Содержание

1 Предельные теоремы и законы больших чисел **3**

1.1 Сходимость последовательностей случайных величин (по вероятности, почти наверное, в среднем, по распределению) 3

1.2 Теорема о связи между сходимстью по вероятности и сходимости к константе 4

1.3 Определение закона больших чисел. Теорема ЗБЧ для н.о.р. величин (с доказательством с помощью х.ф.) 4

1.4 Центральная предельная теорема 5

2 Вариационные ряды и их характеристики **6**

2.1 Генеральная и выборочная совокупности, объём выборки. 6

2.2 Варианционный ряд, варианта, частота. Виды вариационных рядов. Гистограмма, полигон. 7

2.3 Формулы числовых характеристик. Эмперическая функция распределения (ЭФР). Свойства ЭФР. 9

2.4 Эмперическая функция распределения ЭФР. Свойства ЭФР. . . 10

3 Оценки параметров распределения **11**

3.1 Понятия статистики, оценки, выборочной характеристики. . . . 11

3.2 Несмещённые, состоятельные и эффективные оценки 11

3.3 Теорема о несмещённой состоятельной оценке мат. ожидания . 12

3.4 Дополнительная информация 13

3.5 Теорема о несмещённой и состоятельной оценке дисперсии . . . 13

3.6 Теорема о несмещённой и состоятельной оценке функции распределения 14

3.7 Теорема об эффективной оценке математического ожидания . . 16

3.8 Теорема о единственности эффективной оценки 17

3.8.1 Определение регулярной параметрической модели. Информация Фишера. 18

3.9 Неравенство Рао-Крамера. Построение эффективной по Рао-Крамеру оценки для пуассоновского распределения 19

4 Методы построения оценок **19**

4.1 Метод моментов. Пример (биномиальное распределение) 19

4.2 Функция правдоподобия. Метод максимального правдоподобия. Пример (биномиальное распределение) 21

5 Регрессионный анализ **22**

5.1 Постановка задачи 22

5.2 Построение выборочного уравнения регрессии 24

5.3 Выборочный коэффициент корреляции, его свойства 25

5.4 Проверка гипотезы о значимости выборочного коэффициента корреляции 26
5.5 Теорема Гаусса-Маркова 26
5.6 Статические свойства МНК оценок 27
5.7 Анализ вариации 27

1 Предельные теоремы и законы больших чисел

Рассмотрим последовательность $\{\xi_i\}_{i=1}^{\infty}$ случайных величин с $M\xi_i < \infty$.

1.1 Сходимость последовательностей случайных величин (по вероятности, почти наверное, в среднем, по распределению)

Определение:

Последовательность случанйых величин $\{\xi_i\}_{i=1}^{\infty}$ сходится по вероятности к случайной величине ξ ($\xi_i \to^P \xi$), если $\forall \epsilon > 0$

$$P\{\omega : |\xi_i(\omega) - \xi(\omega)| > \epsilon\} \to_{i\to\infty} 0$$

или

$$P\{\omega : |\xi_i(\omega) - \xi(\omega)| < \epsilon\} \to_{i\to\infty} 1$$

Определение:

Последовательность случанйых величин $\{\xi_i\}_{i=1}^{\infty}$ сходится почти наверное к случайной величине ξ ($\xi_i \to^{\text{П.Н.}} \xi$), если

$$P\{\omega : \xi_i(\omega) \to \xi(\omega)\} = 1$$

или

$$P\{\omega : \xi_i(\omega) \nrightarrow \xi(\omega)\} = 0$$

Определение:

Последовательность случанйых величин $\{\xi_i\}_{i=1}^{\infty}$ сходится по распределению к случайной величине ξ, если последовательность функций распределения $F_{\xi_i}(x)$ слабо сходится к $F_\xi(x)$.

То есть

$$F_{\xi_i}(x) \to^{\omega} F_\xi(x) \Rightarrow P\{\omega : \xi_i(\omega) < x\} \to P\{\omega : \xi(\omega) < x\}$$

Определение:

Последовательность случанйых величин $\{\xi_i\}_{i=1}^{\infty}$ сходится в среднем порядка p к случайной величине ξ, если $M|\xi_i - \xi|^p \to_{i\to\infty} 0$.

1.2 Теорема о связи между сходимстью по вероятности и сходимости к константе

Теорема:

Пусть $(\xi_n \to^d \xi)$. Тогда $(\xi_n \to^P \xi)$.

Доказательсво:

Пусть $P\{\xi = c\} = 1$. Тогда $F_c = \begin{cases} 0, & x \leq c \\ 1, & x > c \end{cases}$

Так как $\xi_n \xrightarrow{d} c$, то $F_n(x) \to F_c(x)$ во всех точках непрерывности.

Рассмотрим $\forall \varepsilon > 0$ $P\{-\varepsilon < \xi_n - c < \varepsilon\} = P\{c - \varepsilon < \xi_n < c + \varepsilon\} \leq P\{c - \frac{\varepsilon}{2} < \xi_n < c + \varepsilon\} =$

$$= F_n(c + \varepsilon) - F_n(c - \tfrac{\varepsilon}{2}) \xrightarrow[n \to \infty]{} F_c(c + \varepsilon) - F_c(c - \tfrac{\varepsilon}{2}) = 1 - 0 = 1.$$

Таким образом, $\forall \varepsilon > 0$ $P\{|\xi_n - c| < \varepsilon\} \xrightarrow[n \to \infty]{} 1$, то есть по определению $\xi_n \xrightarrow{P} c$

1.3 Определение закона больших чисел. Теорема ЗБЧ для н.о.р. величин (с доказательством с помощью х.ф.)

Определение:

Пусть $\{\xi_i\}_{i=1}^{\infty}$ — последовательность случайных величин с $M|\xi_i| < +\infty$. Обозначим $M\xi_i = a < +\infty$.

Говорят, что последовательность ξ_i удовлетворяет закону больших чисел, если $\frac{S_n}{n} \to \frac{A_n}{n}$

$S_n = \xi_1 + \cdots + \xi_n; \quad A_n = a_1 + \ldots a_n.$

То есть среднее арифметическое из наблюдений сходится к среднему из математического ожидания.

ЗБЧ для независимых одинаково распределённых случайных величин:

Пусть $\{\xi_i\}$ — независимы одинаково распределённые случаные величины с $M|\xi_i| = M|\xi_1| < +\infty \quad \forall i$. Тогда $\{\xi_i\}$ удовлетворяют закону больших чисел.

Доказательство:

Пусть $S_n = \xi_1 + \ldots \xi_n$; $M\xi_1 = a$; $A_n = na$.

Применим для док-ва метод характеристических функций:

Рассмотрим случайные величины $\frac{S_n}{n} = \frac{\xi_1}{n} + \ldots + \frac{\xi_n}{n}$:

Характеристическая функция $\varphi_{\frac{S_n}{n}}(t) = \varphi_{\frac{\xi_1}{n}}(t) \cdot \ldots \cdot \varphi_{\frac{\xi_n}{n}}(t) = \left[\varphi_{\frac{\xi_1}{n}}(t)\right]^n = \left[\varphi_{\xi_1}\left(\frac{t}{n}\right)\right]^n$

по свой-ву $\varphi_{a\xi+b}(t) = e^{itb}\varphi_{\xi}(at)$

Разложим х/ф $\varphi_{\frac{\xi_1}{n}}\left(\frac{t}{n}\right)$ в ряд: $\varphi_{\frac{\xi_1}{n}}\left(\frac{t}{n}\right) = 1 + \frac{i\left(\frac{t}{n}\right)}{1!} M\xi_1 + o\left(\frac{t^2}{n^2}\right) = 1 + \frac{ita}{n} + o\left(\frac{t}{n}\right)$

Таким образом, $\varphi_{\frac{S_n}{n}}(t) = \left[1 + \frac{ita}{n} + o\left(\frac{t}{n}\right)\right]^n \xrightarrow[n\to\infty]{} e^{ita}$

Заметим, что $\varphi(t) = e^{ita}$ непрерывна в 0 $\varphi(0) = 1 \Rightarrow \varphi(t)$ — х/ф случайной величины $\xi = a$ с $P = 1$.

Получим, что $\varphi_{\frac{S_n}{n}}(t) \xrightarrow[n\to\infty]{} \varphi_a(t)$. По теореме Леви $\Rightarrow F_{\frac{S_n}{n}}(x) \xrightarrow{w} F_a(x)$, то есть

$\frac{S_n}{n} \xrightarrow{d} a$. По теореме о сходимости к константе это означает, что

$\frac{S_n}{n} \xrightarrow{P} a$.

Таким образом, $\forall \varepsilon > 0 \quad P\left\{\left|\frac{S_n}{n} - a\right| \geq \varepsilon\right\} \xrightarrow[n\to\infty]{} 0$ □

1.4 Центральная предельная теорема

Теорема:

Пусть $\{\xi_i\}$ — последовательность независимых одинаково распределённых случайных величин с $M|\xi_i| < +\infty$, $M\xi_1 = a$; $\quad D\xi_1 = \sigma^2$. Тогда $\forall x \in R \quad P\left\{\frac{S_n - na}{\sqrt{DS_n}} < x\right\} \rightarrow_{n\to\infty} \frac{1}{\sqrt{2\pi}} \in_{-\infty}^{x} e^{-\frac{n^2}{2}} du$. То есть $z_n = \frac{S_n - \mu S_n}{\sqrt{DS_n}}$ сходится по распределению с случаной величиной $z \sim N(0, 1)$.

Доказательство:

Построим х/ф z_n:

$$\varphi_{z_n}(t) = \varphi_{\frac{S_n - MS_n}{\sqrt{DS_n}}}(t) = \frac{\varphi_{(\xi_1 - a) + \ldots + (\xi_n - a)}(t)}{\sqrt{\sigma^2 + \ldots + \sigma^2}} =$$

$$= \varphi_{\frac{\xi_1 - a}{\sigma\sqrt{n}}}(t) \cdot \ldots \cdot \varphi_{\frac{\xi_n - a}{\sigma\sqrt{n}}}(t) = \left[\varphi_{\xi_1 - a}\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n \quad (*)$$

Рассмотрим $(\xi_1 - a)$:

$$M(\xi_1 - a) = 0 \qquad D\xi_1 - a = D\xi_1 + Da = \sigma^2$$

$$D(\xi_1 - a) = M(\xi_1 - a)^2 - [M(\xi_1 - a)]^2 = M(\xi_1 - a)^2$$

Рассмотрим $\varphi_{\xi_1 - a}\left(\frac{t}{\sigma\sqrt{n}}\right)$:

$$\varphi_{\xi_1 - a}\left(\frac{t}{\sigma\sqrt{n}}\right) = 1 + \frac{it}{\sigma\sqrt{n}} M(\xi_1 - a) + \frac{(it)^2}{(\sigma\sqrt{n})^2} M(\xi_1 - a)^2 + o\left(\frac{t^2}{\sigma^2 n}\right) =$$

$$= 1 - \frac{t^2}{\sigma^2 n}\sigma^2 + o\left(\frac{t^2}{\sigma^2 n}\right) = 1 - \frac{t^2}{n} + o\left(\frac{t^2}{\sigma^2 n}\right).$$

Подставим в $(*)$:

$$\varphi_{z_n}(t) = \left[1 - \frac{t^2}{n} + \left(\frac{t^2}{\sigma^2 n}\right)\right]^n \xrightarrow[n \to \infty]{} e^{-\frac{t^2}{2}} = \varphi_z(t) \qquad z \sim N(0,1)$$

Таким образом, показали, что $\varphi_{z_n}(t) \xrightarrow[n\to\infty]{} \varphi_z(t)$. По теореме Леви $F_{z_n}(x) \xrightarrow{\omega} F_z(x)$ □

2 Вариационные ряды и их характеристики

2.1 Генеральная и выборочная совокупности, объём выборки.

Рассмотрим постановку задачи математической статистики: по результатам наблюдения за некоторой случайной величиной ξ требуется сделать выводы о неизвстном законе распределения этой величины $\mathcal{L}(x, \theta)$ либо о неизвестных парамерах $\theta_1, \ldots, \theta_n$ известного распределения.

Пусть ξ — случайная величина с некоторой (теоретической) функцией распределения $F_\xi(x) = P\{\xi < x\}, \quad x \in R$.

Определение:

Совокупность n независимых одинаково распределённых случайных величин $X_1, X_2, \ldots, X_n$ называется выборкой (выборочной совокупностью), извлечённой из распределения случайной величины ξ.

Определение:

Под генеральной совокупностью понимается множество всех возможных значений случайной величины ξ.

Определение:

Объёмом совокупности называется количество всех её элементов, объём выборки или выборочной совокупности обозачается n, генеральной совокупности — N.

2.2 Варианционный ряд, варианта, частота. Виды вариационных рядов. Гистограмма, полигон.

Определение:

Вариационный ряд — это последовательность расположенных в порядке неубывания результатов наблюдения $x_1^* \le \dots \le x_n^*$.

x_i^* — варианта.

n_i — частота появления варинты x_i^* в выборке.

Определение:

Точечным вариационным рядом называется:

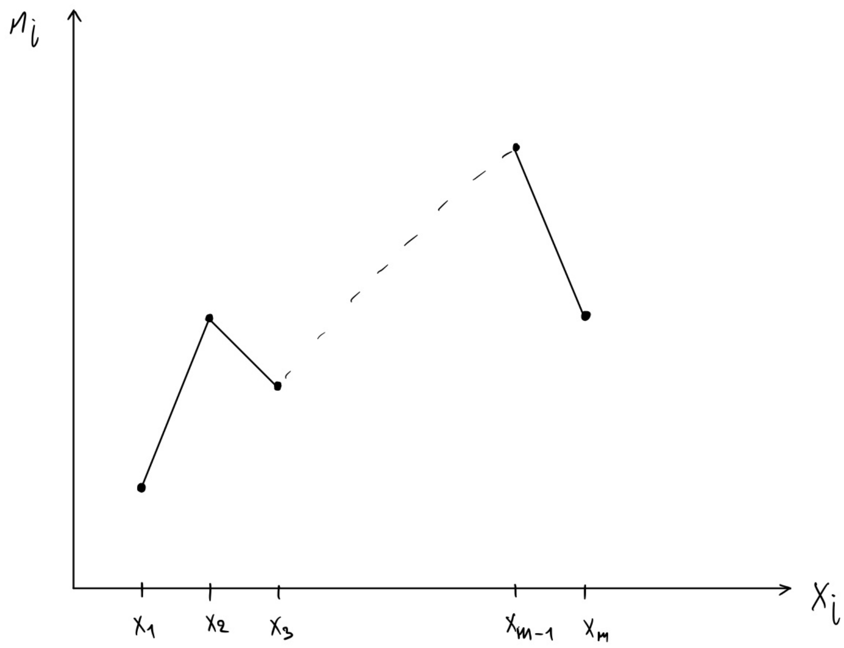
x_i	x_1	x_2	...	x_m
n_i	n_1	n_2	...	n_m

x_i — варианта, n_i — частота соответствующей варианты.

m — количество групп (различных вариант (вариант в таблице)).

$n = \sum\limits_{i=1}^{m} n_i$, где n — объём выборки.

Для графического представления точечных вариационных рядов используется полигон частот — ломанная с вершинами в точках (x_i, n_i).



Определение:

Интервальным вариационным рядом называется:

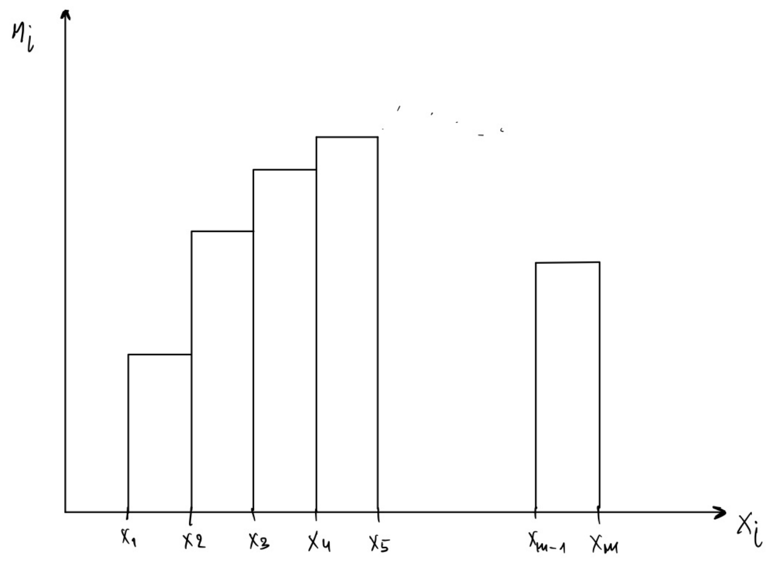
x_i	$[x_1, x_2]$	$(x_2, x_3]$	$\dots$	$(x_m, x_{m+1}]$
n_i	n_1	n_2	$\dots$	n_m

x_i — варианты, n_i — частота.

m — количество групп (интервалов).

$n = \sum\limits_{i=1}^{m} n_i$, где n — объём выборки.

Для графического представления интервальных вариационных рядов используется гистограмма частот — фигура, составленная из прямоугольников, одной стороной которых служат интервалы $(x_i, x_{i+1}]$, а длина второй равна n_i.



2.3 Формулы числовых характеристик. Эмпирическая функция распределения (ЭФР). Свойства ЭФР.

Определение:

Выборочным средним называется величина:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X_i}.$$

Если данные представлены в виде точечного или интервального вариационного ряда, то для вычисления используют формулу:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i * n_i,$$

где m — количество групп в точечном или интервалов в интервальном вариационном ряду, n_i — частота, т.е. количество элементов выборки, принадлежащихх i-той группе или i-тому интервалу, x_i — варианта для точечного ряда и середина i-того интервала для интервального ряда.

Определение:

Выборочной дисперсией (смещённой) называется величина:

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{x})^2.$$

Она характеризует среднее из квадратов отклонений наблюдаемой величины от выборочного среднего. Величина $S = \sqrt{S^2}$ называется выборочным средним квадратическим отклонением (смещённым) величин выборки от выборочного среднего.

Если данные представлены в виде точечного или интервального вариационного ряда, то для вычисления используют формулу:

$$S^2 = \frac{1}{n}\sum_{i=1}^{m}(x_i - \overline{x})^2 n_i.$$

или

$$S^2 = \left(\frac{1}{n}\sum_{i=1}^{m} x_i^2 n_i\right) - \overline{x}^2.$$

Здесь m — количество групп в точечном или интервалов в интервальном вариационных рядах, n_i — частота. т.е. количество элементов выборки, принадежащих i-той группе или i-тому интервалу, x_i — варианта для точечного ряда и середина i-того интервала для интервального ряда.

Определение:

Выборочной дисперсией (несмещённой) называется величина:

$$\overline{\sigma} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{x})^2.$$

Аналогично, величина $\overline{\sigma} = \sqrt{\overline{\sigma}^2}$ называется выборочным несмещённым средним квадратическим отклонением.

Очевидно, что смещённая и несмещённая выборочные дисперсии связаны формулой:

$$\overline{\sigma}^2 = \frac{n}{n-1}S^2.$$

2.4 Эмпирическая функция распределения ЭФР. Свойства ЭФР.

Эмпирическая функция распределения ЭФР

Эмпирической функцией распределения (ЭФР) называется функция

$$\tilde{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} e(x - X_i),$$

где $e(x) = 1$, при $x > 0 \quad e(x) = 0$, при $x \leq 0$.

Таким образом, если $X_i < x$, то $e(x) = 1$, если $X_i \geq x$, то $e(x) = 0$, а сумма $e(x - X_i)$ будет равна количеству элементов выборки, которые приняли значение, строго меньше некоторого $x \in R$.

Пусть $x_1, x_2, \dots, x_n$ — реализация выборки $X_1, X_2, \dots, X_n$, т.е. наблюдавшиеся значения случайной величины ξ.

Обозначим $\mu(x)$ — число элементов выборки, строго меньших $x \in R$. Тогда эмпирическая функция распределения $\tilde{F}_n(x)$ может быть определена как

$$\tilde{F}_n(x) = \frac{\mu(x)}{n}.$$

Свойства ЭФР:

1. $0 \leq \tilde{F}_n(x) \leq 1$, т.к. $0 \leq \mu(x) \leq n$;
2. неубывающая непрерывная слева функция;
3. $\tilde{F}_n(x)$ — ступенчатая функция для всех типов распределений;
4. $\tilde{F}_n(x)$ сходится по распределению к $F_\xi(x)$.

3 Оценки параметров распределения

3.1 Понятия статистики, оценки, выборочной характеристики.

Определение:

Пусть $g(t_1, \dots, t_n)$ — непрерывная функция. Оценкой θ назовём $\tilde{\theta} = g(X_1, \dots, X_2)$. Если $g(X_1, \dots, X_n) = T$ некоторая функция, то T — статистика.

Определение:

Выборочными характеристиками называются функции от наблюдений (точечные оценки), приближённо оценивающие соответствующие числовые характеристики случайной величины.

3.2 Несмещённые, состоятельные и эффективные оценки

Определение:

Оценка $\tilde{\theta}_n$ называется несмещённой оценкой параметра θ, если $M(\tilde{\theta}_n) = \theta$.

Определение:

Оценка $\tilde{\theta}_n$ называется состоятельной оценкой параметра θ, если $\tilde{\theta}_n$ сходится по вероятности к θ.

Определение:

Оценка $\tilde{\theta}_n$ называется эффективной, или оптимальной, или наилучшей несмещённой оценкой с минимальной дисперсией (НОМД), если $M(\tilde{\theta}_n) = \theta$ и $D(\tilde{\theta}_n) = \inf\limits_{\tilde{\theta}_n^*} D\tilde{\theta}_n^*$.

3.3 Теорема о несмещённой состоятельной оценке мат. ожидания

Пусть $\xi \sim L(x, \theta)$, $L(x, \theta)$ — закон распределения известен с точностью до параметра.

$$\overline{\theta} = (\Theta_1, \dots, \Theta_n)$$

По результатам $X_1, \dots, X_n$ наблюдений за ξ требуется построить оценку θ.

Теорема:

Пусть $X_1, \dots, X_n \sim L_\xi(x, \theta)$, где ξ — случайная величина с $M\xi = a < +\infty \quad D\xi = \sigma^2$. Тогда выборочное среднее $\overline{x} = \frac{1}{n}\sum_{i=1}^n x_i$ является несмещённой и состоятельной оценкой $M\xi$.

Доказательство:

$$M\overline{x} = M\left(\frac{1}{n}\sum_{i=1}^n x_i\right) = \frac{1}{n}\sum_{i=1}^n Mx_i =$$

$$= |x_i \text{ — н. о. р. случайной величины} \Rightarrow Mx_i = M\xi \quad \forall i| = \frac{1}{n}\sum_{i=1}^n M\xi =$$

$$= \frac{1}{n} * n * a = a \Rightarrow \overline{x} \text{ — несмещённое}$$

.

Состоятельность $\forall \epsilon > 0 \quad P\{|\overline{x} - a| < \epsilon\} \to_{n\to\infty} 1$

По неравенству Чебышёва:

$$\forall \epsilon > 0 \quad P\{|\overline{x} - a| < \epsilon\} \geq 1 - \frac{D\overline{x}}{\epsilon^2} = 1 - \frac{1}{\epsilon^2} D\left(\frac{1}{n}\sum_{i=1}^n x_i\right) =$$

$$= \left|\text{н. о. р.} \quad Dx_i = \sigma^2\right| =$$

$$= 1 - \frac{1}{n^2}\epsilon^2 D\xi = 1 - \frac{n\sigma^2}{n^2\epsilon^2} = 1 - \frac{\sigma^2}{n\epsilon^2} \to_{n\to\infty} 1 - 0$$

Т.к. $P(A) \leq 1 \quad \forall A$, то $\lim_{n\to\infty} P\{|\overline{x} - a| < \epsilon\} = 1 \Rightarrow \overline{x}$ — состоятельная.

3.4 Дополнительная информация

Определение:

Выборочной дисперсией (смещённой) называется величина $S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$.

Для группированных данных: $S^2 = \frac{1}{n}\sum_{i=1}^{m}(x_i - \overline{x})^2 * n_i$.

Определение:

Выборочной дисперсией (несмещённой) называется оценка $\tilde{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$.

Для группированных данных: $\tilde{\sigma}^2 = \frac{n}{n-1} * S^2$.

3.5 Теорема о несмещённой и состоятельной оценке дисперсии

Теорема:

Пусть $X_1, \ldots, X_n \sim L(x, \theta)$, где $M\xi = a < +\infty$, $D\xi = \sigma^2$. Тогда несмещённой и состоятельной оценкой $D\xi$ является величина $\tilde{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$.

Доказательство:

1. Докажем несмещённость ($M\tilde{\sigma}^2 = \sigma^2$):

$$M\tilde{\sigma}^2 = M\left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x} + \bar{x}^2)\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n}(Mx_i^2 - 2M(x_i\bar{x}) + M\bar{x}^2)\right) =$$

Найдём значения слагаемых:

$$Mx_i^2 = M\xi^2 = D\xi + (M\xi)^2 = \sigma^2 + a^2$$

$$\boxed{\begin{aligned} &D\xi = M\xi^2 - (M\xi)^2 \\ &\sigma^2 = M\xi^2 - a^2 \Rightarrow M\xi^2 = \sigma^2 + a^2 \end{aligned}}$$

$$Mx_i\bar{x} = M\left(x_i \cdot \frac{1}{n}\sum_{j=1}^{n} x_j\right) = \frac{1}{n}\left(M(x_ix_1) + M(x_ix_2) + \dots + Mx_i^2 + \dots + M(x_ix_n)\right) \overset{x_i\ \text{н.о.р.}}{=} \frac{1}{n}\left(Mx_iMx_1 + Mx_iMx_2 + \dots + Mx_i^2 + \dots + Mx_iMx_n\right) =$$

$$= \frac{1}{n}\left(a^2 + a^2 + \dots + (\sigma^2 + a^2) + \dots + a^2\right) = \frac{na^2 + \sigma^2}{n} = a^2 + \frac{\sigma^2}{n}$$

$$M\bar{x}^2 = M\bar{x}\cdot\bar{x} = M\left(\frac{1}{n}\sum_{i=1}^{n} x_i\bar{x}\right) = \frac{1}{n}\sum_{i=1}^{n} Mx_i\bar{x} = \frac{1}{n}\cdot n\cdot\left(a^2 + \frac{\sigma^2}{n}\right) = a^2 + \frac{\sigma^2}{n}$$

$$= \frac{1}{n-1}\cdot n\left(\sigma^2 + a^2 - 2\left(a^2 + \frac{\sigma^2}{n}\right) + \left(a^2 + \frac{\sigma^2}{n}\right)\right) = \frac{n}{n-1}\left(\sigma^2 + a^2 - a^2 - \frac{\sigma^2}{n}\right) =$$

$$= \frac{n}{n-1}\cdot\frac{n-1}{n}\cdot\sigma^2 = \sigma^2$$

Т.к. $M\tilde{\sigma}^2 = \sigma^2$, то оценка несмещённая

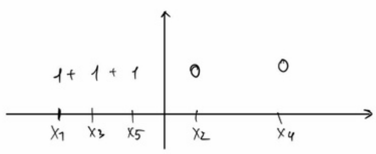
Состоятельность не доказывается.

3.6 Теорема о несмещённой и состоятельной оценке функции распределения

Определение:

Эмпирическая функция распределения: $\tilde{F}_n(x) = \frac{\mu_n(x)}{n}$; $\mu_n(x)$ – кол-во элементов выборки строго меньших x.

Пусть $e(x) = \begin{cases} 1, & x > 0 \\ 0, & x \le 0 \end{cases}$ $\quad \sum_{i=1}^{n} e(x - x_i)$



Таким образом, $\tilde{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} e(x - x_i)$.

Теорема:

Пусть $X_1, \ldots, X_n \sim L_\xi(x, \theta)$, где $F_\xi(x)$ — функция распределения случайной величины ξ. Тогда эмпирическая функция распределения $\tilde{F}_n(x)$ является несмещённой и состоятельной оценкой функции распределения $F_\xi(x)$.

Доказательство:

Покажем несмещённость $\left(M\left(\tilde{F}_n(x)\right) = F_\xi(x) = P\{\xi < x\}\right)$:

$$M\tilde{F}_n(x) = M\left(\frac{1}{n}\sum_{i=1}^{n} e(x - X_i)\right) = \frac{1}{n}\sum_{i=1}^{n} Me(x - X_i) \ \textcircled{=}$$

Построим ряд р-я сл. в. $e(x - X_i)$:

$e(x-X_i)$	0	1
P	$P\{x - X_i \le 0\}$	$P\{x - X_i > 0\}$

$$P\{x - X_i \le 0\} = P\{x \le X_i\} = P\{X_i \ge x\} = 1 - F_X(x)$$
$$P\{x - X_i > 0\} = P\{x > X_i\} = P\{X_i < x\} = F_{X_i}(x)$$

$$M\,e(x - X_i) = 0\cdot(1 - F_{X_i}(x)) + 1\cdot F_{X_i}(x) = F_{X_i}(x) \quad \forall i$$
$$Me^2(x - X_i) = 0^2(1 - F_{X_i}(x)) + 1^2\cdot F_{X_i}(x) = F_{X_i}(x)$$
$$De(x - X_i) = F_{X_i}(x) - F^2_{X_i}(x) = \underbrace{F_{X_i}(x)}_{0\le\cdot\le 1}\left(\underbrace{1 - F_{X_i}(x)}_{0\le\cdot\le 1}\right)$$

$$\textcircled{=}\ \frac{1}{n}\sum_{i=1}^{n} F_{X_i}(x) = \left[\begin{matrix}X_i - \\ \text{н.о.р}\end{matrix}\right] = \frac{1}{n}\cdot n\,F_\xi(x) = F_\xi(x)$$

Покажем состоятельность $\tilde{F}_n(x) \xrightarrow{P} F_\xi(x)$:

По неравенству Чебышева:

$$\forall \varepsilon > 0 \quad P\{|\tilde{F}_n(x) - F_\xi(x)| < \varepsilon\} \ge 1 - \frac{D\tilde{F}_n(x)}{\varepsilon^2} = 1 - \frac{1}{\varepsilon^2}D\left(\frac{1}{n}\sum_{i=1}^{n} e(x - X_i)\right) =$$

$$= 1 - \frac{1}{n^2\varepsilon^2}\cdot n\,F_\xi(x)\left(1 - F_\xi(x)\right) = 1 - \frac{F_\xi(x)\left(1 - F_\xi(x)\right)}{n\,\varepsilon^2} \xrightarrow[n\to\infty]{} 1 - 0$$

Т.К. $P(A) \le 1$, то $\tilde{F}_n(x) \xrightarrow{P} F_\xi(x)$

3.7 Теорема об эффективной оценке математического ожидания

Определение:

Пусть $\tilde{\theta}_n$ — несмещённая оценка. $\tilde{\theta}_n$ — называется эффективной оценкой, если

$$D\tilde{\theta}_n = \inf_{\tilde{\theta}_n^*} D\tilde{\theta}_n^*$$

Теорема:

Пусть $X_1, \dots, X_n \sim L_\xi(x, \theta)$, где ξ — случайная величина с $M\xi = a < +\infty$. Тогда выборочное среднее $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} X_i$ является эффективной оценкой $M\xi$ в классе линейных несмещённых оценок.

Доказательство:

Расс-м произвольную лин оценку $\widetilde{\theta}_n = \sum\limits_{i=1}^{n} \alpha_i X_i$:

$\sum\limits_{i=1}^{n} \alpha_i = 1$; $\alpha_i > 0\ \forall i$

Найдём дисперсию оценки $\widetilde{\theta}_n$:

$$D\widetilde{\theta}_n = D\sum_{i=1}^{n} \alpha_i X_i = \sum_{i=1}^{n} \alpha_i^2 DX_i = \begin{bmatrix} X_i - \\ \text{н.о.р.} \\ \text{сл. в.} \end{bmatrix} = \sum_{i=1}^{n} \alpha_i^2 \sigma^2 = \sigma^2 \sum_{i=1}^{n} \alpha_i^2 = \sigma^2 g(\alpha_1, \dots, \alpha_n)$$

$D\widetilde{\theta}_n$ достигает min на тех α_i, при которых $g(\alpha_1, \dots, \alpha_n) \to \min$

Для док-ва воспользуемся методом Лагранжа поиском условного минимума

$$g(\alpha_i) = \sum_{i=1}^{n} \alpha_i^2$$

константа Лагранжа

$$L(\alpha_i, \lambda) = \sum_{i=1}^{n} \alpha_i^2 + \lambda \left(\sum_{i=1}^{n} \alpha_i - 1 \right)$$

Найдём производные $L(\alpha_i, \lambda)$, λ - const Лагранжа.

$$\begin{cases} \dfrac{\partial L(\alpha_i, \lambda)}{\partial \lambda} = \sum\limits_{i=1}^{n} \alpha_i - 1 = 0 \\ \dfrac{\partial L(\alpha_i, \lambda)}{\partial \alpha_i} = 2\alpha_i + \lambda = 0 \Rightarrow \alpha_i = -\dfrac{\lambda}{2} \end{cases}$$

$$-\sum_{i=1}^{n} \frac{\lambda}{2} - 1 = 0 \Rightarrow$$

$$\Rightarrow \frac{n\lambda}{2} = -1; \quad \lambda = \frac{-2}{n} \Rightarrow \alpha_i = -\left(\frac{-2}{n}\right) \cdot \frac{1}{2} \Rightarrow \alpha_i = \frac{1}{n} \Rightarrow$$

$\Rightarrow g(\alpha_1, \dots, \alpha_n) = \sum\limits_{i=1}^{n} \alpha_i^2$ имеет мин. значение, если $g(\alpha_1, \dots, \alpha_n) = \sum\limits_{i=1}^{n} \left(\frac{1}{n}\right)^2$

$$\sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \frac{1}{n} = 1$$

Следовательно, эффективной оценкой является $\widetilde{\theta}_n = \frac{1}{n} \sum\limits_{i=1}^{n} X_i = \overline{X}$.

3.8 Теорема о единственности эффективной оценки

теорема:

Пусть $X_1, \dots, X_n \sim L_\xi(x, \theta)$, где θ — параметр распределения, и пусть $\tilde{\theta}(\overline{X_n})$ и $\widehat{\theta}_n(\overline{X_n})$ — две эффективные оценки θ. Тогда $\tilde{\theta}_n(\overline{X_n}) = \widehat{\theta}_n(\overline{X_n})$. То есть $P\{(x_1, \dots, x_n) : \tilde{\theta}_n(\overline{X_n}) \neq \widehat{\theta}_n(\overline{X_n})\} = 0$.

Доказательство:

Док-во: т.к. $\tilde{\theta}_1$ и $\tilde{\theta}_2$ — эфф., то по опр. $M\tilde{\theta}_1 = M\tilde{\theta}_2 = \theta$. $D\tilde{\theta}_1$ и $D\tilde{\theta}_2$ есть $\inf D\tilde{\theta}_n$, т.к. $\inf$ единст., то $D\tilde{\theta}_1 = D\tilde{\theta}_2$. Рассм. $\hat{\theta} = \frac{\tilde{\theta}_1 + \tilde{\theta}_2}{2}$:

$M\hat{\theta} = \frac{1}{2}(M\tilde{\theta}_1 + M\tilde{\theta}_2) = \frac{1}{2}(\theta + \theta) = \theta \Rightarrow \hat{\theta}$ — несмещ. оценка

$D\hat{\theta} = \frac{1}{4}(D\tilde{\theta}_1 + D\tilde{\theta}_2 + 2\operatorname{cov}(\tilde{\theta}_1, \tilde{\theta}_2)) = \frac{1}{2}(D\tilde{\theta}_1 + \operatorname{cov}(\tilde{\theta}_1, \tilde{\theta}_2))$ (*)

$|\operatorname{cov}(\tilde{\theta}_1, \tilde{\theta}_2)| = |M(\tilde{\theta}_1 - M\tilde{\theta}_1)(\tilde{\theta}_2 - M\tilde{\theta}_2)| \le \sqrt{M(\tilde{\theta}_1 - M\tilde{\theta}_1)^2} \cdot \sqrt{M(\tilde{\theta}_2 - M\tilde{\theta}_2)^2} =$

$= \sqrt{D\tilde{\theta}_1} \cdot \sqrt{D\tilde{\theta}_2} = D\tilde{\theta}_1$, т.е. $|\operatorname{cov}(\tilde{\theta}_1, \tilde{\theta}_2)| \le D\tilde{\theta}_1$

$D\hat{\theta} = |D\hat{\theta}| = \frac{1}{2}|D\tilde{\theta}_1 + \operatorname{cov}(\tilde{\theta}_1, \tilde{\theta}_2)| \le \frac{1}{2}(|D\tilde{\theta}_1| + |\operatorname{cov}(\tilde{\theta}_1, \tilde{\theta}_2)|) \le \frac{1}{2}(|D\tilde{\theta}_1| + |D\tilde{\theta}_1|) =$

$= D\tilde{\theta}_1 = \inf D\tilde{\theta}_n \Rightarrow D\hat{\theta} = D\tilde{\theta}_1$

Решим (*): $D\tilde{\theta}_1 = \frac{1}{2}(D\tilde{\theta}_1 - \operatorname{cov}(\tilde{\theta}_1, \tilde{\theta}_2)) \Rightarrow D\tilde{\theta}_1 = \operatorname{cov}(\tilde{\theta}_1, \tilde{\theta}_2)$

Коэфф. корреляции $r(\tilde{\theta}_1, \tilde{\theta}_2) = \frac{\operatorname{cov}(\tilde{\theta}_1, \tilde{\theta}_2)}{\sqrt{D\tilde{\theta}_1}\sqrt{D\tilde{\theta}_2}} = \frac{D\tilde{\theta}_1}{D\tilde{\theta}_1} = 1 \Rightarrow \tilde{\theta}_2 = a\tilde{\theta}_1 + b$

$M\tilde{\theta}_2 = aM\tilde{\theta}_1 + b$ (несм.)

$$\left.\begin{array}{l} \theta = a\theta + b \\ 1\cdot\theta + 0 = a\theta + b \end{array}\right\} \begin{array}{l} \Rightarrow a = 1,\ b = 0 \Rightarrow \tilde{\theta}_1 = \tilde{\theta}_2 \Rightarrow \\ \Rightarrow \text{эфф. оценка единст.} \Rightarrow ! \blacksquare \end{array}$$

3.8.1 Определение регулярной параметрической модели. Информация Фишера.

Определение:

Пусть ξ — наблюдаемая случайная величина.

Выборочным пространством $\mathbf{X}_n$ называется $\{\overline{X}_n = (X_1, \ldots, X_n)\}$.

Параметрическим множеством называется Θ — множество значений параметра θ.

Определение:

Параметрической моделью называется

$$\{\mathbf{X}_n; F_\xi(x, \theta) : \theta \in \Theta\}$$

Определение:

Параметрическая модель называется регулярной, если:

1. параметрическое множество Θ — открытое множество;
2. носитель распределён, то есть множество $A = \{x : f(x) > 0\}$ — не зависит от параметров;
3. $\forall \theta$ и $x \in A$ существует производная $\frac{\partial f(x,\theta)}{\partial \theta} < +\infty$;
4. $\forall \theta \in \Theta \quad M\left(\frac{\partial \ln f(x,\theta)}{\partial \theta}\right) = 0 \quad M\left(\frac{\partial \ln f(x,\theta)}{\partial \theta}\right)^2 < +\infty$;

5. Дважды допустимо дифференцирование под знаком интеграла по параметру.

Определение:

Пусть ξ — случайная величина с функцией плотности $f(x, \theta)$. Тогда информацией Фишера для случайного вектора $\overline{X}_n$ называется $I_n(\theta) = M(\frac{\partial}{\partial\theta} \ln f(x, \theta))^2$

3.9 Неравенство Рао-Крамера. Построение эффективной по Рао-Крамеру оценки для пуассоновского распределения

Неравенство Рао-Крамера:

Пусть дана регулярная параметрическая модель. Тогда для любой несмещённой оценки $\tilde{\theta}$ параметра θ $D\tilde{\theta} \geq \frac{1}{nI_n(\theta)}$.

Определение:

Величину $e(\theta) = \frac{1}{nI(\theta) * D\tilde{\theta}}$ называют показаелем эффективности по Рао-Крамеру.

Для любой несмещённой оценки $0 < e(\theta) \leq 1$. Если $e(\theta) = 1$, то $\tilde{\theta}_n$ называется эффективной по Рао-Крамеру, следовательно, просто эффективна.

Построение эффективной по Рао-Крамеру оценки для пуассоновского распределения:

$(X_1, \ldots, X_n) \sim Pois(\lambda)$. Построить эфф. (оптимальную) оценку λ.

$P\{\xi = k\} = \frac{e^{-\lambda}\lambda^k}{k!} \Rightarrow \forall x_i \; P\{x_i = k\} = P\{\xi = k\}$

$M x_1 = \lambda$, $D x_1 = \lambda^*$ (из ТВ) $\quad M x_1^2 = \lambda^2 + \lambda$

Пусть $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ – оценка λ.

$M\overline{X} = \lambda$, $D\overline{X} = D(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2} \cdot \sum_{i=1}^{n} Dx_i^* = \frac{1}{n^2} \cdot n \cdot \lambda = \frac{\lambda}{n}$.

Найдём $I(\lambda)$ – в одном наблюдении:

$I(\lambda) = M\left(\frac{\partial \ln(\frac{\lambda^\xi e^{-\lambda}}{\xi!})}{\partial\lambda}\right)^2 = M(\frac{\partial}{\partial\lambda}(\xi \ln\lambda - \lambda - \ln\xi!)^2) = M(\frac{\xi}{\lambda} - 1)^2 = M(\frac{\xi^2}{\lambda^2} - \frac{2\xi}{\lambda} + 1) = \frac{1}{\lambda^2} M\xi^2 - \frac{2}{\lambda} M\xi + 1 =$

$= \frac{1}{\lambda^2}(\lambda^2 + \lambda) - \frac{2}{\lambda} \cdot \lambda + 1 = 1 + \frac{1}{\lambda} - 2 + 1 = \frac{1}{\lambda}$.

Вычислим $e(\overline{X}) = \frac{\lambda}{n} \cdot n \cdot \frac{1}{\lambda} = 1 \Rightarrow \overline{X}$ – оптимальная оценка параметра λ.

4 Методы построения оценок

4.1 Метод моментов. Пример (биномиальное распределение)

Определение:

Начальным теоретическим моментом k-го порядка называется

$m_k = M\xi^k$, если $M|\xi|^k < +\infty$.

Определение:

Центральным теоретическим моментом k-го порядка называется $\mu_k = M(\xi - M\xi)^k$, если $M|\xi| < +\infty$.

Определение:

Выборочным начальным моментом k-го порядка называется $\tilde{m}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$.

Определение:

Выборочным центральным моментом k-го порядка называется $\tilde{\mu}_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{x})^k$.

Метод моментов:

Метод моментов состоит в том, что за оценку параметров принимается решение системы уравнений:

$$\begin{cases} m_k = \tilde{m}_k \\ \mu_k = \tilde{\mu}_k \end{cases} \qquad (1)$$

Пример (биномиальное распределение):

Пусть $X_1, \ldots, X_n \sim Bin(m; p)$

m - кол-во независ. испытаний Бернулли

p - веро-ть успеха

Получить $\tilde{p}_{мм}$

Решение: $m_1 = \tilde{m}_1 \Rightarrow M\xi = \overline{X} \Rightarrow mp = \overline{X} \Rightarrow p = \frac{\overline{X}}{m}$.

Таким образом, $\tilde{p}_{мм} = \frac{\overline{X}}{m} = \frac{1}{n \cdot m} \sum_{i=1}^{n} X_i$

4.2 Функция правдоподобия. Метод максимального правдоподобия. Пример (биномиальное распределение)

Определение:

Пусть $X_1, \ldots, X_n \sim L_\xi(x, \theta)$, причём $f_\xi(x, \theta)$ — известна, а θ неизвестен. Функцией правдоподобия называют

$L(\overline{X}_n, \theta) = \prod\limits_{i=1}^{n} f(X_i, \theta) = f(X_1, \theta) * \cdots * f(X_n, \theta).$

Метод максимального правдоподобия:

Рассмотрим $\ln L(\overline{X}_n, \theta) = \sum\limits_{i=1}^{n} \ln f(X_i, \theta).$

Метод максимального правдоподобия состоит в том, что за оценку параметра θ принимается точка максимума функции правдоподобия.

Алгоритм:

1. Построить функцию $L(\overline{X}_n, \theta)$;
2. Взять её логарифм $\ln L(\overline{X}_n, \theta)$;
3. $\frac{\partial \ln L(\overline{X}_n, \theta)}{\partial \theta} = 0 \Rightarrow \theta$;
4. Если значение второй производной строго меньше нуля, то оценка параметра $\tilde{\theta}$ и есть θ_0.

$$\frac{\partial^2 \ln L(\overline{X}_n, \theta)}{\partial \theta^2}\Big|_{\theta_0} < 0, \text{ то } \ \tilde{\theta} = \theta_0$$

Пример (биномиальное распределение):

Пример (Построить оценку $\tilde{p}_{\text{МНП}}$ параметра p $Bin(m, p)$):

1. $$L(\overline{X}_n; \theta) = C_m^{X_1} p^{X_1} q^{m-X_1} \cdot C_m^{X_2} p^{X_2} q^{m-X_2} \cdot \ldots \cdot C_m^{X_n} p^{X_n} q^{m-X_n} =$$
$$= \left(\prod_{i=1}^{n} C_m^{X_i}\right) \cdot p^{n\overline{x}} q^{nm - n\overline{x}} \qquad \text{т.к.} \quad x_1 + \ldots + x_n = \sum_{i=1}^{n} x_i = n \cdot \overbrace{\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)}^{\overline{x}} = n\overline{x}$$

2. $$\ln L(\overline{X}_n; p) = \sum_{i=1}^{m} \ln C_m^{X_i} + n\overline{x} \ln p + (nm - n\overline{x}) \ln(1-p)$$

3. $$\frac{\partial}{\partial p} \ln L(\overline{X}_n; p) = \frac{n\overline{x}}{p} - \frac{nm - n\overline{x}}{1-p} = 0 \qquad \boxed{0 < p < 1}$$
$$\overline{x} - p\overline{x} - (m - \overline{x})p = 0$$
$$p_0 = \frac{\overline{x}}{m}.$$

4. $$\frac{\partial^2 L(\overline{X}_n; p)}{\partial p^2} = -\frac{n\overline{x}}{p^2} - \frac{nm - n\overline{x}}{1-p^2} \Bigg|_{p_0} = -n\left(\frac{\overline{x}}{\overline{x}^2/m^2} + \frac{(m-\overline{x})}{(m-\overline{x})^2/m^2}\right) = -m^2 n \left(\frac{1}{\overline{x}} + \frac{1}{m-\overline{x}}\right) =$$
$$= -m^2 n \left(\frac{m - \overline{x} + \overline{x}}{\overline{x}(m-\overline{x})}\right) = -\frac{m^2 n^2}{\overline{x}\underbrace{(m-\overline{x})}_{>0}} < 0 \Rightarrow$$
$$\Rightarrow \tilde{p}_{\text{МНП}} = p_0 = \frac{\overline{x}}{m}.$$

Оценки по методу МНП не явл-ся несмещ по пос-ию, но они состоятельные или асимптотически сос-ые.

5 Регрессионный анализ

5.1 Постановка задачи

подготовительная информрмация:

Выборочные средние имеют вид:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

Выборочные дисперсии (смещённые) имеют вид:

$$S_X^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{x})^2, \quad S_Y^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{y})^2.$$

Величины $S_X = \sqrt{S_x^2}$ и $S_Y = \sqrt{S_Y^2}$ представляют собой выборочные средние квадратические отклонения (смещённые) величин ξ и η.

Выборочное среднее (совместное) есть величина:

$$\overline{xy} = \frac{1}{n}\sum_{i=1}^{n} X_i Y_i$$

Коэффициент корреляции (выборочный):

$$r_{\text{выб}}(\text{дальше} \quad r_v) = \frac{\overline{xy} - \overline{x} * \overline{y}}{\sqrt{S_x^2}\sqrt{S_y^2}}$$

Постановка задачи:

Пусть производятся наблюдения некоторого процесса, который описывается двумя случайными величинами (ξ, η).

Тогда результаты наблюдений представляют собой двумерную выборку $(X_1, Y_1), \ldots, (X_n, Y_n)$.

Множество точек с координатами (X_i, Y_i) называют полем корреляции или диаграммой рассеяния. По виду поля корреляции делают предположение о связи между величинами ξ и η.

Как известно, если коэффициент корреляции $r = Cov(\xi, \eta)/\sqrt{(D\xi D\eta)}$ равен нулю, то величины ξ и η некоррелированы, если $|r| = 1$, то величины ξ и η линейно связаны $\eta = a\xi + b$.

Следовательно, при изучении связи величин необходимо оценить отличие выборочного коэффициента корреляции от нуля.
Проверим гипотезу о значимости коэффициента корреляции, т.е. H_0 :" $r = 0$ " против гипотезы H_1 :" $r \neq 0$ ".

Статистика $t = r \cdot \sqrt{\frac{n-2}{1-r^2}}$ имеет распределение Стьюдента с $n-2$ степенями свободы.

Для заданного уровня значимости критическое значение равно $t_{\text{кр}} = t(\frac{\alpha}{2}; n-2)$.

Если $t_{\text{набл}} < t_{\text{кр}}$, принимают гипотезу H_0 :" $r = 0$ ", в противном случае H_0 отвергают, и делают вывод о значимом отличии коэффициента корреляции от нуля.

Если установлено, что коэффициент значимо отличается от нуля, то линейное уравнение должно достаточно близко описывать имеющуюся между величинами, но неизвестную связь.

5.2 Построение выборочного уравнения регрессии

Пусть наблюдалась выборка $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$, и выборочный коэффициент значимо отличается от нуля.

Найдем коэффициенты уравнения $Y = aX + b$, которое наилучшим образом аппроксимирует $Y = f(X)$ ($\eta = f(\xi)$).

Величина Y называется зависимой переменной, признаком, величина Y называется независимой переменной, фактором, регрессором.

В парной регрессионной модели зависимая переменная зависит только от одного регрессора.

Оценим коэффициенты уравнения методом наименьших квадратов (МНК).
Будем обозначать $\hat{Y}$-- вычисленные (прогнозные) значения.

Согласно МНК требуется найти такие значения оценок параметров $\hat{a}$ и $\hat{b}$, чтобы была минимальной сумма квадратов отклонений прогнозных значений от наблюдаемых:

$$L(\hat{a}, \hat{b}) = \sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2 = \sum_{t=1}^{n}(Y_t - (\hat{a} + \hat{b}X_t))^2 \to min.$$

Значит, для нахождения оценки параметров парной регрессионной модели МНК необходимо найти экстремум (минимум) функции двух аргументов.

Запишем необходимые условия экстремума:

$$\begin{cases} \frac{\partial L}{\partial \hat{a}} = -2\sum\limits_{t=1}^{n}(Y_t - \hat{a} - \hat{b}X_t) = 0, \\ \frac{\partial L}{\partial \hat{b}} = -2\sum\limits_{t=1}^{n}X_t(Y_t - \hat{a} - \hat{b}X_t) = 0, \end{cases}$$

Раскрывая скобки получим:

$$\begin{cases} \sum\limits_{t=1}^{n}Y_t - \hat{a}n - \hat{b}\sum\limits_{t=1}^{n}X_t = 0, \\ \sum\limits_{t=1}^{n}X_tY_t - \hat{a}\sum\limits_{t=1}^{n}X_t - \hat{b}\sum\limits_{t=1}^{n}X_t^2 = 0. \end{cases}$$

Из первого уравнения системы имеем оценку параметра a:

$$\hat{a} = \frac{1}{n}\sum_{t=1}^{n}Y_t - \hat{b}\frac{1}{n}\sum_{t=1}^{n}X_t = \bar{y} - \hat{b}\bar{x}.$$

Преобразуем второе уравнение системы и подставим полученную оценку $\hat{a}$

$$\sum_{t=1}^{n} X_t Y_t - \hat{a} \sum_{t=1}^{n} X_t - \hat{b} \sum_{t=1}^{n} X_t^2 = n\overline{xy} - \hat{a}\bar{x} - n\hat{b}(S_x^2 + \bar{x}^2) = 0$$

$$\overline{xy} - \bar{x}\bar{y} + \hat{b}\bar{x}^2 - \hat{b}S_x^2 - \hat{b}\bar{x}^2 = 0$$

Отсюда получаем оценку параметра b:

$$\hat{b} = \frac{\overline{xy} - \bar{x}\bar{y}}{S_x^2} = \frac{\overline{xy} - \bar{x}\bar{y}}{S_x S_y} \frac{S_y}{S_x} = r\frac{S_y}{S_x},$$

Таким образом решение системы уравнений имеет вид

$$\begin{cases} \hat{b} = r\frac{S_y}{S_x} \\ \hat{a} = \bar{y} - \hat{b}\bar{x} \end{cases}$$

Уравнение регрессии Y на X имеет вид:

$$\bar{y}_x = \bar{y} + r\frac{S_y}{S_x}(x - \bar{x})$$

Уравнение регрессии X на Y имеет вид

$$\bar{x}_y = \bar{x} + 1/r\frac{S_x}{S_y}(y - \bar{y})$$

Заметим, что каждое из уравнений имеет $\bar{x}$ своим решением , т.е графики проходят через точку $(\bar{x}, \bar{y})$

5.3 Выборочный коэффициент корреляции, его свойства

Определения:

$$r_{\text{выб}}(\text{дальше} \quad r_v) = \frac{\overline{xy} - \overline{x} * \overline{y}}{\sqrt{S_x^2}\sqrt{S_y^2}}$$

Свойства выборочного коэффициента корреляции:

1. $|\bar{r}| \leq 1$
2. Если $|\bar{r} \approx 0|$, то x и y слабо коррелируемы. Если известно, что x и y независимы, то $r = 0$.
3. ξ, η линейно связаны $\iff |\bar{r}| = 1$.

5.4 Проверка гипотезы о значимости выборочного коэффициента корреляции

Рассм. выборочный коэфф-т корреляции $r = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{S_x S_y}$, где S_x, S_y — смещённые оценки дисперсии.
Значимость коэфф-та корр. опр. по гипотезе H_0: "$r=0$" против гип. H_1: "$r \neq 0$". Ур. знач-ти $\alpha = (0{,}05; 0{,}01)$.
$t_{набл.} = r\sqrt{\frac{n-2}{1-r^2}}$ $\quad t_{кр.} \sim T(\frac{\alpha}{2}; n-2)$ — р-е Стьюдента
Если $|t_{набл.}| < t_{кр.}$, то гип. H_0 следует принять. Т.о., коэфф-т корр. незначимо отл. от 0 $\Rightarrow$ ур-е регрессии незначимо.
Иначе ($|t_{набл.}| > t_{кр.}$) принять H_1 (H_0 отвергнуть): r значимо отл. от 0.

5.5 Теорема Гаусса-Маркова

Теорема Гаусса-Маркова: пусть (X_t, Y_t), $t=\overline{1,n}$, — рез-т набл.. Предп., что $Y=f(X)$, но в рез-те набл. знач. Y_t измер. с ошибкой, т.е. $Y_t = f(X_t) + \varepsilon_t$. Пусть $f(X)$ — линейная ф-ция. Тогда
$Y_t = a + bX_t + \varepsilon_t$, $t=\overline{1,n}$, Y_t — зависимая пер. (объясняемая), X_t — независ. пер. (регрессор), ε_t — сл. вел. (сл. ошибка).

[redacted]

1) Спецификация регрессионной модели
$Y_t = a + bX_t + \varepsilon_t$, $t=\overline{1,n}$
2) X_t — детерминированная вел.
3) $M\varepsilon_t = 0$ $\quad M\varepsilon_t^2 = D\varepsilon_t = \sigma^2$ не зав. от t
$M\varepsilon_t\varepsilon_s = 0$, $s \neq t$ (некорр-ть ошибок модели)
незав-ть $\Rightarrow$ корр-ть $\not\Leftarrow$ обратное неверно
3*) $\varepsilon_t \sim N(0;\sigma^2)$ — норм. модель

Теорема Гаусса-Маркова: в предп-ях 1,2,3 оценки $\hat{a}$ и $\hat{b}$, полученные МНК, имеют наим. дисперсию в классе несм. оценок.

5.6 Статические свойства МНК оценок

Статистические свойства МНК – оценок.

Пусть рассм. норм. регрессионная модель, т.е. $\varepsilon_t \sim N(0; \sigma^2)$.
Тогда $\bar{\varepsilon} \sim N(0; \sigma^2 I_n)$, где $\bar{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)$ – вектор ошибок модели, ε_t – н.о.р. сл.в.. (I_n – един. мат.)

1) $y_t = a + bx_t + \varepsilon_t \Rightarrow My_t = a + bx_t$

$\vec{y} \sim N(a + b\bar{x}; \sigma^2 I_n)$, $\vec{y} = (y_1, \dots, y_n)$

$\hat{a} \sim N\left(a; \sigma^2 \frac{\Sigma x_t^2}{n \Sigma x_t^2}\right)$ $\hat{b} \sim N\left(b; \sigma^2 \frac{1}{\sum_{t=1}^{n} x_t^2}\right)$

2) Дисперсия ошибок ε_t

$\bar{\varepsilon} \sim N(0; \sigma^2 I_n)$; $\frac{(n-2)S^2}{\sigma^2} \sim \chi^2(n-2)$

$S^2 = \sum_{t=1}^{n} (\varepsilon_t - \bar{\varepsilon})^2$; $\bar{\varepsilon} = \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t$

3) Дисперсия ошибок S^2, оценки $\hat{a}$ и $\hat{b}$ – независимые сл.в..

5.7 Анализ вариации

Анализ вариации зависимой переменной в регрессии.

$y_t = a + bx_t + \varepsilon_t$ наблюд.

$\hat{y} = \hat{a} + \hat{b} x_t$ предсказ. (прогноз.)

Рассм. дисперсию

$$\sum_{t=1}^{n} (y_t - \bar{y})^2 = \sum_{t=1}^{n} (y_t - \hat{y}_t + \hat{y}_t - \bar{y})^2 = \underbrace{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}_{\text{откл. от рег.}} + \underbrace{\sum_{t=1}^{n} (\hat{y}_t - \bar{y})^2}_{\text{рег. откл. от ср.}} + \underbrace{2 \sum (y_t - \hat{y}_t)(\hat{y}_t - \bar{y})}_{=0}$$

$$\underbrace{\sum (y_t - \bar{y})^2}_{TSS} = \underbrace{\sum (y_t - \hat{y}_t)^2}_{ESS} + \underbrace{\sum (\hat{y}_t - \bar{y})^2}_{RSS}$$

TSS – Total Sum of Squares, ESS – Error Sum of Squares, RSS – Regression Sum of Squares

Коэфф-т детерминации $R^2 = \frac{RSS}{TSS} = 1 - \frac{ESS}{TSS}$, $0 \le R^2 \le 1$.

$R^2 = 1$ – ур-е почти полностью объясняет вариацию

$R^2 \approx 0$ – ур-е незначимо

R^2 имеет распр-е Фишера $F(n, m) = \frac{\frac{1}{n}(\xi_1^2 + \dots + \xi_n^2)}{\frac{1}{m}(\eta_1^2 + \dots + \eta_m^2)}$