<a_0.\end{matrix}$ Аналогично рассуждениям в примере 3.4 получим выражение:

$$sl = \Phi\left(\frac{\overline{x} - a_0}{\sigma/\sqrt{N}}\right), \tag{39}$$

где Φ(.) – функция распределения нормального стандартного распределения. Найдём функцию мощности этого критерия уровня α по формуле (38):

$$w(\alpha, a_1) = P\left\{\Phi\left(\frac{\overline{X} - a_0}{\sigma/\sqrt{N}}\right) < \alpha \,|\, a_1\right\} = P\left\{\Phi^{-1}\left[\Phi\left(\frac{\overline{X} - a_0}{\sigma/\sqrt{N}}\right)\right] < u_\alpha \,|\, a_1\right\} =$$

$$= P\left\{\frac{\overline{X} - a_0}{\sigma/\sqrt{N}} < u_\alpha \,|\, a_1\right\} = P\left\{\frac{\overline{X} - a_1 + a_1 - a_0}{\sigma/\sqrt{N}} < u_\alpha \,|\, a_1\right\} =$$

$$= P\left\{\frac{\overline{X} - a_1}{\sigma/\sqrt{N}} < u_\alpha + \frac{a_0 - a_1}{\sigma/\sqrt{N}} \,|\, a_1\right\}.$$

Окончательное выражение:

$$w(\alpha, a_1) = \Phi\left(u_\alpha + \frac{a_0 - a_1}{\sigma/\sqrt{N}}\right). \tag{40}$$

В гл. 4 с помощью леммы Неймана–Пирсона будет показано, что критерий (39) является оптимальным, т. е. имеет максимальную мощность среди всех критериев уровня α.

Критерий (39) является *равномерно наиболее мощным*, поскольку он не зависит от $\theta_1 = a_1$.

Напомним, что мощность $w(\alpha, \theta_1)$ показывает, с какой вероятностью реальная ситуация θ_1 будет зафиксирована критерием уровня α, т. е. данные выборки объёма N окажутся значимыми. Можно поставить вопрос по-другому: сколь велика должна быть выборка, чтобы ситуация θ_1 фиксировалась критерием уровня α с

заданной (высокой) вероятностью w? При этом в задаче проверки гипотез $\begin{matrix} H_0 : a = a_0, \\ H_1 : a > a_0 \end{matrix}$ о математическом ожидании нормальной случайной величины при известной дисперсии получается формула

$$N = \sigma^2 \frac{(u_{1-\alpha} + u_w)^2}{(a_1 - a_0)^2}. \tag{41}$$

Случай, когда дисперсия неизвестна. Критерий проверки гипотез $\begin{matrix} H_0 : a = a_0, \\ H_1 : a < a_0 \end{matrix}$ имеет вид (32)–(33). Проводя вычисления подобно тем, что привели к (40), получим

$$w(\alpha, a_1) = \Phi^{\text{Ст}}_{N-1,\, \frac{a_1 - a_0}{\sigma_1 / \sqrt{N}}} \left(t_{\alpha,\, N-1} \right), \tag{42}$$

где $\Phi^{\text{Ст}}_{N,\,c}(.)$ – функция распределения случайной величины $T_N^c = \dfrac{N(0,1) + c}{\sqrt{\chi_N^2 / N}}$ (с нецентральным распределением Стьюдента с N степенями свободы и параметром нецентральности c), σ_1 – предполагаемое в H_1 стандартное отклонение случайной величины.

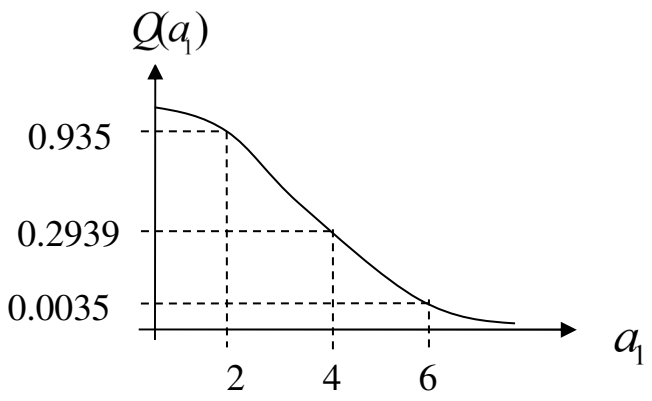
Определим *функцию чувствительности* критерия: $Q\left(\theta_1\right) = w\left(sl, \theta_1\right)$ – вероятность получить полученные фактически (sl) доводы против H_0 при $H_1^{'} : \theta = \theta_1$. $Q\left(\theta_1\right)$ позволяет указать (подобрать) те значения θ_1, при которых полученные доводы вероятны ($Q \geq 0.8$), т. е. высказаться о характере $H_1^{'}$.

Пример 3.7 (продолжение примера 3.4)

В примере 3.4 получили $sl = 0.0035$.

Найдём $Q(4) = w(0.0035, 4) = \Phi^{\text{Ст}}_{9,\frac{4-6}{2.37/\sqrt{10}}}(-3.474) = 0.2939$, т. е. если бы среднее снижение массы тела вследствие приёма лекарства было 4 кг, то вероятность получить $sl = 0.0035$ была бы невысока (порядка 0.3), значит, вряд ли среднее снижение массы есть 4 кг! $Q(2) = w(0.0035, 2) = 0.935$, т. е. только при Mξ < ~ 2 имеется большая вероятность получить наблюдённые результаты, проверяя гипотезу H_0.

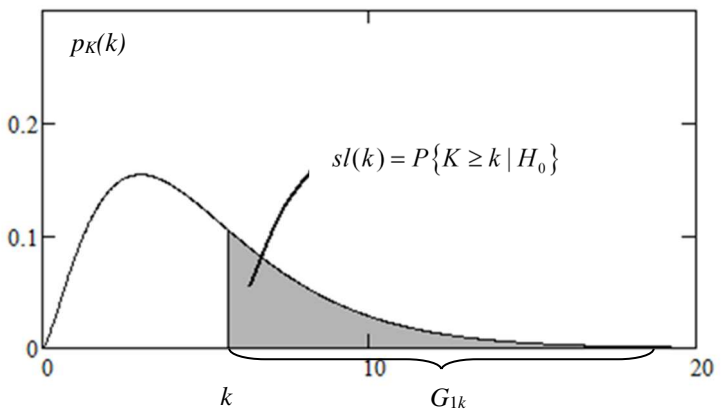
В нашем примере график функции чувствительности критерия от a_1 – среднего снижения массы тела вследствие приёма лекарства, имеет следующий вид.



Смысл величины α в критерии. Рассмотрим вероятность следующей ошибки

$$\alpha_{10} = P\{\text{Сказать } H_1 \mid \text{Фактически } H_0\} = P\{sl(K) \le \alpha \mid H_0\}. \qquad (43)$$

На следующем рисунке видно, что функция *sl*(.) монотонно убывает (при односторонней альтернативе).



Поэтому применение к обеим частям неравенства (43) обратного к *sl*(.) преобразования даёт:

$$P\left\{K \geq sl^{-1}(\alpha) \mid H_0\right\} = sl[sl^{-1}(\alpha)] = \alpha .$$

Таким образом, получили $\alpha = \alpha_{10}$ – вероятность *ошибки I рода*: сказать H_1, когда верна H_0. Вероятность *ошибки II рода*: сказать H_0, когда верна H_1': $\alpha_{01'} = 1 - \alpha_{11'} = 1 - w\left(\alpha, \theta_1\right)$.

3.9. Различные критерии на одних и тех же данных

Пример 3.8

Рассмотрим два критерия проверки гипотез $\begin{matrix} H_0 : a = a_0, \\ H_1 : a > a_0 \end{matrix}$ о математическом ожидании нормальной случайной величины при известной дисперсии: стандартный, параметры которого (аналогично формулам (39), (40)) составляют: $sl = 1 - \Phi\left(\dfrac{\overline{x} - a_0}{\sigma / \sqrt{N}}\right)$ $w(\alpha, a_1) = 1 - \Phi\left\{u_{1-\alpha} - \dfrac{a_1 - a_0}{\sigma / \sqrt{N}}\right\}$, и «новый», предписывающий из имеющихся N наблюдений использовать только первое и последнее:

$$\tilde{x} = \frac{x^1 + x^N}{2},$$

$$\widetilde{sl} = P\left\{\widetilde{X} \geq \tilde{x} \mid H_0\right\} = 1 - P\left\{\widetilde{X} \leq \tilde{x} \mid H_0\right\} = 1 - P\left\{\frac{\widetilde{X} - a_0}{\sigma/\sqrt{2}} \leq \frac{\tilde{x} - a_0}{\sigma/\sqrt{2}}\right\} =$$

$$= 1 - \Phi\left(\frac{\tilde{x} - a_0}{\sigma/\sqrt{2}}\right).$$

Если $\overline{x} \approx \tilde{x}$, что вполне может быть, то $\widetilde{sl} > sl$, т. е. стандартный критерий будет показывать значимость данных выше. Но для некоторых данных ситуация может быть и обратной. Мощность «нового» критерия меньше:

$$\widetilde{w}(\alpha, a_1) = P\left\{\widetilde{SL} < \alpha \mid a_1\right\} = 1 - \Phi\left\{u_{1-\alpha} - \frac{a_1 - a_0}{\sigma/\sqrt{2}}\right\}, \widetilde{w}(\alpha, a_1) < w(\alpha, a_1).$$

Эти два (подобные по сути) критерия одного уровня α имеют различную мощность, хотя на одних и тех же данных могут давать один и тот же уровень значимости.

Если рассматривать критерии проверки одних и тех же гипотез на различных по типу статистиках $K = k(\left\{X^n\right\})$, то они будут улавливать различные отклонения в данных от H_0. Т. о., чтобы убедиться в справедливости H_0, нужно, чтобы все критерии ей не противоречили. Чтобы отвергнуть H_0, достаточно отвергнуть одним критерием.

3.10. Критерии при сложной основной гипотезе

Пусть конкурируют гипотезы: $\begin{matrix} H_0 : M\xi \geq a_0, \\ H_1 : M\xi < a_0. \end{matrix}$ Если считать, как обычно, $sl = P\left\{\overline{X} \leq \overline{x} \mid H_0\right\}$, то поскольку H_0 не определена, то sl не определён. Единственное, что можно в таком случае, так это рассчитать конкретный уровень значимости при граничном значении математического ожидания $\widetilde{sl} = P\left\{\overline{X} \leq \overline{x} \mid \widetilde{H_0} : M\xi = a_0\right\}$ и применить следующий критерий уровня α:

$$\widetilde{sl} \underset{<}{>} \alpha \Rightarrow \begin{matrix} H_0 \\ H_1 \end{matrix}. \qquad (44)$$

Смысл α здесь состоит в следующем: рассмотрим

$$\alpha_{\tilde{1}0} = P\{\text{сказать } H_1 \text{ по } \widetilde{SL} \mid \text{верна } H_0\} = P\{\widetilde{SL} < \alpha \mid H_0\} \leq P\{SL < \alpha \mid H_0\} = \alpha.$$

Использовали соотношение $SL \leq \widetilde{SL}$ (т. к. при $M\xi$, больших a_0, вероятность получить $\overline{X} \leq \overline{x}$ меньше). Значит, α при сложной основной гипотезе имеет смысл максимальной вероятности ошибки *I* рода при применении критерия (44) уровня α.

Мощность критерия (44):

$w(\alpha, H_1^{'}) = w(\alpha, a_1) = P\left\{\overset{\text{сказать}}{H_1} \mid H_1^{'}\right\} = P\left\{\widetilde{SL} < \alpha \mid H_1^{'}\right\}$ – по старой формуле,

т. е. мощность не зависит от сложности гипотезы H_0.

В связи с этим в пакете программ *STATISTICA* компании *StatSoft, Inc.* в модуле *Power Analysis* при расчёте мощности рассматриваются случаи: $\begin{matrix} H_0 : M\xi \geq a, \\ H_1 : M\xi < a, \end{matrix}$

$\begin{matrix} H_0 : M\xi \leq a, \\ H_1 : M\xi > a \end{matrix}$ со сложной основной гипотезой и $\begin{matrix} H_0 : M\xi = a, \\ H_1 : M\xi \neq a. \end{matrix}$

В последнем случае α равен точно α_{10}.

Упражнения

Упражнение 3.1

Количество несоответствующих деталей в большой партии не должно превышать 4 %. В результате проверки 100 деталей из этой партии обнаружено 9 несоответствующих. Можно ли считать, что процент несоответствующих деталей превосходит допустимый при $\alpha = 0.01$?

Упражнение 3.2

При 600 подбрасываниях игральной кости шестёрка появилась 75 раз. Верна ли гипотеза о том, что вероятность появления шестёрки меньше, чем 1/6, если $\alpha = 0.01$?

Упражнение 3.3

Производится аттестация по настроенности процесса производства резисторов. Номинал резисторов равен 10 кОм. Дисперсия процесса неизвестна. После проверки 16 резисторов получено $\overline{x} = 8.3$ кОм, несмещённая оценка дисперсии $= 6.25$ кОм^2. Решить вопрос об аттестации процесса на уровне $\alpha = 0.01$.

Упражнение 3.4

Два штурмана определили пеленг цели по нескольким замерам, используя различные пеленгаторы. Результаты замеров первого штурмана: число замеров 4, $\overline{x} = 70.2^0$, $s = 0.5^0$; второго: число замеров 9, $\overline{x} = 71.5^0$, $s = 1^0$. Это одна цель или две? $\alpha = 0.05$.

Упражнение 3.5

Найти уровень значимости данных в упр. 1.8. Что если бы такая же оценка была получена при $N = 10$?

Упражнение 3.6

При исследовании стабилизатора напряжения самолёта на стенде проведено 9 независимых испытаний и получена несмещённая оценка дисперсии выходного напряжения, равная 0.08 $В^2$. В полёте проведено ещё 15 испытаний, в результате которых оценка дисперсии выходного напряжения оказалась равной 0.13 $В^2$. Есть ли основания полагать, что факторы, воздействующие на стабилизатор в полёте, оказывают существенное влияние на его точность? Принять $\alpha = 0.1$.

Упражнение 3.7

Вывести формулу (37).

Упражнение 3.8

Для определённой процедуры проверки коэффициента трения шины по мокрому асфальту утверждается, что дисперсия результатов измерения этого коэффициента составляет 0.1. Выборочное значение дисперсии, вычисленное по результатам 25 измерений коэффициента трения, оказалось равным 0.2. При $\alpha = 0.1$ проверить гипотезу о том, что дисперсия результатов измерений коэффициента трения составляет 0.1 против двусторонней альтернативы.

Упражнение 3.9

Заново рассмотреть данные примера 3.6, выдвинув более подходящую альтернативу.

Упражнение 3.10

Вывести формулу (41).

Упражнение 3.11

Из продукции автомата, выпускающего болты с номинальным значением диаметра 40 мм, была взята выборка болтов в количестве 20 шт. Выборочное среднее диаметра оказалось 40.459 мм. Результаты предыдущих измерений дают основание предполагать, что действительные диаметры болтов образуют нормально распределённую совокупность с дисперсией 1 мм^2. Можно ли утверждать, что диаметр продукции автомата не имеет положительного смещения по отношению к номинальному значению?

1. Построить критерий.
2. Каков уровень значимости данных *sl*?
3. Принять решение на уровне $\alpha = 0.01$.
4. Вычислить мощность этого критерия при фактическом среднем диаметра 40.2, 40.6, 41.0, 40 мм. Построить график функции мощности от a_1. Рассмотреть также случай $\alpha = 0.1$.
5. Построить функцию чувствительности. Дать интерпретацию её значений.
6. Каков должен быть объём выборки, чтобы критерий уровня $\alpha = 0.01$ фиксировал положительное смещение на 0.6 мм с вероятностью 0.9?
7. Построить график мощности из п. 4 данного упражнения как функцию от N.

Упражнение 3.12

Вывести формулу (42) для мощности критерия проверки гипотезы о математическом ожидании нормальной случайной величины при неизвестной дисперсии.

Упражнение 3.13

Придумать простой критерий с малой вероятностью ошибки *I* рода (равной 0.01), но плохой мощностью (взять колоду из 100 карточек, на одной написать…). Какова мощность? Придумать аналогичный критерий с противоположными свойствами.

Упражнение 3.14

Каково распределение *SL* при справедливости H_0?

4. Связь проверки гипотез с распознаванием образов

Необходимые сведения о постановке задач распознавания образов см., например, [5].

4.1. Построение моделей распознавания без информации об априорных вероятностях. Лемма Неймана–Пирсона

Пусть априорные вероятности $p_1, \ldots, p_Q$ классов образов неизвестны. Обозначим α_{kq} – вероятность (ошибки) отнести образ к классу k, когда он был из класса q. Пусть есть всего 2 класса. Применим метод последовательных уступок в многокритериальной задаче $\begin{matrix} \alpha_{12} \to \min, \\ \alpha_{21} \to \min, \end{matrix}$ считая первый критерий важнейшим. Отсюда следует определение: оптимальное правило распознавания реализует минимум вероятности ошибки одного вида при заданной вероятности ошибки другого вида –

$$\begin{aligned} &\alpha_{21} \to \min, \\ &\alpha_{12} = \alpha = \text{const}. \end{aligned} \tag{45}$$

Имея в виду второе равенство в (45), говорят, что это правило (критерий) уровня α.

Лемма (Неймана–Пирсона)

Оптимальным (наиболее мощным) критерием уровня α является критерий отношения правдоподобий.

Доказательство. Для решения задачи условной оптимизации (45) применим метод неопределённых множителей Лагранжа.

$$L = \alpha_{21}(A_1) + \lambda\alpha_{12}(A_1) \to \min_{A_1},$$

$$\alpha_{12}(A_1) - \alpha = 0,$$

где A_1 – та область, при попадании в которую относим образ к классу 1.

Выражая вероятности ошибок через плотности распределения информативных признаков в классах, получим $L = \int\limits_{A_2} p(\vec{x}\,|\,1)d\vec{x} + \lambda \int\limits_{A_1} p(\vec{x}\,|\,2)d\vec{x} \to \min\limits_{A_1}$.

Поскольку $A_2 = \Omega \setminus A_1$ и плотность распределения информативных признаков для класса q $p(\vec{x}\,|\,q)$ нормирована в Ω, постольку $L = 1 - \int\limits_{A_1} [\,p(\vec{x}\,|\,1) - \lambda p(\vec{x}\,|\,2)]d\vec{x} \to \min\limits_{A_1}$. Для достижения минимума следует интегрировать подынтегральное выражение по всем точкам, где оно положительно. Отсюда находим область A_1:

$$A_1 = \{\vec{x} : p(\vec{x}\,|\,1) - \lambda p(\vec{x}\,|\,2) > 0\}. \tag{46}$$

Добавляя к этому условие $\int\limits_{A_1} p(\vec{x}\,|\,2)d\vec{x} = \alpha$, находим окончательно оптимальные значения A_1^*, λ^*.

Если имеем наблюдённое значение $\vec{x}^0$, то найденное оптимальное решающее правило (46), очевидно, эквивалентно следующему правилу:

$$\frac{p(\vec{x}^0\,|\,1)}{p(\vec{x}^0\,|\,2)} \begin{matrix} > \\ < \end{matrix} \lambda^* \to \begin{matrix} \hat{k} = 1 \\ \hat{k} = 2 \end{matrix}. \tag{47}$$

Дробь в левой части (47) называется отношением правдоподобий.

Умножая неравенство в (46) на произвольное $\tilde{p}_1 > 0$, имеем

$$A^*{}_1 = \{\vec{x} : \tilde{p}_1 p(\vec{x}\,|\,1) - \tilde{p}_2 p(\vec{x}\,|\,2) > 0\}, \tag{48}$$

где $\tilde{p}_2 = \lambda^* \tilde{p}_1$.

Таким образом, формула (46) имеет такой же вид, как в задаче распознавания образов с известными априорными вероятностями ($\tilde{p}_1$, $\tilde{p}_2$).

4.2. Наиболее мощный критерий уровня α проверки гипотезы о математическом ожидании нормальной случайной величины при известной дисперсии

Задачу проверки простой гипотезы о параметре при простой альтернативе $\begin{matrix} H_0 : \theta = \theta_0, \\ H_1 : \theta = \theta_1 \end{matrix}$ можно интерпретировать как задачу распознавания образов двух классов при неизвестных (несуществующих) вероятностях p_1 и p_2, причём $p(\vec{x}\,|1) = p(\vec{x}\,|\,\theta_1)$, $p(\vec{x}\,|2) = p(\vec{x}\,|\,\theta_0)$. В отличие от распознавания образов, где в $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_r \end{pmatrix}$, x_i – различные информативные признаки образа, в проверке гипотез в качестве $\vec{x}$ обычно выступает выборка $\vec{x} = \begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^N \end{pmatrix}$.

Пример 4.1

Найти наиболее мощный критерий уровня α проверки гипотезы $\begin{matrix} H_0 : M\xi = a_0, \\ H_1 : M\xi = a_1, \end{matrix}$ $\xi \sim N(?, \sigma), a_1 > a_0$.

Эта задача сводится к распознаванию 2-х классов в модели Фишера по выборке. По лемме Неймана–Пирсона имеем (46) или через фиктивные априорные вероятности (48). Из формулы (48) по результатам упр. 2.4 из [5] получим, устранив обратно фиктивные априорные вероятности:

$$A^*_1 = \left\{ \vec{x} : \overline{x} > \overline{a} - \frac{\sigma^2 \ln(1/\lambda^*)}{N(a_1 - a_2)} \right\}. \qquad (49)$$

Найдём λ^*. Определяющее его условие $\alpha_{10} = \alpha$ будет выглядеть так: $c = a_0 + u_{1-\alpha}\dfrac{\sigma}{\sqrt{N}}$, где c – правая часть неравенства в формуле (49). Заменяя c этим выражением, имеем наиболее мощный критерий уровня α:

если $\overline{x} > a_0 + u_{1-\alpha}\dfrac{\sigma}{\sqrt{N}}$, то принять H_1, в противном случае – H_0. (50)

Поскольку этот наиболее мощный критерий (50) не зависит от a_1, постольку он является равномерно наиболее мощным.

Отметим, что критерий (50) совпал с критерием проверки гипотез

$$H_0 : M\xi = a_0,$$
$$H_1 : M\xi > a_0,$$

$$sl \begin{matrix} > \\ < \end{matrix} \alpha \to \begin{matrix} H_0 \\ H_1 \end{matrix}, \quad \text{где } sl = 1 - \Phi(\frac{\overline{x} - a_0}{\sigma / \sqrt{N}}), \tag{51}$$

поскольку $1 - \Phi(\dfrac{\overline{x} - a_0}{\sigma / \sqrt{N}}) < \alpha$ ведёт к (50) . Т. о., получили, что критерий (51) является наиболее мощным критерием проверки гипотез $\begin{matrix} H_0 : M\xi = a_0, \\ H_1 : M\xi = a_1, \end{matrix}$ $\xi \sim N(?, \sigma),\ a_1 > a_0$. Мощность его уже знаем: $w(\alpha, a_1) = 1 - \Phi(u_{1-\alpha} - \dfrac{a_1 - a_0}{\sigma / \sqrt{N}})$.

4.3. Последовательный анализ

Пусть проверяется простая гипотеза о параметре против простой альтернативы $\begin{matrix} H_0 : \theta = \theta_0, \\ H_1 : \theta = \theta_1. \end{matrix}$ $\vec{x} = (x^1, x^2, ..., x^n, ..., x^N)^T$ – наблюдённая выборка для

принятия решения. По лемме Неймана–Пирсона $l_N = \dfrac{p(\vec{x}\,|\,\theta_1)}{p(\vec{x}\,|\,\theta_0)} \begin{matrix} > \\ < \end{matrix} \lambda \to \begin{matrix} H_1 \\ H_0 \end{matrix}$.

Венгерский статистик А. Вальд в 1949 г. разработал последовательный критерий проверки гипотез:

1) задаются вероятности α и β ошибок I и II рода критерия;

2) рассчитываются постоянные A и B ($A > B$), которые называют *границами остановки*, такие, что после m наблюдений дальнейшие действия определяются правилами:

- если $l_m \le B$, то прекратить наблюдения и принять H_0;
- если $l_m \ge A$, то прекратить наблюдения и принять H_1;
- если $B < l_m < A$, то продолжить наблюдения.

Теорема Вальда

Границы A и B удовлетворяют условиям: $A \le \dfrac{1-\beta}{\alpha}$ и $B \ge \dfrac{\beta}{1-\alpha}$.

Доказательство. Пусть H_1 принимается именно при m-м наблюдении, т. е. $B < l_n < A$, n=1, 2, …, m - 1 и $l_m \ge A$. Последнее неравенство эквивалентно $p(x^1, ..., x^m\,|\,\theta_1) \ge Ap(x^1, ..., x^m\,|\,\theta_0)$. Поскольку выборка случайна, постольку l_m есть реализация случайного L_m и вероятность ошибки I рода

$$\alpha = P\{L_m \ge A\,|\,\theta_0\} = \int\limits_{R_m^A} p(x^1, ..., x^m\,|\,\theta_0)dx^1 \cdot ... \cdot dx^m =$$

$$\le \frac{1}{A}\int\limits_{R_m^A} p(x^1, ..., x^m\,|\,\theta_1)dx^1 \cdot ... \cdot dx^m = \frac{1}{A}P\{L_m \ge A\,|\,\theta_1\} = \frac{1}{A}(1-\beta).$$

Аналогично этому доказывается и второе неравенство $B \ge \dfrac{\beta}{1-\alpha}$.

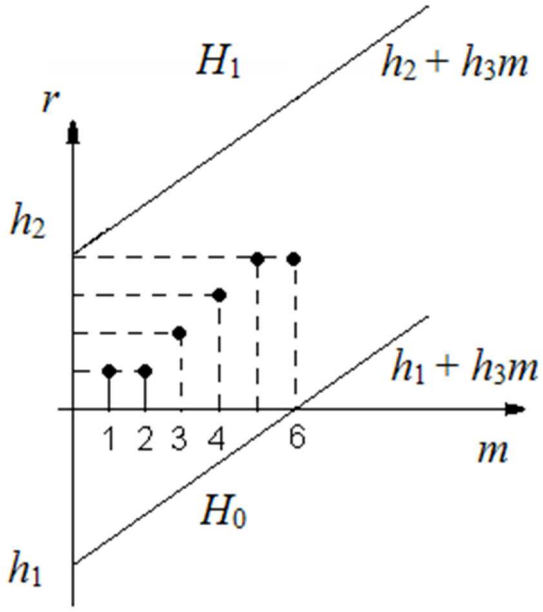
Если возьмём в качестве границы вместо A значение $\frac{1-\beta}{\alpha}$, то позже примем решение H_1; если вместо B примем $\frac{\beta}{1-\alpha}$, то позже примем H_0. При таком выборе границ риски мы не ухудшим.

Биномиальный последовательный критерий отношения правдоподобий. Пусть выборка извлекается из распределения Бернулли: $P\{\xi=1|\,p\}=p, P\{\xi=0|\,p\}=1-p=q$. Проверяемые гипотезы: H_0: $p=p_0$, H_1: $p=p_1$, $p_1>p_0$. Составим отношение правдоподобий: $l_m=\frac{p_1^r q_1^{m-r}}{p_0^r q_0^{m-r}}$, где r – число успехов из m испытаний. Сравним отношение правдоподобий с A и B: $B<\frac{p_1^r q_1^{m-r}}{p_0^r q_0^{m-r}}<A$. Прологарифмируем неравенство и после преобразований получим: $h_1+h_3m<r<h_2+h_3m$, где $h_1=\frac{\ln B}{\ln(p_1q_0/p_0q_1)}$; $h_2=\frac{\ln A}{\ln(p_1q_0/p_0q_1)}$; $h_3=\frac{\ln(q_0/q_1)}{\ln(p_1q_0/p_0q_1)}$.

Пример 4.2

По заданным параметрам α, β, p_0, p_1 строятся две прямые (см. рисунок). Затем проводятся последовательные наблюдения (пусть они будут 1, 0, 1, 1, 1, 0, …), которые откладываются на графике. Пока продолжаем последовательные наблюдения.

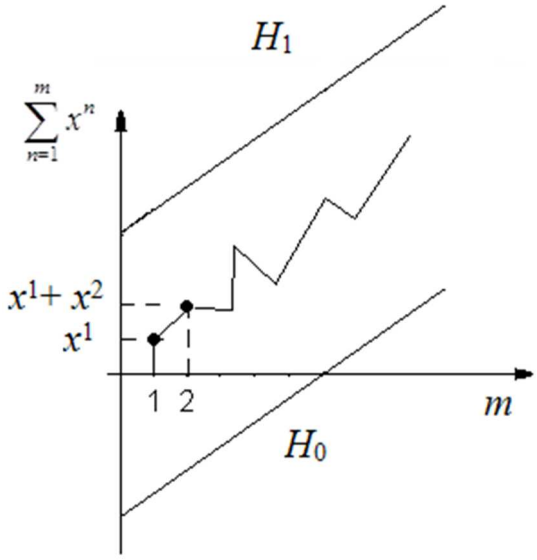
Можно рассчитать ожидаемое число m, при котором будет принята одна из гипотез. Обычно оно составляет около ½ объёма выборки N обычного критерия с теми же α и β.



Нормальный последовательный критерий отношения правдоподобий. Выборка извлекается из распределения $\xi \sim N(a, \sigma)$, дисперсия известна. Проверяемые гипотезы: H_0: $a = a_0$, H_1: $a = a_1$. Пусть $a_1 > a_0$. Запишем отношение правдоподобий

$$l_m = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^m e^{-\frac{1}{2\sigma^2}\sum_{n=1}^{m}(x^n - a_1)^2} \Big/ \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^m e^{-\frac{1}{2\sigma^2}\sum_{n=1}^{m}(x^n - a_0)^2},$$

подобно предыдущему разделу получим условие продолжения наблюдения $h_1 + h_3m < m\overline{x} < h_2 + h_3m$,



где $h_1 = \dfrac{\sigma^2}{a_1 - a_0}\ln B$, $h_2 = \dfrac{\sigma^2}{a_1 - a_0}\ln A$, $h_3 = \dfrac{a_0 + a_1}{2}$. Затем построим график (см. рисунок). На некотором шаге ломаная линия достигнет области принятия гипотезы H_1 или H_0.

Пример 4.3

Партия капронового волокна испытывается на прочность. Характеристика прочности, измеряемая в [г/денье] (удельная прочность), подчиняется закону нормального распределения со стандартным отклонением 0.8 г/денье, причём партия считается хорошей, если среднее значение прочности больше либо равно 5.4 г/денье, и негодной, если среднее значение прочности меньше либо равно 4.9 г/денье. Составить последовательный план контроля капронового волокна, обеспечивающий $\alpha = 0.1$, $\beta = 0.01$.

$a_0 = 5.4$, $a_1 = 4.9$, $\sigma = 0.8$. Так как $(a_1 - a_0) < 0$, то $h_2 + h_3 m < m\bar{x} < h_1 + h_3 m$.
$h_1 = 5.76$, $h_2 = -2.93$, $h_3 = 5.15$.

Последовательные наблюдения пусть будут 5, 4, 5, 4, 3, … На графике, подобном предыдущему, обнаружим, что на пятом наблюдении выходим за нижнюю границу и принимаем гипотезу H_1.

Упражнение 4.1

Известно, что если интенсивность отказов составляет 0.01, то партия гироскопов считается надёжной, если 0.02, то партия ненадёжна и должна быть забракована. Принимая $\alpha = \beta = 0.001$, составить план последовательного контроля. Найти ожидаемое число испытанных гироскопов по формуле
$M(M|\lambda_{худш}) = - h_1 h_2/(h_3)^2$.

5. Планирование эксперимента и дисперсионный анализ

5.1. Оптимальный выбор матрицы плана

Необходимые сведения о постановке задач регрессионного анализа см., например, [6].

Рассмотрим активный эксперимент, когда значения факторов x_1, …, x_K для каждого n-го опыта выбирает исследователь. Какие же значения факторов выбрать?

В качестве критерия оптимальности естественно выбрать величину дисперсии оценок коэффициентов регрессии $D\hat{\mathrm{B}}_k = \sigma^2[(z^T z)^{-1}]_{kk} \to \min$, т. е. задача сводится к выбору такой матрицы плана z, чтобы диагональные элементы матрицы ошибок $P = (z^T z)^{-1}$ были минимальны.

Если значения факторов выбирать совсем произвольно, то все элементы матрицы P можно сделать как угодно малыми. Действительно, если z заменить на $z^{\text{нов}} = az$, где a – число, то P заменится на $P^{\text{нов}} = a^{-2}P \xrightarrow[a\to\infty]{} 0$. При этом новая оценка коэффициентов $\hat{\vec{\beta}}^{\text{нов}} = a^{-2}(z^T z)^{-1} az^T a\vec{y} \approx \hat{\vec{\beta}}^{\text{старое}}$. Таким образом, увеличивая все факторы в a раз, мы получаем те же оценки β_k, а их стандартные ошибки уменьшаются в a раз, что неограниченно улучшает точность: $\dfrac{\hat{\sigma}_k^{\text{нов}}}{\hat{\beta}_k^{\text{нов}}} = \dfrac{a^{-1}\hat{\sigma}_k^{\text{старое}}}{\hat{\beta}_k^{\text{старое}}} \xrightarrow[a\to\infty]{} 0$. Это естественно если зафиксирован уровень шума σ^2, то нужно работать с всё более сильными сигналами (факторами). Однако это обычно нереально, поэтому рассмотрим случай, когда значения факторов можно менять в ограниченных областях:

$$z^{\langle k\rangle^T} z^{\langle k\rangle} = \begin{pmatrix} x_k^0 & x_k^1 & \dots & x_k^{N-1} \end{pmatrix} \begin{pmatrix} x_k^0 \\ x_k^1 \\ \dots \\ x_k^{N-1} \end{pmatrix} = \sum_{n=0}^{N-1} \left(x_k^n \right)^2 = a_k^2 = \text{const} \tag{52}$$

,

где $k = 0, \dots, K$.

Модули векторов столбцов матрицы плана фиксированы, т. е. если сделать x_k больше в одном наблюдении, то в других наблюдениях придётся сделать меньше. Это характерно, например, для распределения денежного ресурса.

Теорема

При условии (52) имеет место неравенство $D\hat{\mathrm{B}}_k \geq \dfrac{\sigma^2}{a_k^2}$, $k = 0, \dots, K$, и минимум (равенство) достигается тогда и только тогда, когда столбцы матрицы плана попарно ортогональны, т. е.

$$z^{\langle k\rangle^T} z^{\langle k'\rangle} = 0, k \neq k'. \tag{53}$$

Пример 5.1

Если по приведённым ниже данным найти стандартные ошибки (оценок) коэффициентов парной регрессии, то получим $\hat{\sigma}_0 = \hat{\sigma} \cdot \sqrt{4.771} = 1.507, \hat{\sigma}_1 = \hat{\sigma} \cdot \sqrt{0.0714} = 0.184$, где $\hat{\sigma}$ – стандартная ошибка аппроксимации, $\hat{\sigma} = 0.69$.

x^n	8	10	5	8	9
y^n	1	3	1	2	3

Если применить оптимальное (ортогональное) планирование $z = \begin{pmatrix} 1 & 8 \\ 1 & 10 \\ 1 & -8 \\ 1 & -5 \\ 1 & -5 \end{pmatrix}$,

которое приведёт к наблюдениям результирующей переменной $\vec{y} = \begin{pmatrix} 1 \\ 3 \\ -4 \\ -3 \\ -3 \end{pmatrix}$, то

при подобной стандартной ошибке аппроксимации $\hat{\sigma} = 0.53$ получим $\hat{\sigma}_0 = 0.23, \hat{\sigma}_1 = 0.03$.

Для случая (53) матрица $(z^T z)$ диагональная с элементами a_k^2, а формулы оценки коэффициентов получаются следующими:

$$\hat{\beta}_k = \frac{z^{\langle k \rangle T} \vec{y}}{a_k^2},\ D\hat{\mathrm{B}}_k = \frac{\sigma^2}{a_k^2},\ \mathrm{cov}(\hat{\mathrm{B}}_k, \hat{\mathrm{B}}_{k'}) = 0,\ k \neq k'. \qquad (54)$$

5.2. Однофакторный дисперсионный анализ

Пример 5.2

В условиях серийного производства изделия часто производятся на нескольких I однотипных станках. Является ли работа станков идентичной? Можно ли смешивать полученные изделия в одну партию? Есть ли эффект в контролируемом параметре при переходе от станка к станку?

Собраны выборочные данные о размере деталей, выпускаемых на машинах трёх типов.

n_i	Машины		
	KRUPP (i = 1)	*FUJI* (i = 2)	УЗТМ (i = 3)
1	4	4	3
2	3	2	3
3	4	2	
4	4		
5	4		

Модель дисперсионного анализа:

$$Y^{n_i} = a_i + E^{n_i},$$

где Y^{n_i} – размер детали в n_i-м наблюдении, $n_i = 1, \ldots, N_i$.

N_i – число наблюдений в i-й группе (деталей с i-й машины); i – номер машины, $i = 1, \ldots, I$;

a_i – неслучайное (должное) значение размера детали i-й машины; E^{n_i} – особенность размера в n_i-м наблюдении (ошибка изготовления).

Считаем, что $ME^{n_i} = 0$ и что дисперсии ошибок одинаковы для всех наблюдений: $DE^{n_i} = \sigma^2$.

Обозначим μ – генеральное среднее, $\mu = \frac{1}{I}\sum_{i=1}^{I} a_i$. α_i – главные эффекты уровней, $\alpha_i = a_i - \mu$. Легко заметить, что $\sum_{i=1}^{I} \alpha_i = 0$.

Задача дисперсионного анализа – исследовать влияние качественного фактора (тип машины) на количественный признак (размер детали), т. е. оценить эффекты уровней α_i, и проверить значимость полученного результата. Так как a_i есть математические ожидания контролируемого параметра в I популяциях, то фактически это проверка гипотезы равенства нескольких математических ожиданий.

Оценки a_i находим методом наименьших квадратов

$\hat{\vec{a}}_{мнк} = \arg\min_{\hat{\vec{a}}} \sum_{i=1}^{I}\sum_{n_i=1}^{N_i}(y^{n_i} - \hat{a}_i)^2$. Обозначим $\Delta = \sum_{i=1}^{I}\sum_{n_i=1}^{N_i}(y^{n_i} - \hat{a}_i)^2$.

$\frac{\delta\Delta}{\delta\hat{a}_j} = -2\sum_{i=1}^{I}\sum_{n_i=1}^{N_i}(y^{n_i} - \hat{a}_i)\delta_{ij} = -2\sum_{n_j}(y^{n_j} - \hat{a}_j) = 0,$ отсюда получается, что

$$\hat{a}_j = \frac{1}{N_j}\sum_{n_j} y^{n_j} = \overline{y}_j$$

– выборочное среднее в j-й группе.

В нашем примере находим:

$$\hat{a}_1 = \frac{1}{5}\cdot(4+3+4+4+4) = 3.8,$$
$$\hat{a}_2 = 2.667, \qquad .$$
$$\hat{a}_3 = 3;$$

оценки эффектов уровней (машин)

$$\hat{\mu} = \frac{1}{3}\cdot(3.8+2.66+3) = 3.155,$$
$$\hat{\alpha}_1 = \hat{a}_1 - \hat{\mu} = 3.8 - 3.155 = 0.644,$$
$$\hat{\alpha}_2 = -0.488,$$
$$\hat{\alpha}_3 = -0.155.$$

Проверка гипотез о главных эффектах: H_0: $\alpha_i = 0$, ($i = 1, \ldots, I$) – означает, что эффектов типа (марки) машины нет, все изделия можно смешать в одну партию. H_1: $\alpha_i \neq 0$ – наоборот, изделия нельзя смешивать в одну партию.

Обозначим $\overline{Y} = \frac{1}{N}\sum_{i=1}^{I}\sum_{n_i=1}^{N_i} Y^{n_i}$, $N = \sum_i N_i$, тогда $\overline{Y} = \frac{1}{N}\sum_{i=1}^{I} N_i\overline{Y}_i$.

Рассмотрим

$$SS_{tot} = \sum_i \sum_{n_i} (Y^{n_i} - \bar{Y})^2 = \sum_i \sum_{n_i} [(Y^{n_i} - \bar{Y}_i) + (\bar{Y}_i - \bar{Y})]^2 =$$

$$= \sum_i \sum_{n_i} (Y^{n_i} - \bar{Y}_i)^2 + \sum_i \sum_{n_i} (\bar{Y}_i - \bar{Y})^2 + 2\sum_i \sum_{n_i} (Y^{n_i} - \bar{Y}_i)(\bar{Y}_i - \bar{Y}).$$

Покажем, что последнее слагаемое равно нулю.

$$2\sum_i (\bar{Y}_i - \bar{Y}) \sum_{n_i} (Y^{n_i} - \bar{Y}_i) = 2\sum_i (\bar{Y}_i - \bar{Y})(\sum_{n_i=1}^{N_i} Y^{n_i} - \sum_{n_i=1}^{N_i} \bar{Y}_i) =$$

$$= 2\sum_i (\bar{Y}_i - \bar{Y})(N_i \bar{Y}_i - \bar{Y}_i N_i) = 0.$$

Следовательно, $SS_{tot} = \sum_i \sum_{n_i} (Y^{n_i} - \bar{Y}_i)^2 + \sum_i \sum_{n_i} (\bar{Y}_i - \bar{Y})^2$, или $SS_{tot} = SS_w + SS_b$,

где $SS_w = \sum_i \sum_{n_i} (Y^{n_i} - \bar{Y}_i)^2$ и $SS_b = \sum_i \sum_{n_i} (\bar{Y}_i - \bar{Y})^2$.

Найдём

$$MSS_w = M\sum_i \sum_{n_i} (Y^{n_i} - \bar{Y}_i)^2 = \sum_i M\frac{N_i - 1}{N_i - 1} \sum_{n_i} (Y^{n_i} - \bar{Y}_i)^2 = \sum_i (N_i - 1)\sigma^2 =$$

$$= \sigma^2 (\sum_i N_i - \sum_i 1) = \sigma^2 (N - I),$$

где $\frac{1}{N_i - 1} \sum_{n_i} (Y^{n_i} - \bar{Y}_i)^2$ – несмещённая оценка дисперсии σ^2, взятая по наблюдениям только i-й группы.

Получили $\frac{MSS_w}{N - I} = \sigma^2$, что означает, что $\frac{ss_w}{N - I} = \hat{\sigma}^2$ является несмещённой оценкой дисперсии ошибок.

При гипотезе H_1 разброс между группами ожидается больше разброса внутри (влияет различие средних значений), т. е. $SS_b > SS_w$. Поэтому

$$sl = \mathrm{P}\left\{\frac{SS_b}{SS_w} \geq \frac{ss_b}{ss_w} \mid H_0\right\} = P\left\{F_{I-1,N-I} \geq \frac{ss_b/(I-1)}{ss_w/(N-I)}\right\}.$$

В последнем равенстве использовано, что SS_b и SS_w независимы и $\frac{SS_b/(I-1)}{SS_w/(N-I)} \sim F_{I-1,\,N-I}$ – распределение Фишера.

Можно преобразовать: $ss_b = \sum_i N_i \overline{y}_i^2 - N\overline{y}^2$, $ss_w = \sum_i \sum_{n_i} (y^{n_i})^2 - \sum_i N_i \overline{y}_i^2$.

В нашем примере

$$ss_b = 5\cdot(3.8)^2 + 3\cdot(2.667)^2 + 2\cdot(3)^2 - 10\cdot(3.3)^2 = 2.63,$$

$$ss_w = 4^2 + 3^2 + 4^2 + 4^2 + 4^2 + 4^2 + 2^2 + 2^2 + 3^2 + 3^2 - 111.53 = 3.47.$$

$$sl = P\left\{F_{2,7} \geq 2.66\right\} = 0.138.$$

Данные согласуются с гипотезой H_0 (эффекта типа (марки) машины нет, все изделия можно смешать в одну партию).

5.3. Полный многофакторный эксперимент

Модель двухфакторного дисперсионного анализа без взаимодействий:

$$Y^{n_{ij}} = \beta_0 + \alpha_{1i} + \alpha_{2j} + E^{n_{ij}},$$

где $Y^{n_{ij}}$ – отклик (количественно зависимая переменная), полученная в n_{ij}-м наблюдении; $n_{ij} = 1, 2, ..., N_{ij}$, N_{ij} – число наблюдений при i-м уровне первого фактора и j-м уровне второго фактора; i – уровни первого качественного фактора, $i = 1, \ldots, I$; j – уровни второго качественного фактора, $j = 1, \ldots, J$.

Если первый фактор установлен на i-м уровне, а второй на j-м, то мы должны получить отклик $\beta_0 + \alpha_{1i} + \alpha_{2j}$, если бы не ошибка (или особенность) $E^{n_{ij}}$ в n_{ij}-м наблюдении.

Дисперсионный анализ с К двухуровневыми факторами. У каждого фактора только два уровня, один уровень можно обозначить -1, другой +1. Обозначим

$\alpha_{11} = \beta_1(-1)$, $\alpha_{12} = \beta_1(+1)$ или $\alpha_{1i} = \beta_1 x_1^{n_{i,\ldots,j}}$, где $x_1^{n_{i,\ldots,j}} = \begin{cases} -1, \text{при } i = 1; \\ +1, \text{при } i = 2. \end{cases}$

$\sum_{i=1}^{2} \alpha_{1i} = \beta_1 \cdot (-1) + \beta_1 \cdot (+1) = 0$, как и должно быть, т. е. α_{1i} – эффекты (главные) уровней первого фактора. Перепишем так же эффекты остальных факторов вплоть до последнего $\alpha_{Kj} = \beta_K x_K^{n_{i,\ldots,j}}$ и получим классическую линейную модель множественной регрессии $Y^n = \beta_0 + \beta_1 x_1^n + \beta_2 x_2^n + \ldots + \beta_K x_K^n + E^n$, $n = 0, \ldots, (N - 1)$. Значения всех K факторов во всех N экспериментах (наблюдениях) составляют матрицу плана эксперимента

$$z = \begin{pmatrix} 1 & x_1^0 & x_2^0 & \ldots & x_K^0 \\ 1 & x_1^1 & x_2^1 & \ldots & x_K^1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_1^{N-1} & x_2^{N-1} & \ldots & x_K^{N-1} \end{pmatrix}.$$

Как следует из предыдущего пункта, наименьшие ошибки в оценке коэффициентов β_k классической линейной модели множественной регрессии будут при ортогональности столбцов матрицы плана. Этому условию удовлетворит матрица

$$z = \begin{pmatrix} + & - & - & .. & - \\ + & + & - & .. & - \\ + & - & + & .. & - \\ + & + & + & .. & - \\ .. & .. & .. & .. & .. \\ + & + & + & .. & + \end{pmatrix}, \qquad (55)$$

у которой 2^{k-1} начальных минусов в столбце фактора k. Последний столбец должен наполовину состоять из (-), наполовину из (+), т. е. $2^{K-1} = N/2$. Отсюда число наблюдений составляет $N = 2^K$.

План эксперимента, в котором полностью использованы все возможные сочетания уровней всех факторов, называют *полным факторным экспериментом*.

Если есть K двухуровневых факторов, то существует $\underbrace{2 \cdot 2 \cdot \ldots \cdot 2}_{K} = 2^K = N$ возможных сочетаний. Полный факторный эксперимент (55) ортогонален.

Пример 5.3

Провести анализ полного факторного эксперимента в пакете *STATISTICA*. Рассмотрим файл данных, поставляемый с пакетом: *File* → *Open* → *Examples* → *Textile.sta*.

Процедура: *Statistics*→ *Industrial statistics and six* σ → *DOE* (*design of experiment*) → 2(*k-p*) *standard design* → *OK*.

В этом разделе можно либо спланировать эксперимент: → *design experiment* (построить матрицу плана *z*), либо проанализировать план (данные) в файле: → *analyze design*.

Анализируем данные в файле *Textile.sta*. Выполняем в диалоге процедуры: *Analyze design* → *Var* → *dependent* (*STRENGTH*), *independent* (1÷6) → *OK*→ *Model* → *No interaction* → *Quick* → *Summary* → *effect estimates* (оценки эффектов) → *mean effect* ($\hat{\beta}_0$) = 11.12. Проверим этот результат по формуле (54):

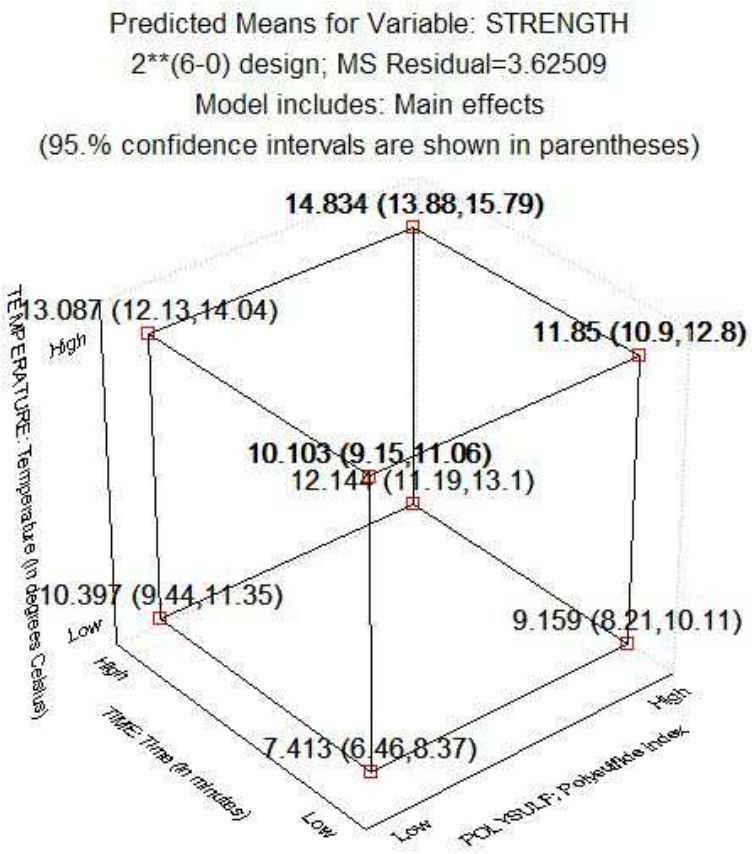
$$\hat{\beta}_0 = \frac{z^{\langle 0 \rangle^T} \vec{y}}{a_0^2} = \frac{(1, 1, \ldots, 1)\begin{pmatrix} y_0 \\ \ldots \\ y_{63} \end{pmatrix}}{64} = \overline{y} = 11.12.$$

POLYSULF effect = 1.764. Проверим этот результат: если посмотреть фактическую матрицу плана *Design* → *display design* и взять столбец фактора *POLYSULF*, то получим

$\hat{\beta}_{polysulf} = \dfrac{(-1,-1,...,+1,+1,...)\vec{y}}{64} = \dfrac{-328+383.8}{64} = 0.873$. Эффектом в *STATISTICA* называют $\alpha_{k+} - \alpha_{k-} = 2\beta_k$, поэтому 2·0.873 = 1.764.

По столбцу уровней значимости (столбец *p*) видим, что значимы эффекты факторов: 1 *POLYSULF*, 4 – *TIME*, 6 – *TEMPERATURE*. Наглядное представление значимости эффектов → *Pareto Chart of Effects*.

График → *Cube plot of predicted means* содержит прогноз для результирующей количественной переменной *STRENGTH* с доверительным интервалом при разнообразных сочетаниях уровней важнейших факторов.



На графике → *observed marginal means* можно посмотреть зависимость от отдельных факторов → *means plot* → выбрать *POLYSULF*, *TIME*, *TEMPERATURE.*

Проведём проверку итоговой модели. Сначала включим все незначимые факторы в член ошибки: *Model* → *Ignore some effects* → 2, 3, 5 → *OK* → *prediction*

and profiling → *predicted vs observed* – модель неплохо описывает эксперимент, т. к. точки на графике лежат вдоль прямой.

Проверка гипотезы о нормальном распределении ошибок: → *residual plots* → *Normal plot*. Можно считать, что ошибки распределены нормально, т. к. точки на графике лежат вдоль прямой.

5.4. Дробные реплики

Пусть число факторов $K = 3$ и нужно оценить 4 коэффициента $\beta_0 \div \beta_3$. Полный факторный эксперимент должен содержать $2^K = 8$ экспериментов (измерений). Эта избыточность нарастает с увеличением K (при $K = 6$ коэффициентов 7 и 64 эксперимента). Можно ли уменьшить число экспериментов? Как при этом запланировать уровни факторов?

Пример 5.4

$K = 3$. Достаточно 4 эксперимента. Для факторов x_1, x_2 сохраним полный факторный ортогональный план. Уровни x_3 свяжем с уровнями предыдущих факторов генерирующим соотношением $x_3 = x_1 x_2$.

x_0	x_1	x_2	$x_3 = x_1 x_2$
+	-	-	+
+	+	-	-
+	-	+	-
+	+	+	+

Легко проверить, что столбцы такой матрицы ортогональны. Планирование по такой схеме называют планированием типа 2^{3-1}, т. е. для трёх факторов, но в основе плана лежит полный факторный эксперимент для (3–1) факторов, поэтому $2^{3-1} = 2^3 \cdot ½$ – число экспериментов – полуреплика.

В общем виде $2^{K-p} = 2^K \frac{1}{2^p}$, где $\frac{1}{2^p}$ – доля полного факторного плана 2^K.

При p = 1, $\frac{1}{2}$ – реплика (доля) полного факторного плана; p = 2, $\frac{1}{4}$ – реплика (доля) полного факторного плана, и т. д.

Можно использовать другое генерирующее соотношение $x_3 = -x_1 x_2$. Получится другая полуреплика, которая вместе с первой даст полный факторный эксперимент.

Дробная реплика $2^{K-p} = 2^K \frac{1}{2^p}$ строится на основе полного факторного эксперимента с (K – p) переменными. Уровни остальных p переменных определяются генерирующими соотношениями с сохранением ортогональности и с учётом возможного взаимодействия факторов.

Поскольку число наблюдений должно быть не меньше числа параметров, постольку имеют смысл реплики вида $2^{K-p} \geq K+1$.

Пример 5.5 (продолжение примера 5.3)

Оценить эффекты тех же шести факторов на ту же зависимую количественную переменную по восьми экспериментам.

Спланируем 1/8 реплику: *Design experiment* → 6 факторов → 6/1/8 →*OK* → *order of runs* → *standard order* → *display design* → *design* 2*(6-3).

+	-	-	-	+	+	+
+	+	-	-	-	-	+
+	-	+	-	-	+	-
+	+	+	-	+	-	-
+	-	-	+	+	-	-
+	+	-	+	-	+	-
+	-	+	+	-	-	+
+	+	+	+	+	+	+

Получим результаты этих 8 экспериментов, а фактически выберем из 64 имеющихся строк нужные с уровнями всех факторов, как в только что спланированной реплике. *Analyze design* → *select cases* (из таблицы выбираем наблюдения: 57; 34; 19; 12; 13; 22; 39; 64) → *OK* → *dependent* → *STRENGTH* → *independent* → 1-6 → *OK* → *OK* → *model* → *no interaction* → *quick* → *Effect estimates*. Получим оценки среднего значения и эффектов факторов с указанием стандартных ошибок и уровней значимости:

	effect	*St. error*	*sl*
Mean/intercept	11.5625	0.66	0.036
POLYSULF	2.225	1.32	0.342
…	…	…	…

При полном факторном эксперименте эти же результаты были лучше (но были получены по числу экспериментов в 8 раз больше):

	effect	*St. error*	*sl*
Mean/intercept	11.12	0.238	0.000
POLYSULF	1.747	0.476	0.00054
…	…	…	…

Вернёмся к *select cases* и выберем наугад любые другие 8 наблюдений, например: с 7 по 12; 37; 61. Программа предупредит, что план не ортогонален, но рассчитает оценки, которые, как и ожидалось, хуже, чем по ортогональной реплике.

	effect	*St. error*	*sl*
Mean/intercept	10.7	3.16	0.077
POLYSULF	1.6	3.99	0.727
…	…	…	…

Вернёмся к ортогональной реплике. В рассматриваемой процедуре *DOE* (*design of experiment*) производится оценка функции регрессии по фактическим

значениям факторов на верхнем и нижнем уровне. В нашем примере она имеет вид:

$$\hat{y}^n = -58.535 + 2.225x_1^n + 0.144x_2^n - 2.29x_3^n + 0.398x_4^n - 0.081x_5^n + 0.227x_6^n.$$

Её коэффициенты можно увидеть в закладке *Prediction and profiling → Predict dependent variable* или *ANOVA → Effects → regression coefficients*. Если сюда подставить все нижние уровни: 6, 150, 1.8, 24, 30, 120, то получим прогноз $\hat{y}^n = 6.6625$. То же самое будет, если подставить все $\hat{\alpha}_{k-}$ в модель дисперсионного анализа:

$$\hat{y}^n = 11.5625 + \hat{\alpha}_{1-} + \hat{\alpha}_{2-} + \hat{\alpha}_{3-} + \hat{\alpha}_{4-} + \hat{\alpha}_{5-} + \hat{\alpha}_{6-} =$$
$$= 11.5625 - 1.112 - 1.44 + 0.688 - 2.39 + 0.49 - 1.137 = 6.6625.$$

Латинские квадраты. Пусть число факторов $K = 3$, все факторы имеют по I уровней. Неполный план, в котором при каждой комбинации уровней двух любых факторов наблюдение проводится только для одного уровня третьего фактора, называют *планом латинского квадрата*.

Латинский квадрат – это решётка с I строками и I столбцами, в которой каждый из I символов (латинские буквы A, B, C, …) встречается по одному разу в каждой строке и каждом столбце.

Пример 5.6

Пусть $I = 4$.

A	*B*	*C*	*D*
B	*C*	*D*	*A*
C	*D*	*A*	*B*
D	*A*	*B*	*C*

Первый фактор – номер строки, 2-й фактор – номер столбца, 3-й фактор – буква. При полном факторном эксперименте было бы I^3 наблюдений, в плане латинского квадрата – I^2.

Пример 5.7

Исследуется влияние четырёх добавок к топливу на снижение содержания окиси углерода в выхлопных газах. Желательно выявить чистый эффект, не зависящий от стиля вождения водителя, марки машины, а также сразу изучить влияние всех факторов по отдельности.

1-й фактор – шофёр (стиль вождения четырёх разных шофёров);

2-й – автомобиль (четыре марки автомобилей);

3-й – добавка (*A*, *B*, *C*, *D*).

Рассмотрим модель $Y^{n_{ijk}} = \beta_0 + \alpha_{1i} + \alpha_{2j} + \alpha_{3k} + E^{n_{ijk}}$, где $Y^{n_{ijk}}$ – снижение содержания окиси углерода, $i, j, k = 1, \ldots, 4$. Примем план латинского квадрата.

авто/шофёр	1	2	3	4
1	*A*	*B*	*D*	*C*
2	*D*	*C*	*A*	*B*
3	*B*	*D*	*C*	*A*
4	*C*	*A*	*B*	*D*

Решение в STATISTICA:

Открываем файл с данными: *File* → *Open* → *Examples* → *Latin.sta*.

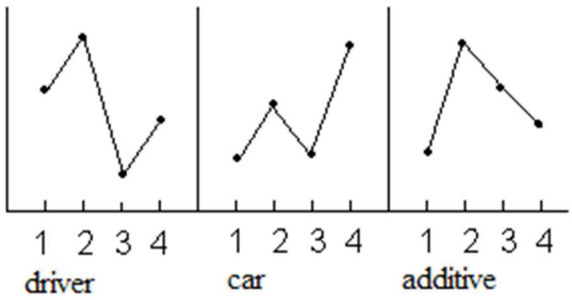
Анализ данных: *Statistics* → *Industrial statistics*…→ *Destine of experiment* (*DOE*) → *advanced tab* → *Latin* → *Analyze design* → *var dependent* (*REDUCTION*), *Independent* (*DRIVERS*, *CAR*, *ADDITIVE*) → *OK*.

Quick → *Marginal means*. Оценка среднего снижения окиси углерода $\hat{\beta}_0 = 20$; оценка стандартного отклонения ошибок $\hat{\sigma} = 4.44$.

$\hat{a}_{31} = A_one = \dfrac{21+20+18+15}{4} = 18$ (среднее внутри группы (уровня) 1), $\hat{\mu}_3 = 20$, $\hat{\alpha}_{31} = \hat{a}_{31} - \hat{\mu}_3$.

	Level	*Means*	*Parameter estimate*
ADDITIVE	*A_one* (*A*)	18	-2
	A_two (*B*)	22	+2
	A_three (*C*)	21	+1
	A_four (*D*)	19	-1

Means plot показывает графики зависимости снижения окиси углерода от уровней факторов:



Summary → *ANOVA* – проверка значимости эффектов уровней:

Effect	*SL*
DRIVER	0.00069 (значимый)
CAR	0.116 (незначимый)
ADDITIVE	0.045 (значимый)

Таким образом, вторая добавка обеспечивает значимое снижение окиси углерода (на 2 единицы от среднего значения).

Греко-латинские квадраты. Два латинских квадрата можно наложить друг на друга.

$$\begin{matrix} A & B & C \\ B & C & A \\ C & A & B \end{matrix} \quad \text{и} \quad \begin{matrix} \alpha & \beta & \gamma \\ \gamma & \alpha & \beta \\ \beta & \gamma & \alpha \end{matrix} \quad \rightarrow \quad \begin{matrix} A\alpha & B\beta & C\gamma \\ B\gamma & C\alpha & A\beta \\ C\beta & A\gamma & B\alpha \end{matrix}$$

Важно следить, чтобы каждая греческая буква встречалась один раз с каждой латинской буквой. По греко-латинскому плану можно оценить главные эффекты четырёх факторов (фактор строки, фактор столбца, латинская буква, греческая буква) на произвольном числе I уровней.

Есть таблицы греко-латинских квадратов всех нужных размеров $I \times I$, кроме квадрата размером 6×6.

6. Статистические решающие правила

Напомним суть «игр против природы»: оперирующая сторона или лицо, принимающее решение, (ЛПР) не обладает полной информацией о состоянии природы (конъюнктура рынка, замыслы конкурентов, погодные условия, состояние оборудования и т. п.) – о стратегии природы θ. Однако ЛПР должен предпринять какое-то действие (стратегию-константу) a. Известна функция потерь $L(a, \theta)$, которые несёт оперирующая сторона, предприняв действие a в условиях природы θ.

Новым в «статистических играх» будет следующая *информационная гипотеза*: оперирующей стороне разрешается провести эксперимент, в котором она получает значение x величины X, которая связана с θ, причём известно условное распределение $p_X(x,\theta) \triangleq p(x|\theta)$.

Функция $a = d(x)$, определяющая действия ЛПР при всех возможных x, называется *решающей функцией* (решающим правилом, стратегией-функцией).

Применение решающего правила d даже при фиксированном θ в силу случайности X ведёт к случайным потерям $L(d(X),\theta)$. Среднее (условное) их значение называется *функцией риска*

$$R(d,\theta) = M\{L[d(X),\theta]|\theta\} = \int L[d(x),\theta]p(x|\theta)dx. \qquad (56)$$

Риск – это ожидаемые, средние потери. Функция риска показывает их зависимость от применяемого решающего правила и состояния природы. Т. о. при использовании стратегий-функций акцент переводится с функции потерь на функцию риска, и поиск оптимальной функции d будет связан с минимизацией риска.

6.1. Минимаксное решающее правило

Применим стандартный подход – критерий Вальда, который состоит в следующем: т. к. θ неизвестен, то любое правило $d(x)$ лучше охарактеризовать максимально возможным при нём риском – $r_m(d) = \max_{\theta} R(d,\theta)$ (максимальный риск правила *d*). Следует остановиться на правиле, где этот риск минимален. Это определяет *минимаксное решающее правило*:

$$d_{mm} = \arg\min_{d} \max_{\theta} R(d,\theta). \qquad (57)$$

Часто его можно найти перебором всех возможных *d* по минимуму $r_m(d)$.

$r_m{}^* = r_m(d_{mm})$ – минимаксный риск.

Пример 6.1

Обычная цена месторождения 2 у. е. На торги выставлено 3 месторождения (пакет) за 4 у. е. Эта дешевизна неспроста: из них одно (состояние «природы» θ_1) или два (состояние «природы» θ_2) исчерпаны. У вас есть 2 варианта действий: a_1 – покупать, a_2 – не покупать. Функция потерь $L(a,\theta) = \begin{array}{c} \\ a_1 \\ a_2 \end{array}\!\!\begin{array}{c} \begin{array}{cc} \theta_1 & \theta_2 \end{array} \\ \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \end{array}$. Ваши потери в 1 у. е. будут, если неплохие (при θ_1) месторождения достанутся конкурентам.

Продавец согласен допустить вас проверить одно любое месторождение. Результатом может быть: x_1 – месторождение годное или x_2 – месторождение исчерпанное. Распределение результата X в зависимости от состояния природы очевидно таково:

$$p(x\,|\,\theta) = \begin{array}{c} \\ x_1 \\ x_2 \end{array}\!\!\begin{array}{c} \begin{array}{cc} \theta_1 & \theta_2 \end{array} \\ \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \end{array}.$$

Возможно всего 4 различных решающих правила:

$$d_1(x) = \begin{cases} a_1, x = x_1, \\ a_1, x = x_2; \end{cases} \quad d_2(x) = \begin{cases} a_1, x = x_1, \\ a_2, x = x_2; \end{cases} \quad d_3(x) = \begin{cases} a_2, x = x_1, \\ a_1, x = x_2; \end{cases} \quad d_4(x) = \begin{cases} a_2, x = x_1, \\ a_2, x = x_2. \end{cases}$$

Можно обозначить $d_1 = 1{,}1$; $d_2 = 1{,}2$; $d_3 = 2{,}1$; $d_4 = 2{,}2$.

В общем случае если число возможных действий – Na и число возможных исходов эксперимента – Nx, то имеется $(Na)^{Nx}$ различных решающих правил (для одного исхода – Na, для двух – $Na \cdot Na$ вариантов и т. д.).

Найдём значения функции риска правила d_1 при всех θ. По формуле (56)

$$R(d_1, \theta) = \sum_x L(d_1(x), \theta)\, p(x \mid \theta) = L(a_1, \theta)\, p(x_1 \mid \theta) + L(a_1, \theta)\, p(x_2 \mid \theta) =$$

$$= \left| \text{сразу для двух} \theta \right| = (0, 2) \cdot (\tfrac{2}{3}, \tfrac{1}{3}) + (0, 2) \cdot (\tfrac{1}{3}, \tfrac{2}{3}) = (0, \tfrac{2}{3}) + (0, \tfrac{4}{3}) = (0, 2).$$

Поэтому максимальный риск этого правила $r_m(d_1) = \max\limits_{\theta}(0, 2) = 2$. Точно так же находим $r_m(d_2) = 2/3$, $r_m(d_3) = 4/3$, $r_m(d_4) = 1$.

По (57) минимаксное решающее правило $d_{mm} = \arg\min\limits_{d} \begin{pmatrix} 2 \\ 2/3 \\ 4/3 \\ 1 \end{pmatrix} = d_2$.

Минимаксный риск $r_m{}^* = r_m(d_{mm}) = r_m(d_2) = 2/3$.

В отличие от стратегических игр, минимаксный риск 2/3 в нашей статистической игре не является гарантированным результатом при применении минимаксной стратегии d_2. Например, получив в эксперименте x_1, будем покупать пакет месторождений, но при этом может быть $\theta = \theta_2$ и потери будут 2 у. е.

6.2. Байесовское решающее правило

Будем предполагать, что состояние природы Θ – случайная величина с известным априорным распределением $p(\theta)$, тогда любое правило $d(x)$ можно охарактеризовать средним по состояниям природы риском (*байесовский риск правила d*):

$$r_B(d) = MR(d, \Theta) = \int R(d, \theta)\, p(\theta) d\theta \qquad (58)$$

(ранее каждое d характеризовали худшим (максимальным) по состояниям природы риском).

Следует остановиться на правиле, где этот риск минимален, – *байесовское решающее правило*: $d_B = \arg\min_d r_B(d)$.

$r_B{}^* = r_B(d_B)$ – минимальный байесовский риск.

Подчеркнём, что байесовское решающее правило зависит от предположенного априорного распределения $p(\theta)$.

Пример 6.2 (продолжение 6.1)

Будем считать, что $p(\theta_1) = \alpha,\ p(\theta_2) = 1-\alpha$ составляет априорное распределение. Найдём байесовские риски всех решающих правил. Функцию риска уже рассчитывали.

$$r_B(d_1) = \sum_\theta R(d_1,\theta)\,p(\theta) = R(d_1,\theta_1)\alpha + R(d_1,\theta_2)(1-\alpha) =$$

$$= 0\alpha + 2\cdot(1-\alpha) = 2\cdot(1-\alpha);$$

$$r_B(d_2) = \ldots = \tfrac{1}{3}\alpha + \tfrac{2}{3}\cdot(1-\alpha) = \tfrac{1}{3}\cdot(2-\alpha);$$

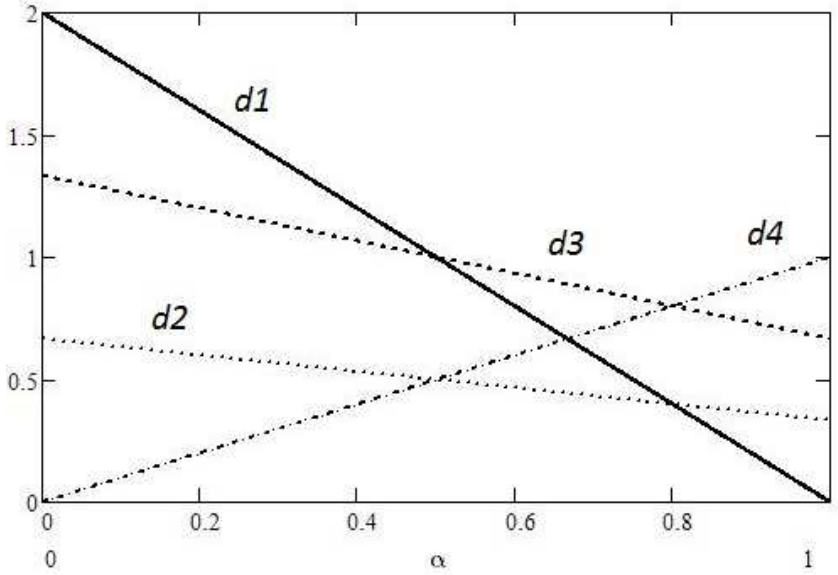
$$r_B(d_3) = \ldots = \tfrac{2}{3}\alpha + \tfrac{4}{3}\cdot(1-\alpha) = \tfrac{2}{3}\cdot(2-\alpha);$$

$$r_B(d_4) = \ldots = 1\alpha + 0\cdot(1-\alpha) = \alpha.$$

Пусть α = 0.4 (т.е. вероятность «хорошего» пакета меньше половины),

$r_B(d) = \begin{pmatrix} 1.2 \\ 0.533 \\ 1.06 \\ 0.4 \end{pmatrix}$, $d_B = \arg\min_d r_B(d) = \arg\min_d \begin{pmatrix} 1.2 \\ 0.533 \\ 1.06 \\ 0.4 \end{pmatrix} = d_4$, минимальный байесовский риск $r_B{}^* = r_B(d_B) = 0.4$ у. е.

Можно изобразить байесовские риски всех решающих правил в зависимости от априорного распределения.



Таким образом, в зависимости от α байесовским является либо 4-е, либо 2-е, либо 1-е решающее правило:

$$d_B = \begin{cases} d_4, \alpha < 1/2; \\ d_2, 1/2 \le \alpha < 4/5; \\ d_1, \alpha \ge 4/5. \end{cases} \tag{59}$$

Применённый подход легко реализуем, когда стратегий-функций не много и они явно известны. Однако хотелось бы уметь автоматически строить d_B как функцию от *x*, т. е. для каждого возможного значения *x* заготовить $a = d_B(x)$.

Теорема (байесовский принцип)

Байесовская решающая функция имеет вид:

$$a = d_B(x) = \arg\min_a EMV(a \mid x), \tag{60}$$

где $EMV(a \mid x)$ – ожидаемая денежная ценность (потери) действия *a* (условная при известном *x*). Т. о. следует делать то, что несёт наименьшие ожидаемые при данном *x* потери.

$$EMV(a \mid x) = M[L(a,\Theta) \mid x] = \int L(a,\theta) p(\theta \mid x) d\theta.$$

Здесь $p(\theta \mid x)$ – апостериорное распределение состояния природы,

$$p(\theta \mid x) = \frac{p(\theta)\, p(x \mid \theta)}{\int p(\theta)\, p(x \mid \theta) d\theta} \text{ – формула Байеса.}$$

Доказательство. Подставляем в байесовский риск (58) функцию риска (56), меняем местами интегрирование, и, т. к. по формуле Байеса $p(\theta)\, p(x \mid \theta) = p(x)\, p(\theta \mid x)$, имеем,

$$r_B(d) = \int \{ \int L[d(x), \theta] p(\theta \mid x) d\theta \} p(x) dx . \qquad (*)$$

Минимизация $r_B(d)$ достигается минимизацией внутреннего интеграла при каждом x, значит нужно брать действие $d(x) = a$ дающее интеграл поменьше, что и означает (60).

Пример 6.3 (продолжение 6.2)

Построим байесовскую решающую функцию по формуле (60). Вначале рассчитаем ожидаемую денежную ценность каждого действия при каждом x:

$EMV(a \mid x) = \sum_{\theta} L(a, \theta) \frac{p(\theta)\, p(x \mid \theta)}{p(x)}$. Пусть $x = x_2$.

$$EMV(a \mid x_2) = \frac{1}{p(x_2)} [L(a, \theta_1) \alpha p(x_2 \mid \theta_1) + L(a, \theta_2)(1 - \alpha)\, p(x_2 \mid \theta_2) =$$

$$= \frac{1}{p(x_2)} \begin{bmatrix} 0... + 2 \cdot (1 - \alpha) \cdot \frac{2}{3} \\ 1\alpha \cdot \frac{1}{3} + 0... \end{bmatrix} = \frac{1}{p(x_2)} \begin{bmatrix} \frac{4}{3} \cdot (1 - \alpha) \\ \frac{1}{3} \cdot \alpha \end{bmatrix}.$$

Пусть α = 0.4, $EMV(a \mid x_2) = \frac{1}{p(x_2)} \begin{bmatrix} 0.8 \\ 0.133 \end{bmatrix}$, байесовское действие при этом x:

$d_B(x_2) = \arg\min_a EMV(a \mid x_2) = \arg\min_a \frac{1}{p(x_2)} \begin{bmatrix} 0.8 \\ 0.133 \end{bmatrix} = a_2$. Аналогично

$d_B(x_1) = \arg\min_a \frac{1}{p(x_1)} \begin{bmatrix} 0.4 \\ 0.266 \end{bmatrix} = a_2$. Т. е. байесовская решающая функция

$d_B(x) = \begin{cases} a_2, x = x_1 \\ a_2, x = x_2 \end{cases} = d_4$. Минимальный байесовский риск по формуле (*)

$$r_B{}^* = r_B(d_B) = \sum_x EMV(d_B(x) \mid x)\, p(x) = \sum_x (\min_a EMV(a \mid x))\, p(x) = 0.133 + 0.266 = 0.4$$

у.е. Результаты совпадают с прежним решением (путём перебора байесовских рисков всех решающих правил).

Или при произвольном α

$$d_B(x_2) = \arg\min_a EMV(a \mid x_2) = \begin{cases} a_1, \alpha > \frac{4}{5}, \\ a_2, \alpha \le \frac{4}{5}, \end{cases} \qquad (61)$$

точно так же находим $d_B(x_1) = \begin{cases} a_1, \alpha > \frac{1}{2}, \\ a_2, \alpha \le \frac{1}{2}. \end{cases}$ Вместе с (61) это совпадает с (59).

Т. о., при исчерпанном проверенном месторождении покупать ($a_B = a_1$) следует, только если априорная вероятность того, что в пакете всего одно исчерпанное месторождение, больше 4/5. Оптимальный байесовский риск, например, при α = 0.6 равен 0.466 67 (см. рисунок перед формулой (59)).

Если априорное распределение неизвестно, то можно считать его равномерным, т. е. α = 1/2.

Свойства минимаксных и байесовских стратегий. Байесовское решение, построенное для любого априорного распределения $p(\theta)$, является допустимым, т. е. нет другого лучшего правила с риском, меньшим при всех θ.

Если существует априорное распределение $p(\theta)$, для которого байесовское правило d_B имеет постоянную функцию риска $R(d_B, \theta) = \text{const}$, то такое d_B есть минимаксное правило.

Пример 6.4

На технологическую линию может поступать сырье с малым θ_1 и большим θ_2 количеством примесей. Известно, что в среднем поступает 60 % сырья первого вида и 40 % сырья второго вида. Для использования различных видов сырья предусмотрено три режима работы технологической линии: a_1, a_2, a_3. Потери, связанные с качеством выпускаемой продукции и расходом сырья, в зависимости от качества сырья и режима работы технологической линии составляют:

	θ_1	θ_2
a_1	0	5
a_2	1	3
a_3	3	2

Допускается произвести грубый предварительный анализ содержания примесей (точный лабораторный анализ невозможен, т. к. требует затраты значительного времени, а значит, простоя оборудования). Результаты эксперимента могут быть следующими: x_1 – примесей не обнаружено; x_2 – примеси в небольшом количестве; x_3 – примесей много. Условное распределение результатов известно:

	θ_1	θ_2
x_1	0.6	0.2
x_2	0.25	0.3
x_3	0.15	0.5

Найдём минимаксное решающее правило.

Всего имеется 27 решающих правил: $d_1 = 1,1,1; d_2 = 1,1,2; ...; d_{27} = 3,3,3$ (в каждом указаны действия при исходах эксперимента x_1, x_2, x_3 соответственно). Максимальные риски всех правил:

5 4 3.5 4.4 3.4 2.9 4.1 3.1 2.6 4.6 3.6 3.1 4 3 2.5 3.7 2.7 2.2 4.4 3.4 2.9 3.8 2.8 2.5 3.5 2.7 3

Поэтому минимаксным оказывается решающее правило $d_{mm} = d_{18} = 2,3,3$. Минимаксный риск равен 2.2 у.е.

Имея функцию риска (для всех стратегий-функций и состояний природы), можно, как и в обычной теории игр, найти оптимальную смешанную стратегию вместо чистой (минимаксной). Результат будет: применять d_{18} и d_{27} с вероятностями 5/7 и 2/7. Средний риск будет при этом 2 1/7 = 2.142 86.

Найдём байесовское решающее правило.

По формуле (60) при заданном априорном распределении (0.6, 0.4) получим байесовскую стратегию $d_B = 1,2,3$. Минимальный байесовский риск равен 1.58 у. е. При «оптимальной» смешанной стратегии природы – априорном распределении (1/7, 6/7) – минимальный байесовский риск становится наибольшим и равным 2.142 86. Он соответствует «наихудшему» априорному распределению.

6.3. Выборочный контроль качества продукции

Поставщик должен поставить потребителю партию изделий объёмом M = 100 шт. По договору поставщик обязуется заменять дефектные изделия, выявленные в процессе эксплуатации партии за свой счёт. Издержки замены одного дефектного изделия c_2 = 2000 р. Перед отправкой поставщик проводит выборочную проверку партии. Объем выборки N = 30 шт. Издержки проверки одного изделия c_1 = 180 р. Если число обнаруженных дефектных изделий в выборке $X \le m$, где m – приёмное число, то партия принимается (решение a = 1). Выявленные в выборке дефектные изделия заменяются на годные, это не даёт дополнительных издержек. Если $X > m$, партия отвергается (решение a = 2), все изделия в ней проверяются, годные остаются в резерве (например, для замены дефектных в выборке в принятых партиях).

Требуется найти оптимальное приёмное число m, минимизирующее риск (средние потери (издержки) в случае применения этого решающего правила).

Обозначим θ – доля дефектных изделий в партии (характеризует состояние «природы»). $0 \le \theta \le 1$. Функция потерь

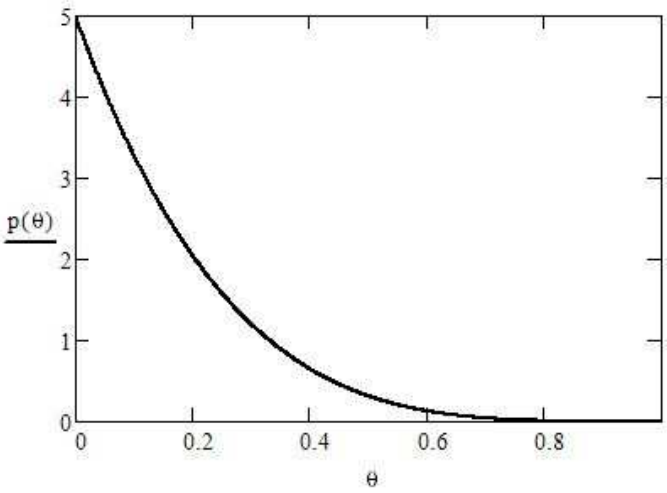
$$L(a, \theta) = \begin{cases} c_1 N + c_2 (M - N)\theta, \text{ при } a = 1, \\ c_1 N + c_1 (M - N), \text{ при } a = 2. \end{cases}$$

Мыслимые решающие правила:

$a = d_m(x) = \begin{cases} 1, x \le m, \\ 2, x > m, \end{cases}$ $m = 0,1,...,N$. Каждое из них нумеруется применяемым в нём приёмным числом, всего (N + 1) правил.

Функция риска правила m: $R(m,\theta) = \sum_{x=0}^{N} L[d_m(x),\theta]\, p(x \mid \theta)$; условное распределение числа дефектных изделий в выборке $p(x \mid \theta)$ для большой партии можно принять биномиальным: $p(x \mid \theta) = C_N^x \theta^x (1-\theta)^{N-x}$.

Применим байесовский подход для нахождения оптимального решающего правила. Считаем известным априорное распределение θ в партиях $p(\theta)$. При отлаженном производстве доля дефектных изделий обычно мала, поэтому можно для него принять бета-распределение с параметрами формы 1 и 5:



Байесовские риски (58) всех правил: $r_B(m) = \int_0^1 R(m,\theta)\, p(\theta)\, d\theta$, и байесовское решающее правило (оптимальное приёмное число m_B): $m_B = \arg\min_m r_B(m)$.

Расчёт в *MathCAD* по этим формулам дал оптимальное приёмное число m_B = 2. Средние потери (издержки) в случае применения решающего правила с приёмным числом 2 составят 16 040 р. на партию (минимальный байесовский риск).

7. Непараметрические методы

До сих пор мы считали, что распределение наблюдений принадлежит некоторому конкретному типу (нормальное, пуассоновское и т. п.), и методы оценивания и проверки гипотез позволяли нам делать выводы о неизвестных параметрах таких распределений. Поэтому справедливость результатов зависела от адекватности исходного предположения о параметрическом семействе распределений.

Очевидно, было бы полезно иметь подходы, не связанные с определённым видом распределения, свободные от распределения. В главе рассмотрены некоторые непараметрические методы.

7.1. Критерий однородности Уилкоксона–Манна–Уитни

Пусть $x^1, x^2, ..., x^{N_1}$ – выборка из популяции с непрерывным распределением $F_X(x)$, а выборка $y^1, y^2, ..., y^{N_2}$ – из популяции с непрерывным распределением $F_Y(y)$. Однородны ли эти популяции, т. е. верна ли гипотеза $H_0 : F_X(z) = F_Y(z), \forall z$ против альтернативы $H_1 : F_X(z) \neq F_Y(z)$ (разные распределения)?

Пример 7.1

Группа из 6 испытуемых получала с 50 лет до смерти плацебо, а группа из 9 испытуемых получала по 1 г витамина *C*. Срок жизни тех и других:

x^n	65	71	67	78	84	69	–	–	–
y^n	70	63	89	83	76	87	84	80	85

Можно ли считать, что приём витамина *C* увеличивает продолжительность жизни или эти популяции однородны?

Составим общий вариационный ряд:

63	65	67	69	70	71	76	78	80	83	84	84	85	87	89
1	2	3	4	5	6	7	8	9	10	11.5	11.5	13	14	15

Каждому элементу поставлен в соответствие его номер в ряду – ранг. Если несколько элементов совпадают, то каждому из них присваивается ранг, равный среднему арифметическому их номеров.

Обозначим W – сумма рангов элементов более короткой выборки (длины N_1). При справедливости H_0 W имеет симметричное распределение с $MW = \frac{1}{2}N_1(N_1 + N_2 + 1)$. Большие отклонения w от MW свидетельствуют против H_0 (т. е. данные из меньшей выборки сбились в начало (при малом w) или в конец (при большом w)). Поэтому уровень значимости составит

$$sl = P\{W \le w \text{ или } W \ge 2MW - w \mid H_0\} = 2P\{W \le w \mid H_0\} \text{ при } w < MW.$$

При противоположном последнем неравенстве – подобная формула.

В нашем примере $MW = \frac{1}{2} \cdot 6 \cdot (6 + 9 + 1) = 48$, $w = 2 + 3 + 4 + 6 + 8 + 11.5 = 34.5$, поэтому $sl = 2P\{W \le 34.5 \mid H_0\} = 2 \cdot (0.05 \div 0.1) = (0.1 \div 0.2)$ (см. таблицу квантилей распределения величины W [3]). Такие значения sl интерпретируются как согласие с H_0, т. е. изменение продолжительности жизни не доказано. Если конкурирующая гипотеза H_1: $F_X(z)$ сдвинута влево, т. е. в сторону уменьшения лет жизни, то данные более значимы против H_0, $sl = P\{W \le 34.5 \mid H_0\} = (0.05 \div 0.1)$.

Если N_1, $N_2 > 7$, то W распределена приближённо нормально с дисперсией $DW = \frac{1}{12}N_1N_2(N_1 + N_2 + 1)$. Это ведёт к приближенному уровню значимости для двусторонней альтернативы $sl(z) = 2 \cdot [1 - \Phi(|z|)]$,

где $z = \dfrac{w - \frac{1}{2}N_1(N_1 + N_2 + 1)}{\sqrt{\frac{1}{12}N_1N_2(N_1 + N_2 + 1)}}$.

7.2. Критерий Краскела–Уоллиса

Этот критерий ещё называют дисперсионным анализом на основе рангов. Цель его – определить, являются ли средние в K (3-х или более) популяциях одинаковыми. Обозначим N_k – число наблюдений в k-й популяции, $k = 1, 2, ..., K$. Общее число наблюдений $\sum_{k=1}^{K} N_k = N$. Наблюдения должны быть независимыми. В отличие от обычного дисперсионного анализа, не предполагается нормальная распределённость и равенство дисперсий в популяциях.

Проверяемые гипотезы:

H_0 : распределения всех K популяций одинаковы.

H_1 : распределения K популяций не одинаковы.

Составим объединённый вариационный ряд выборок. Аналогично п. 7.1 найдём ранги. Обозначим за W_k сумму рангов наблюдений из k-й популяции; W_k / N_k – средние ранги. Если H_0 верна, то все средние должны быть близкими и близки к общему среднему рангу $\frac{1+2+...+N}{N} = \frac{1}{2}(N+1)$.

Статистика критерия – взвешенная сумма квадратов отклонений, $V = \frac{12}{N(N+1)} \sum_{k=1}^{K} N_k [\frac{W_k}{N_k} - \frac{1}{2}(N+1)]^2$. Большие значения V свидетельствуют против H_0 (в пользу H_1), поэтому $sl = P\{V \geq v \mid H_0\}$, v – наблюдённое значение статистики. Таблица распределения V приведена в [8]. Если все N_k > 5, то V при H_0 имеет приближённо распределение хи-квадрат с (K-1) степенями свободы, тогда $sl \approx 1 - \Phi_{K-1}^{\chi^2}(v)$.

Пример 7.2

Планируется семинар по менеджменту для коммерсантов из сферы производства (5 человек), финансов (6 человек) и торговли (5 человек). Нужно решить,

проводить его для всех вместе или по отдельности. Всем слушателям был предложен тест по основам менеджмента, полученные баллы представлены в таблице:

Баллы слушателей и их ранги (в скобках)		
Производство	Финансы	Торговля
50 (8)	48 (5)	39 (1)
54 (11)	49 (6.5)	66 (16)
59 (14.5)	49 (6.5)	44 (3)
59 (14.5)	52 (10)	47 (4)
40 (2)	56 (12)	51 (9)
–	57 (13)	–

Есть ли различие у групп в исходном уровне подготовки?

$w_1 = 8 + 11 + 14.5 + 14.5 + 2 = 50,\ w_2 = 53,\ w_3 = 33$, значение статистики критерия $v = 1.32$, $sl \approx 1 - \Phi_2^{\chi^2}(1.32) = 0.516$. Т. о., данные не значимы против H_0, значит, можно распределения (в т. ч. средние) считать одинаковыми и семинар проводить общий.

По точному распределению V при объёмах выборок 6, 5, 5 верхняя критическая точка порядка 0.1 равна 4.56, а порядка 0.05 – 5.73.

Другие критерии непараметрической статистики (критерии согласия Колмогорова и χ^2, критерий серий случайности) см. [4]. Ранговая корреляция Спирмена, конкордация Кендалла и критерии для категоризованных переменных – см. [5].

7.3. Бутстреп-процедуры

Пусть есть небольшая выборка:

$$x^1, x^2, ..., x^n, ..., x^N. \qquad (62)$$

Рассмотрим *i.i.d.* (независимые одинаково распределённые) X_*^n с $P\{X_*^n = x^m\} = 1/N,\ \forall n,\ m$. Совокупность $X_*^1, X_*^2, ..., X_*^N$ – бутстреп-выборка. Такие выборки получаются взятием наудачу значений из набора (62), при этом возможны повторения. Что это может дать – некоторые примеры рассмотрим в следующих двух пунктах.

Бутстреп-оценка стандартного отклонения (бутстреп-стандартная ошибка) оценки S^2 для дисперсии. Пусть мы имеем выборку наблюдений случайной величины. По статистике S^2 мы можем оценить дисперсию этой случайной величины. Какова стандартная ошибка данной оценки? Чтобы её найти нужно взять корень квадратный из какой-либо оценки дисперсии оценки S^2. Для оценки дисперсии по статистике «*s*-квадрат» нужны несколько значений S^2, а у нас всего одно! В этих условиях мы размножаем одну реальную выборку в *M* бутстреп-выборок, для каждой рассчитываем S^2 и можем найти корень квадратный из «*s*-квадрата».

Пример 7.3 (продолжение примера 1.1)

Найти бутстреп-стандартную ошибку оценки дисперсии случайной величины по выборке: 3, 7, 7, 0, 2, 1, 2, 3, 5, 4.

При *M* = 1000 получено *S.E.* S^2 = 1.753. Эта величина даёт представление о точности ранее полученной оценки дисперсии случайной величины s^2 = 5.6.

Как проверить правильность бутстреп-оценки? Используем выборку из нормального распределения с известной дисперсией. В этом случае известно стандартное отклонение оценки S^2 (26).

Пример 7.4

Выборка из $N(2, 4)$: 4.563, -0.19, -8.621, -1.814, 7.498·10^-3, 5.827, 6.722, -0.19, 1.955, -7.194. При M = 1000 получено *S.E.* S^2 = 8.729. Стандартное отклонение оценки S^2= 7.542. При увеличении N совпадение улучшается [11].

Уточнённые доверительные интервалы. Каковы характеристики бутстреп-распределения? $MX_*^n = \sum_{m=1}^{N} x^m \cdot \frac{1}{N} = \overline{x}$, $DX_*^n = \sum_{m=1}^{N} (x^m - \overline{x})^2 \cdot \frac{1}{N} = \frac{N-1}{N} s^2$. По следствию из ЦПТ в применении к бутстреп-выборке имеем сходимость:

$$\frac{\overline{X}_* - \overline{x}}{S_* / \sqrt{N}} \xrightarrow{N \to \infty} N(0,1) \qquad (63)$$

так же, как и по ЦПТ, $\frac{\overline{X} - a}{S / \sqrt{N}} \xrightarrow{N \to \infty} N(0,1)$. Если предполагать нормальное распределение, то левая часть последнего имеет распределение Стьюдента, что ранее привело нас к точному доверительному интервалу (28). Если распределение неизвестно и N невелико, то границы доверительного интервала неточны. Основная идея бутстреп-процедуры состоит в том, чтобы считать распределения левых частей обеих последних формулах близкими всегда, даже при фактическом малом N. Тогда это общее распределение, нужное во второй формуле для построения доверительного интервала для a, можно изучать по поведению левой части первой формулы. В частности, можно оценить квантиль нужного порядка с помощью многократного генерирования бутстреп-выборок и расчёта левой части (63), причём числитель и знаменатель оцениваются независимо.

Пример 7.5

Проведём имитационное исследование вероятности накрытия математического ожидания стандартным интервалом (28) и интервалом при уточнённых квантилях. Распределение $M(1/3)$, $N = 9$, доверительная вероятность 0.9,

$M = 200$. При 2000 прогонах (имитациях, исходных выборках) частоты накрытия составили: для стандартного интервала 0.849, для бутстреп-уточнённого 0.857. При других распределениях (нормальном, равномерном) данная бутстреп-процедура не даёт улучшения.

Упражнение 7.1

Проверить с помощью имитационного моделирования в *MathCAD* основную идею бутстреп-процедур. Например: 1) сгенерировать небольшую ($N \sim 10$) выборку из $M(1/a)$; 2) найти γ-доверительный интервал для a с использованием асимптотических нормальных стандартных квантилей; 3) с помощью бутстреп-процедуры найти улучшенный γ-доверительный интервал для a, оценив нужную квантиль выборочной квантилью

$$\hat{x}_{p,N} = \begin{cases} x^{(ceil(Np))}, Np - \text{дробное}; \\ px^{(Np)} + (1-p)x^{(Np+1)}, Np - \text{целое}; \end{cases}$$

4) повторяя многократно п. 1÷3 упр. 7.1, оценить вероятности накрытия a интервалами в п. 2 и 3, сравнить с γ.

8. Лабораторный практикум

(выполняется с использованием пакета программ *STATISTICA*)

Упражнение 8.1

1. Повторить вычисление оценок математических ожиданий и стандартных отклонений из упр. 1.8.
2. Выполнить упражнение 9.1 и найти доверительный интервал для математического ожидания из упр. 9.2.

Меню *Statistics* → *Basic Statistics and Tables* → *Descriptive Statistics.*

Упражнение 8.2

1. Повторить пример 3.4.
2. Выполнить упр. 9.5.

Basic Statistics and Tables → *Difference Tests*: *R*, *%*, *Means* → *Difference between two means* (*normal distribution*) → *Single mean*…

Упражнение 8.3

1. Проверить квантили и значения функций распределения из примеров и упражнений гл. 1÷5 для распределений: нормального, Стьюдента, χ^2, Фишера.
2. Найти точное значение *sl* в упр. 9.10.

Basic Statistics and Tables → *Probability calculator*.

Упражнение 8.4

Выполнить упр. 9.9, 9.11 и 9.10 (с учётом того, что в процедуре альтернативная гипотеза H_1: $\sigma_2^2 \neq \sigma_1^2$, т. е. в ней *sl* в 2 раза больше, чем при односторонней альтернативе).

Basic Statistics and Tables → *T-test for Independent Samples by Variables*.

Упражнение 8.5

1. Выполнить упр. 3.4.
2. Выполнить упр. 9.9 и 9.8, предварительно сведя его к типу упр. 9.9.

Difference Tests: *R*, *%*, *Means* → *Difference between two means* (*normal distribution*).

Упражнение 8.6

1. Выполнить упр. 3.5.
2. Выполнить упр. 9.13.

Basic Statistics and Tables → *Correlation Matrices* → *One variable list*…*Options*: *Display r*, *p-levels*, *and N's*.

Упражнение 8.7

1. Повторить в упр. 3.11 расчёт *sl* (с помощью *Probability calculator*).
2. Найти, каков должен быть объем выборки, чтобы критерий уровня $\alpha = 0.01$ фиксировал положительное смещение 0.6 мм с вероятностью 0.9, 0.95?

Industrial Statistics & Six Sigma → *Process Analysis* → *Sampling plans for means*, *proportions*, & *Poisson frequencies* → *Advanced*. *Distribution*: *normal means* → *One-sided* (*right*). Задав $\alpha = 0.01$ и $\beta = \alpha 01 = 1 - \alpha 11 = 1 - w = 0.1$, найдём *N*.

3. Рассчитать функцию мощности критерия уровня $\alpha = 0.01$. В частности, $w(0.01, 40.6) = ?$ Изменив *N* на нужное, по графику *Operating characteristic curve* найти мощность *w* (рассчитываемое процедурой значение β не всегда точно!).
4. Построить график мощности $w(0.01, 40.6)$ как функцию от *N* по 4÷5 точкам.

Выполнить упр. 9.16, 9.17.

Упражнение 8.8

Повторить пример 3.7.

Power Analysis → *Power Calculation* → *One Mean*, *t-Test* → 1-*tailed* (*Mu* ≥ *Mu*0).

1. Найти чувствительность критерия при $a_1 = 4$.

2. Построить (сохраняя остальные параметры, как в задании 1) графики: $w(N)$, $w(\alpha)$, $w(a1)$. Последний – с помощью выражения w через стандартизированный эффект E_s. Чему равно $w(a1 = 2)$?

3. Найти объём выборки, при котором t-критерий уровня 0.0035 будет иметь против альтернативы $N(4, 2.366)$ мощность 0.8.

4. Сравнить мощность $w(0.0035, 4, 10)$ t-критерия с мощностью критерия при известной дисперсии (совпадающей с её оценкой).

5. Подтвердить результат мощности $w(0.0035, 4, 10)$ t-критерия ручным расчётом с использованием нецентрального распределения Стьюдента.

Power Analysis → *Probability Distributions* → *Noncentral t Distribution*.

Выполнить упр. 9.18, 9.19.

Упражнение 8.9

В упр. 3.5 выполнить задания.

Power Analysis → *Power Calculation* → *One Correlation, t-Test.*

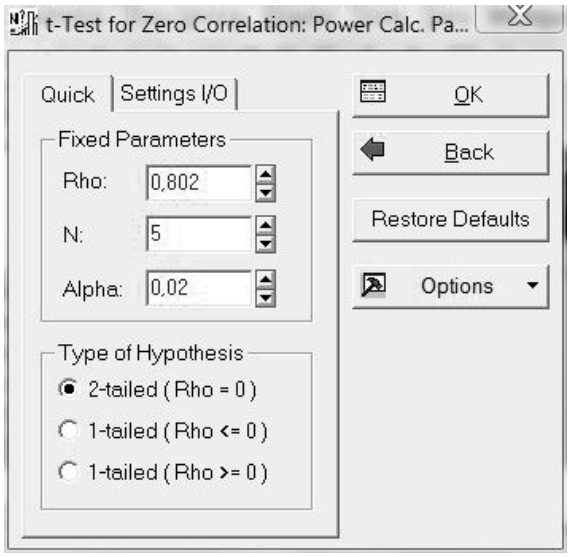
1. Найти, какую мощность имел бы критерий уровня 0.02 против альтернативы, совпадающей с оценкой?

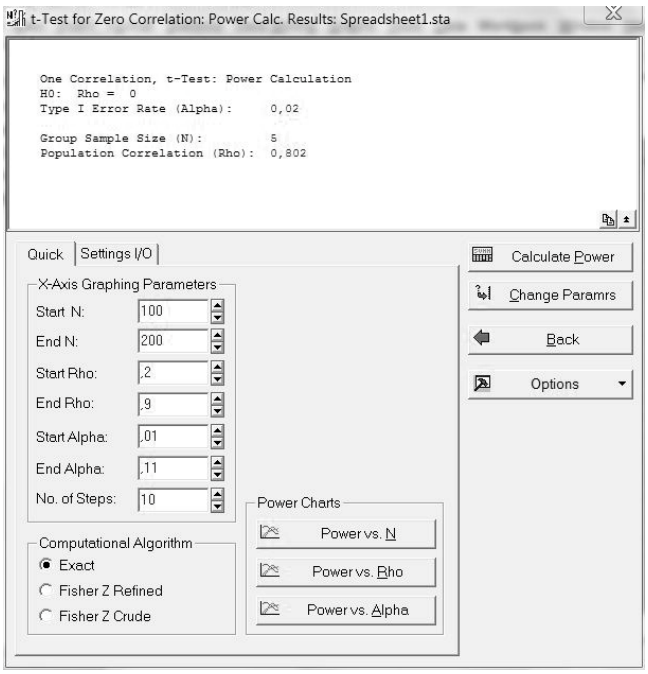
Calculate Power, ответ: 0.2334. См. рисунок.

2. По графику функции чувствительности определить, какое, предположительно, было истинное ρ (чтобы вероятность получить имеющиеся доводы против H_0 была > 0.8)? Использовать точное распределение коэффициента корреляции.

Устанавливаем в первом диалоге альфа, равное 0.104, по графику *Power vs. Rho* находим ответ: > 0.89.

Выполнить упр. 9.14.





Упражнение 8.10

Повторить пример 4.3.

Statistics → *Industrial statistics* & *six* σ → *Process analyses* → *sampling plans*… → *distribution* (*normal means*) → *test criterion* (*one-sided* (*left*) *test*) → α, β, a_0, a_1, σ → *OK*→ *sequential sampling plans* . Подготавливаем данные: *var* 5, 4, 5, 4, 3 (удалить пустые строки) → *variable containing data* → *plot.*

Учесть, что на графике по оси ординат откладывается Δ_m – накопленное отклонение наблюдений от a_0.

Упражнение 8.11

Повторить пример 5.1 – сравнение произвольной и ортогональной матрицы плана в регрессионном анализе. Оценить в обоих случаях коэффициенты регрессии и найти их стандартные ошибки. Подтвердить оценки во втором случае вручную по формулам (54). Для обоих случаев построить графики оценок функции регрессии на диаграмме рассеяния вместе с границами 95 % интервалов для предсказаний. Дать точечные и интервальные прогнозы при $x = 7$. Меню *Statistics*

→ *Multiple Regression* … Построение графика в меню: *Graphs* → *Scatterplots* → *Quick* → *Regression bands* → *prediction*.

Упражнение 8.12

var 1	*var* 2
4	1
3	1
4	1
4	1
4	1
4	2
2	2
2	2
3	3
3	3

Повторить пример 5.2. Подготовить данные в следующей форме.

Statistics → *ANOVA*→ *One-way ANOVA* → *dependent* (*var* 1); *categorical factor* (*var* 2) → *OK*. *Means* → *Observed weighted* (выборочные средние в группах). *Summary* → *between effects* → *coefficients* (среднее средних и эффекты по числу групп – 1). *All effects* (уровень значимости данных *sl* против гипотезы H_0).

Упражнение 8.13

Повторить пример 5.3. Подтвердить проверку оценок $\widehat{\beta}_0$, $\widehat{\beta}_1$, проверить $\widehat{\beta}_2$. Объяснить значение 14.834 у вершины куба на рисунке в том примере с помощью найденных оценок эффектов. Включить незначимые факторы в ошибку. Предсказать прочность ткани при разнообразных сочетаниях уровней факторов через диалог *Predict dependent variable values* (где требуются некодированные значения уровней, которые можно найти, кликнув правой кнопкой на имя фактора, *Variable specs*. → *Text labels*). Опять получить значение 14.834. В том же диалоге предсказать прочность ткани при значениях факторов, отличных от значений, задающих в экспериментах уровни. В частности, при значениях *POLYSULF* = 7, *TIME* = 36, *TEMPERATURE* = 200 получить значение прочности 33.67.

Упражнение 8.14

Повторить пример 5.5. Объяснить связь оценённых коэффициентов регрессионной и дисперсионной моделей. Почему коэффициенты у *POLYSULF* совпадают? Сравнить результаты с результатами из примера 5.3, объяснить улучшение коэффициента детерминации. Оставить только 6 первых экспериментов (каков R^2?).

Провести анализ 1/8 дробного факторного эксперимента при выборе произвольного сочетания уровней. Результаты сравнить с результатами при ортогональном планировании.

Упражнение 8.15

Повторить пример 5.7 анализа плана латинского квадрата. Вручную подтвердить средние и эффекты уровней фактора *DRIVER*.

9. Индивидуальная самостоятельная работа (типовой расчёт)

Исходные данные к упражнениям приведены в приложении. В упр. 9.2–9.3, 9.5–9.11, 9.13–9.19 считать, что выборки *A* и *B* значений случайных величин ξ_1 и ξ_2 взяты из генеральной совокупности с двумерным нормальным распределением.

Упражнение 9.1

По выборке *A* значений случайной величины ξ_1 найти несмещённые оценки математического ожидания $M\xi_1$ и дисперсии $D\xi_1$, оценки стандартного отклонения σ и медианы [4]. Какова стандартная ошибка оценки математического ожидания?

Упражнение 9.2

По выборке *A* найти доверительные интервалы для математического ожидания и дисперсии случайной величины ξ_1 при доверительной вероятности γ.

Упражнение 9.3

Считая выборку *A* пробной, определить минимальный объем выборки для нахождения доверительного интервала 2Δ среднего значения случайной величины ξ_1 при доверительной вероятности γ_1:

$$\gamma_1 = \begin{cases} 0.8, & V \le 5; \\ 0.9, & 5 < V \le 10; \\ 0.95, & 10 < V \le 15; \\ 0.98, & 15 < V \le 25; \end{cases} \qquad \Delta = \begin{cases} V+1, & V \le 5, \\ V-2, & 5 < V \le 10, \\ V/2, & 10 < V \le 15, \\ V/4, & 15 < V \le 25, \end{cases}$$

где *V* – номер варианта.

Упражнение 9.4

Из партии b изделий извлечена бесповторная выборка объёмом n штук, из которых m оказались дефектными. Найти оценку процента дефектных изделий в партии и её стандартную ошибку.

Упражнение 9.5

По выборке A при уровне значимости α проверить гипотезу H_0: $M\xi_1 = a$ при альтернативной гипотезе H_1: $M\xi_1$ * a. Каков уровень значимости данных?

Упражнение 9.6

По выборке A при уровне значимости α проверить гипотезу H_0: $D\xi_1 = b$ при альтернативной гипотезе H_1: $D\xi_1 \neq b$.

Упражнение 9.7

По выборкам A и B найти доверительный интервал для разности средних ($M\xi_2$ - $M\xi_1$) при доверительной вероятности γ, считая дисперсии равными.

Упражнение 9.8

По выборкам A и B при уровне значимости α_1 проверить гипотезу H_0: $M\xi_2$ - $M\xi_1 = c$ при альтернативной гипотезе H_1: $M\xi_2$ - $M\xi_1 \neq c$, считая дисперсии равными. Каков уровень значимости данных?

Упражнение 9.9

По выборкам A и B при уровне значимости 0.5 проверить гипотезу H_0: $M\xi_1 =$ $= M\xi_2$ при альтернативной гипотезе H_1: $M\xi_1 \neq M\xi_2$, считая дисперсии равными. Каков уровень значимости данных?

Упражнение 9.10

По выборкам A и B при уровне значимости α проверить гипотезу о равенстве дисперсий H_0: $\sigma_1^2 = \sigma_2^2$ при альтернативной гипотезе H_1: $\sigma_2^2 > \sigma_1^2$. Каков уровень значимости данных?

Упражнение 9.11

По выборкам A и B при уровне значимости α проверить гипотезу о равенстве дисперсий H_0: $\sigma_1^2 = \sigma_2^2$ при альтернативной гипотезе H_1: $\sigma_1^2 \neq \sigma_2^2$. Каков уровень значимости данных?

Упражнение 9.12

В серии из $100n$ выстрелов по мишени наблюдалось $100m$ попаданий. Найти приближенный доверительный интервал для вероятности p попадания в мишень при доверительной вероятности γ.

Упражнение 9.13

Пусть выборка A – наблюдённые значения случайной величины ξ_1, а такое же число первых элементов выборки B – соответствующие наблюдённые значения случайной величины ξ_2. Найти точечную оценку коэффициента корреляции Пирсона ρ величин ξ_1 и ξ_2. Каков уровень значимости данных против гипотезы H_0: $\rho = 0$ (в пользу гипотезы H_1: $\rho \neq 0$)?

Упражнение 9.14

По условию упр. 9.13 найти, какую мощность имел бы критерий уровня α_1 против альтернативы, совпадающей с оценкой? Использовать точное распределение коэффициента корреляции.

Упражнение 9.15

По условию упр. 9.13 найти оценку линейной функции регрессии ξ_2 на ξ_1. Построить её график с 95 % доверительными интервалами для предсказаний. Найти точечное и интервальное предсказание величины ξ_2 при значении ξ_1, равном $\max\{A\} + 3$.

Упражнение 9.16

По выборке A, пользуясь оптимальным критерием Неймана–Пирсона уровня α, проверить гипотезу H_0: $M\xi_1 = a$ при альтернативной гипотезе H_1: $M\xi_1 > a$. Какова мощность критерия при альтернативе H'_1: $M\xi_1 = a + \Delta$, где Δ см. по условию упр. 9.3? Считать стандартное отклонение равным $\sqrt{s^2}$, где s^2 – несмещённая оценка дисперсии по выборке A.

Упражнение 9.17

Из условия упр. 9.16 найти размер выборки, при котором критерий уровня α проверки гипотезы H_0 против альтернативы H'_1 будет иметь мощность γ.

Упражнение 9.18

По выборке A, пользуясь t-критерием Стьюдента уровня α, проверить гипотезу H_0: $M\xi_1 = a$ при альтернативной гипотезе H_1: $M\xi_1 > a$. Какова мощность критерия при альтернативе H'_1: $M\xi_1 = a + \Delta$, где Δ см. по условию упр. 9.3 и стандартное отклонение равно $\sqrt{s^2}$. Здесь s^2 – несмещённая оценка дисперсии по выборке A? Сравнить результаты с результатами упр. 9.16. Сделать выводы.

Упражнение 9.19

Из условия упр. 9.18 найти размер выборки, при котором критерий уровня α проверки гипотезы H_0 против альтернативы H'_1 будет иметь мощность γ. Сравнить результат с результатом упр. 9.17. Сделать выводы.

Упражнение 9.20

Производится проверка идентичности работы однотипного оборудования (машин) по выпуску некоторых изделий. Собраны выборочные данные о контролируемом параметре изделий, выпущенных на различных машинах (выборки *A, B, C* (и *D* для вариантов 13÷25)). Оценить эффекты уровней. Каков *sl* данных против гипотезы об отсутствии эффектов уровней? Можно ли смешивать изделия в однородную партию (принять решение на уровне α)?

Упражнение 9.21

Выполняется в пакете *STATISTICA*.

Спланировать и провести эксперимент $2^{(5-2)}$ для чётных вариантов и $2^{(4-1)}$ для нечётных. Чтобы была возможность регрессионного анализа, по крайней мере один фактор должен за уровнями иметь количественные значения (см. ниже тепло и жарко). Сгенерировать матрицу плана. Скопировать ее в файл данных. Перекодировать в *Variable Specs*…, *Text Labels*… Фактические значения уровней факторов должны быть в *Numeric* (например, 20 и 30), а тепло и жарко – в *Text Labels*. Для этого сначала уровням необходимо приписать текстовые метки (тепло и жарко), а затем заменить (-1 и +1) на (20 и 30). Для чисто качественного фактора, например «пол», должно быть соответственно (-1 и +1), (мужчина и женщина). Провести эксперимент (собрать значения результирующей переменной). Оценить эффекты уровней, проверить их значимость. Составить и записать итоговую дисперсионную модель, описать подробно эффекты уровней (для какого больше и насколько), визуально проверить нормальность ошибок, указать оптимальное сочетание уровней. По итоговой дисперсионной модели вручную и с помощью регрессионной модели в *STATISTICA* получить одинаковый прогноз результирующей величины для нового объекта (желательно с неиспользовавшимся сочетанием уровней и по возможности сравнить его с фактическим значением). В *STATISTICA* получить прогноз результирующей величины при значе-

ниях факторов, если возможно, отличных от значений, задающих в экспериментах уровни. По возможности сравнить его с фактическим значением. Сделать выводы.

Упражнение 9.22

Выполняется в пакете *STATISTICA*.

Подобрать данные и провести анализ плана латинского или греко-латинского квадрата (3x3) или (4x4). Возможно продолжение дробного эксперимента $2^{(k-p)}$ из упр. 9.28 с расширением числа уровней для важнейших факторов. Оценить эффекты уровней, выявить значимые факторы. Составить и записать итоговую модель, описать подробно эффекты уровней (для какого больше и насколько), указать оптимальное сочетание уровней.

Упражнение 9.23

Собираясь утром на работу, вы принимаете решение пойти без плаща ($a = 1$), надеть плащ ($a = 2$). Состояние погоды днём может быть: сухо ($\theta = 1$), дождь ($\theta = 2$), ливень ($\theta = 3$). Матрица потерь L. Выглянув в окно, можно увидеть, что с утра ясно ($x = 1$), облака ($x = 2$), дождевые тучи ($x = 3$). Известно условное распределение x в зависимости от θ $p(x|\theta)$.

Найти минимаксную решающую функцию и минимаксный риск.

Считая справедливым априорное распределение $p(\theta)$, найти байесовскую решающую функцию и минимальный байесовский риск.

Упражнение 9.24

Для контроля было взято 200 узлов, собранных на ученическом конвейере. Они распределились по числу пропущенных при сборке операций следующим образом:

Число операций	0	1	2	3	4	5	6	7
Число узлов	22	71	45	32	16	8	4	2

Согласуются ли эти данные по критерию χ^2 при уровне значимости α с распределением Пуассона с параметром d для случайного числа пропущенных операций? Каков sl? Указание: последние 2 интервала следует объединить.

Упражнение 9.25

Используя предельное распределение Колмогорова, найти приближённо уровень значимости выборки A против гипотезы H_0. Дать его традиционную интерпретацию. В вариантах 1÷8, 14, 15, 17, 18, 20 считать H_0 – равномерное распределение генеральной совокупности на отрезке от минимального значения выборки до максимального. В вариантах 9÷13, 16, 19, 21÷25 считать H_0 – экспоненциальное распределение с параметром, совпадающим с оценкой по выборке.

Упражнение 9.26

Проверить нормальность объединённой выборки A, B и C по критериям Колмогорова (используя предельное распределение) и Шапиро–Уилка на уровне значимости α. Указать sl и «да – нет» гипотезе нормальности (H_0) для обоих критериев.

Упражнение 9.27

Пользуясь критерием серий, проверить гипотезу случайности кортежа: выборка A, продолженная упорядоченной в порядке возрастания выборкой B (т. е. оценить уровень значимости и дать его традиционную интерпретацию). Каков уровень значимости, рассчитанный по статистике для большой выборки?

Упражнение 9.28

По ранговому критерию Уилкоксона–Манна–Уитни проверить на уровне значимости (1 - γ) гипотезу однородности выборок A и B при двусторонней альтернативе. Каков уровень значимости, рассчитанный по нормальной аппроксимации статистики критерия?

Упражнение 9.29

По условию упр. 9.13 найти оценку коэффициента ранговой корреляции Спирмена. Аппроксимируя точное распределение распределением Стьюдента, проверить на уровне значимости (1 - γ) гипотезу H_0: $\rho_s(\xi_1, \xi_2) = 0$ против односторонней альтернативы, на которую указывают данные. Каков достигаемый уровень значимости?

Библиографический список

1. Айвазян С. А. Прикладная статистика и основы эконометрики / С. А. Айвазян, В. С. Мхитарян. М. : ЮНИТИ, 2001. 1022 с.

2. Андронов А. М. Теория вероятностей и математическая статистика / А. М. Андронов, Е. А. Копытов, Л. Я. Гринглаз. СПб. : Питер, 2004. 461 с.

3. Большев Л. Н. Таблицы математической статистики / Л. Н. Большев, Н. В. Смирнов. М. : Наука, 1983. 416 с.

4. Бородачёв С. М. Имитационное моделирование в экономике : учебное пособие / С. М. Бородачёв. Екатеринбург : УрФУ, 2010. 77 с.

5. Бородачёв С. М. Многомерные статистические методы : учебное пособие / С. М. Бородачёв. Екатеринбург : ГОУ ВПО УГТУ-УПИ, 2009. 74 с.

6. Бородачёв С. М. Эконометрика : учебное пособие / С. М. Бородачёв. Екатеринбург : УрФУ, 2011. 76 с.

7. Ивченко Г. И. Математическая статистика / Г. И. Ивченко, Ю. И. Медведев. М. : Высш. шк., 1984. 248 с.

8. Ликеш И. Основные таблицы математической статистики. / И. Ликеш, Й. Ляга. М. : Финансы и статистика, 1985. 356 с.

9. Пугачев В. С. Теория вероятностей и математическая статистика / В. С. Пугачев. М. : ФИЗМАТЛИТ, 2002. 496 с.

10. Справочник по прикладной статистике : в 2 т. / ред. Э. Ллойд, У. Ледерман. М. : Финансы и статистика, 1990. Т. 1. 510 с.; Т. 2. 526 с.

11. Bickel P. J. Some asymptotic theory for the bootstrap / P. J. Bickel, D. A. Freedman. Ann. Stat. 1981. Vol. 9. P.1196–1217.

12. StatSoft, Inc. (2001) [Электрон. ресурс] : электронный учебник по статистике. М. : StatSoft. Режим доступа: http://www.statsoft.ru/home/textbook/default.htm.

Приложение

Исходные данные к упр. 9.1÷9.20, 9.24÷9.29.

Для каждого варианта указаны выборки *A*, *B*, *C* (и *D*), а также столбец П параметров: γ, α, *a*, *, *b*, α_1, *c*, *n*, *m*, *d*.

Вариант 1				Вариант 2				Вариант 3			
A	*B*	*C*	П	*A*	*B*	*C*	П	*A*	*B*	*C*	П
40	51	28	0.95	41	34	67	0.9	88	70	50	0.5
38	44	69	0.1	47	38	89	0.025	100	92	215	0.05
37	47	132	35	46	41	141	49	94	110	139	90
52	24	45	≠	41	41	10	<	99	113	321	>
42	43	109	20	47	50	139	100		110		8
	34	54	0.05	49	59	49	0.002				0.2
	60		13		45	126	12				19
			7		23		8				5
			5				6				4
			2				1.82				1.87

Вариант 4				Вариант 5				Вариант 6			
A	*B*	*C*	П	*A*	*B*	*C*	П	*A*	*B*	*C*	П
-26	-30	-54	0.8	45	57	167	0.95	-14	-30	-92	0.9
-22	-20	-60	0.1	46	49	58	0.001	-17	-31	-37	0.1
-28	-21	-47	-16.5	45	46	36	47	-23	-28	-24	-27
-23	-18	-31	≠	44	47	135	<	-22	-17	-64	>
-16	-12	-30	90	43	43	111	30	-24	-22	-15	150
-13	-18	-18	0.1	46	43	163	0.002	-30	-19	-85	0.05
-18	-19	4	-3		45	72	-9	-31	-34	-67	5
	-18	-32	9		35		8		-40	1	9
	-10		6				7		-9		5
			1.98				1.74				1.91

Вариант 7				Вариант 8				Вариант 9			
A	*B*	*C*	П	*A*	*B*	*C*	П	*A*	*B*	*C*	П
-16	-27	-34	0.5	58	47	146	0.95	113	97	108	0.8
-17	-24	-23	0.1	57	49	53	0.001	120	104	108	0.1
-14	-18	-40	-10	60	59	146	63	113	105	149	109
-8	-15	-26	≠	58	57	68	<	109	103	123	>
-20	-5	-33	9		65	194	200	111	122	131	14
	-7	-14	0.2		41		0.1	102	128	232	0.05
	-30		4				2	116	113		9
			7				6				10
			4				4				7
			1.99				1.77				1.97

Вариант 10				Вариант 11				Вариант 12			
A	*B*	*C*	П	*A*	*B*	*C*	П	*A*	*B*	*C*	П
49	57	59	0.8	104	87	225	0.9	35	38	57	0.95
51	58	122	0.1	93	91	146	0.001	65	33	90	0.1
46	50	61	54	98	90	107	105	50	37	43	37
49	51	59	≠	96	95	101	<	46	65	74	>
56	46	54	70	100	94	203	9		78	98	61
	39	79	0.05	100	122	122	0.2		66	119	0.1
	67		-6	96	158	219	-4		31		20
			7		130		8				7
			5				7				4
			1.92				1.8				2.1

Вариант 13					Вариант 14				
A	*B*	*C*	*D*	П	*A*	*B*	*C*	*D*	П
127	175	208	160	0.5	48	31	33	18	0.95
127	157	241	136	0.025	61	38	39	12	0.01
132	103	109	74	115	60	44	48	-5	63
143	139	179	99	≠	61	52	80	45	<
121	122	139	46	19	63	66	77	30	10
	97			0.2	65	65	80	50	0.2
				18	53	76	96	49	7
				6		81			8
				5					4
				1.82					1.83

Вариант 15					Вариант 16				
A	*B*	*C*	*D*	П	*A*	*B*	*C*	*D*	П
61	66	106	42	0.8	70	83	85	76	0.9
65	49	70	27	0.01	79	77	77	64	0.1
66	79	124	62	60	77	76	108	54	82
71	71	71	67	>	71	66	75	43	≠
64	73	117	70	120		60	64	56	6
64	83			0.1		90			0.05
				-6					13
				10					6
				6					4
				1.84					1.85

Вариант 17					Вариант 18				
A	*B*	*C*	*D*	П	*A*	*B*	*C*	*D*	П
69	100	130	74	0.95	58	50	51	35	0.5
97	142	151	129	0.05	56	54	61	42	0.005
92	102	107	65	100	53	51	56	30	50
91	54	73	47	<	53	56	63	31	>
91	89	114	48	60	56	53	74	50	115
105	101			0.5	52	70	91	55	0.01
				-4	52	68			-8
				11					10
				6					7
				1.86					1.73

Вариант 19					Вариант 20				
A	*B*	*C*	*D*	П	*A*	*B*	*C*	*D*	П
203	200	220	183	0.5	141	155	156	153	0.8
214	233	287	219	0.1	128	138	145	128	0.1
190	169	245	110	180	134	149	149	127	138
195	194	215	149	≠	140	130	145	120	<
	170			1000	126	120	133	90	20
				0.2	139	137			0.1
				9					-5
				11					15
				4					6
				1.88					1.89

Вариант 21					Вариант 22				
A	*B*	*C*	*D*	П	*A*	*B*	*C*	*D*	П
93	93	93	84	0.9	35	65	71	62	0.8
92	94	110	89	0.1	34	45	55	42	0.1
103	93	101	90	91	47	34	35	30	35
95	92	105	90	>	46	53	61	48	≠
	90	95	83	10		68	72	60	19
	65	81	54	0.02		70	74	69	0.01
	87			5		70			-9
				8					15
				3					7
				1.9					1.94

Вариант 23					Вариант 24					Вариант 25				
A	*B*	*C*	*D*	П	*A*	*B*	*C*	*D*	П	*A*	*B*	*C*	*D*	П
93	99	107	90	0.5	84	83	83	82	0.9	47	55	56	54	0.8
96	97	100	94	0.1	83	83	85	82	0.25	59	55	57	54	0.25
94	95	99	91	96	83	83	84	82	82.5	61	61	63	60	53
95	94	98	93	<	83	82	82	80	>	60	62	63	61	≠
	90	90	89	15	83	82	85	81	8		67	69	65	20
	91	95	81	0.1	81	81	83	79	0.04		75	77	74	0.2
	89			3		85	85	84	2		64			0
				10		84			6					9
				6		86			2					7
				1.96					1.95					2.2

Исходные данные к упр. 9.23 указаны в форме:

$p(\theta = 1)$	$p(\theta = 2)$	$p(\theta = 3)$	$p(x = 1\|\theta = 1)$	$p(x = 1\|\theta = 2)$	$p(x = 1\|\theta = 3)$
L_{11}	L_{12}	L_{13}	$p(x = 2\|\theta = 1)$	$p(x = 2\|\theta = 2)$	$p(x = 2\|\theta = 3)$
L_{21}	L_{22}	L_{23}	$p(x = 3\|\theta = 1)$	$p(x = 3\|\theta = 2)$	$p(x = 3\|\theta = 3)$

Вариант 1

0.4	0.34	0.26	0.579	0.2	0.109
0	1.8	6	0.252	0.47	0.305
3	4	2	0.169	0.33	0.586

Вариант 2

0.14	0.72	0.14	0.254	0.2	0.149
0	2.2	6	0.359	0.47	0.236
5.1	4	2	0.387	0.33	0.615

Вариант 3

0.44	0.23	0.33	0.356	0.2	0.15
0	2.8	6	0.452	0.47	0.383
4.6	4	2	0.192	0.33	0.467

Вариант 4

0.1	0.72	0.18	0.444	0.2	0.109
0	3.7	6	0.423	0.47	0.451
4.1	4	2	0.133	0.33	0.44

Вариант 5

0.34	0.34	0.32	0.42	0.2	0.107
0	4	6	0.355	0.47	0.245
4.4	4	2	0.225	0.33	0.648

Вариант 6

0.3	0.39	0.31	0.527	0.2	0.065
0	1.7	6	0.328	0.47	0.242
5.3	4	2	0.145	0.33	0.693

Вариант 7

0.43	0.41	0.16	0.543	0.2	0.132
0	1.6	6	0.284	0.47	0.247
5.9	4	2	0.173	0.33	0.621

Вариант 8

0.28	0.33	0.39	0.586	0.2	0.124
0	4.4	6	0.35	0.47	0.259
4.4	4	2	0.064	0.33	0.617

Вариант 9

0.31	0.34	0.35	0.294	0.2	0.116
0	3.3	6	0.294	0.47	0.453
3.1	4	2	0.412	0.33	0.431

Вариант 10

0.43	0.29	0.28	0.572	0.2	0.075
0	1.6	6	0.263	0.47	0.2
3.9	4	2	0.165	0.33	0.725

Вариант 11

0.4	0.34	0.26	0.425	0.2	0.094
0	1.5	6	0.373	0.47	0.409
4.7	4	2	0.202	0.33	0.497

Вариант 12

0.38	0.46	0.16	0.289	0.2	0.068
0	3	6	0.228	0.47	0.337
4.9	4	2	0.483	0.33	0.595

Вариант 13

0.19	0.59	0.22	0.489	0.2	0.113
0	4.6	6	0.456	0.47	0.335
5.8	4	2	0.055	0.33	0.552

Вариант 14

0.17	0.56	0.27	0.484	0.2	0.074
0	3.3	6	0.348	0.47	0.381
4.9	4	2	0.168	0.33	0.545

Вариант 15

0.42	0.31	0.27	0.517	0.2	0.141
0	3.5	6	0.295	0.47	0.418
5.2	4	2	0.188	0.33	0.441

Вариант 16

0.44	0.41	0.15	0.468	0.2	0.058
0	1.4	6	0.323	0.47	0.392
3.9	4	2	0.209	0.33	0.55

Вариант 17

0.4	0.25	0.35	0.301	0.2	0.137
0	1.6	6	0.435	0.47	0.29
4.4	4	2	0.264	0.33	0.573

Вариант 18

0.36	0.51	0.13	0.515	0.2	0.073
0	1.3	6	0.348	0.47	0.476
4.8	4	2	0.137	0.33	0.451

Вариант 19

0.48	0.37	0.15	0.261	0.2	0.063
0	3.2	6	0.217	0.47	0.486
4.1	4	2	0.522	0.33	0.451

Вариант 20

0.39	0.51	0.1	0.318	0.2	0.121
0	4.5	6	0.302	0.47	0.265
3.4	4	2	0.38	0.33	0.614

Вариант 21

0.31	0.35	0.34	0.305	0.2	0.064
0	4.2	6	0.22	0.47	0.321
4.1	4	2	0.475	0.33	0.615

Вариант 22

0.22	0.44	0.34	0.548	0.2	0.078
0	3.2	6	0.308	0.47	0.469
4.7	4	2	0.144	0.33	0.453

Вариант 23

0.32	0.56	0.12	0.607	0.2	0.102
0	1.9	6	0.244	0.47	0.228
3.7	4	2	0.149	0.33	0.67

Вариант 24

0.23	0.55	0.22	0.527	0.2	0.145
0	1.2	6	0.278	0.47	0.36
5.8	4	2	0.195	0.33	0.495

Вариант 25

0.33	0.36	0.31	0.491	0.2	0.094
0	2.6	6	0.373	0.47	0.202
4.6	4	2	0.136	0.33	0.704

Оглавление

ВВЕДЕНИЕ 3
1. ТОЧЕЧНЫЕ ОЦЕНКИ 5
1.1. Примеры статистик и их распределений 6
1.2. Оценка 7
1.3. Оценка математического ожидания и дисперсии 8
1.4. Теорема Фишера 10
1.5. Оценка математического ожидания и ковариационной матрицы случайного вектора 12
1.6. Оценка стандартного отклонения через размах выборки 14
1.7. Оценка параметра *p* гипергеометрического распределения 15
1.8. Оценка параметра *p* биномиального распределения 17
1.9. Оценка функции распределения 17
1.10. Метод максимального правдоподобия (МП) 18
1.11. Неравенство Рао–Крамера. Эффективные оценки 20
1.12. Достаточные статистики 22
1.13. Теоретико-игровой подход к оцениванию 24
Упражнения 26
2. ИНТЕРВАЛЬНЫЕ ОЦЕНКИ (ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ) 31
2.1. Доверительный интервал для математического ожидания нормальной случайной величины 31
2.2. Доверительный интервал для дисперсии нормальной случайной величины 34
2.3. Односторонние доверительные интервалы 36
2.4. Решение двух основных задач в теории ошибок измерений 37
Упражнения 38
3. СТАТИСТИЧЕСКАЯ ПРОВЕРКА ГИПОТЕЗ 40
3.1. Общая схема проверки гипотез 42

3.2. Проверка гипотезы о математическом ожидании нормальной случайной величины 43
3.3. Критерий равенства двух математических ожиданий 44
3.4. Проверка гипотезы о коэффициенте корреляции 46
3.5. Критерий равенства двух дисперсий 46
3.6. Связь проверки гипотез с интервальным оцениванием 48
3.7. Мощность критерия 48
3.8. Мощность критерия проверки гипотезы о математическом ожидании нормальной случайной величины 49
3.9. Различные критерии на одних и тех же данных 53
3.10. Критерии при сложной основной гипотезе 55
Упражнения 56
4. СВЯЗЬ ПРОВЕРКИ ГИПОТЕЗ С РАСПОЗНАВАНИЕМ ОБРАЗОВ 60
4.1. Построение моделей распознавания без информации об априорных вероятностях. Лемма Неймана–Пирсона 60
4.2. Наиболее мощный критерий уровня α проверки гипотезы о математическом ожидании нормальной случайной величины при известной дисперсии 62
4.3. Последовательный анализ 63
5. ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА И ДИСПЕРСИОННЫЙ АНАЛИЗ 68
5.1. Оптимальный выбор матрицы плана 68
5.2. Однофакторный дисперсионный анализ 70
5.3. Полный многофакторный эксперимент 74
5.4. Дробные реплики 78
6. СТАТИСТИЧЕСКИЕ РЕШАЮЩИЕ ПРАВИЛА 85
6.1. Минимаксное решающее правило 85
6.2. Байесовское решающее правило 87
6.3. Выборочный контроль качества продукции 93

7. НЕПАРАМЕТРИЧЕСКИЕ МЕТОДЫ 95
7.1. Критерий однородности Уилкоксона–Манна–Уитни 95
7.2. Критерий Краскела–Уоллиса 96
7.3. Бутстреп-процедуры 99
8. ЛАБОРАТОРНЫЙ ПРАКТИКУМ 102
9. ИНДИВИДУАЛЬНАЯ САМОСТОЯТЕЛЬНАЯ РАБОТА (ТИПОВОЙ РАСЧЁТ) 108
БИБЛИОГРАФИЧЕСКИЙ СПИСОК 116
ПРИЛОЖЕНИЕ 117

Учебное издание

Бородачёв Сергей Михайлович

МЕТОДЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Редактор *И. В. Меркурьева*

Компьютерная вёрстка *авторская*

Подписано в печать 16.11.2012. Формат 60 x 84 1/16.

Бумага писчая. Плоская печать. Усл. печ. л. 7.32.

Уч.-изд. л. 5.1. Тираж 51 экз. Заказ .

Редакционно-издательский отдел УрФУ

620002, Екатеринбург, ул. Мира, 19

rio@ustu.ru

Ризография НИЧ УрФУ

620002, Екатеринбург, ул. Мира, 19