Содержание

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 сходимостью в среднем и по вероятности

Теорема:

Из сходимости среднеквадратичного следует сходимость по вероятности.

Доказательство:

Так как $\xi_i \xrightarrow{\text{по ср-к-ому}} \xi$, то $M(\xi_i - \xi)^2 \xrightarrow[i\to\infty]{} 0$

По неравенству Чебышёва

$$\forall \varepsilon > 0 \quad P\{|\xi - M\xi| < \varepsilon\} \geq 1 - \frac{D\xi}{\varepsilon^2}$$

Тогда

$$\forall \varepsilon > 0 \quad P\{|\xi_i - \xi| < \varepsilon\} \geq 1 - \frac{M|\xi_i - \xi|^2}{\varepsilon^2} \xrightarrow[i\to\infty]{} 1$$

$$\xi_i \xrightarrow{\text{ср. квдр.}} \xi \Rightarrow \xi_i \xrightarrow{P} \xi$$

1.3 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 $\xi_n \xrightarrow{P} c$

1.4 Теорема о связи между сходимостями по вероятности, в среднем, по распределению

Теорема:

Из сходимости почти наверное вытекает сходимость по вероятности $\{\xi_n\} \to^{\text{п.н.}} \xi \Rightarrow \{\xi_n\} \to^P \xi$.

Из схоимости по вероятности вытекает сходимость по распределению $\{\xi_n\} \to^P \Rightarrow \{\xi_n\} \to^d \xi$.

Из сходимости в среднем по порядку вытекает сходимость по вероятности $\{\xi_n\} \to \xi \Rightarrow \{\xi_n\} \to^P \xi$.

1.5 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 $\frac{S_n}{n} \to \frac{A_n}{n}$
$S_n = \xi_1 + \dots + \xi_n; \quad A_n = a_1 + \dots a_n$.

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

Теорема 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, $\forall \varepsilon > 0 \quad P\{|\frac{S_n}{n} - a| \geq \varepsilon\} \xrightarrow[n\to\infty]{} 0$ $\square$

1.6 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 $\forall x \in R \quad P\left\{\frac{S_n - na}{\sqrt{DS_n}} < x\right\} \to_{n\to\infty} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{n^2}{2}}du$. То есть $z_n = \frac{S_n - MS_n}{\sqrt{DS_n}}$ сходится по распределению с случаной величиной $z \sim N(0,1)$.

Доказательство:

Построим х/ф z_n:

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

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

Рассмотрим $(\xi_1 - a)$:

$$D\xi_1 - a = D\xi_1 + Da = \sigma^2$$

$$M(\xi_1 - a) = 0 \qquad D(\xi_1-a) = M(\xi_1-a)^2 - [M(\xi_1-a)]^2 = M(\xi_1-a)^2$$

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 $x_1^* \leq \dots \leq x_n^*$.

x_i^* — варианта.

n_i — частота появления варинты x_i^* в выборке.

Определение:

Точечным вариационным рядом называется:

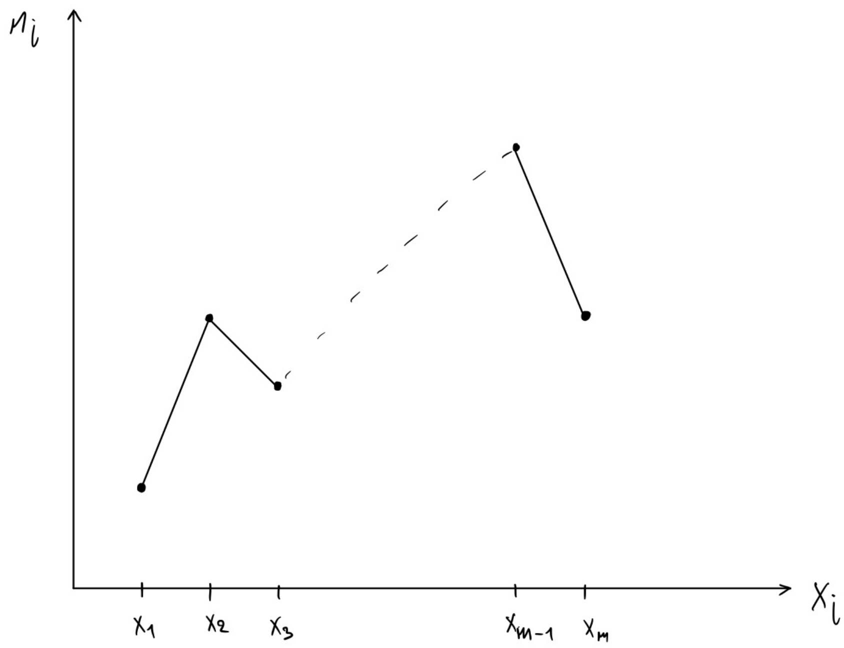
x_i	x_1	x_2	...	x_m
n_i	n_1	n_2	...	n_m

x_i — варианта, n_i — частота соответствуующей 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:

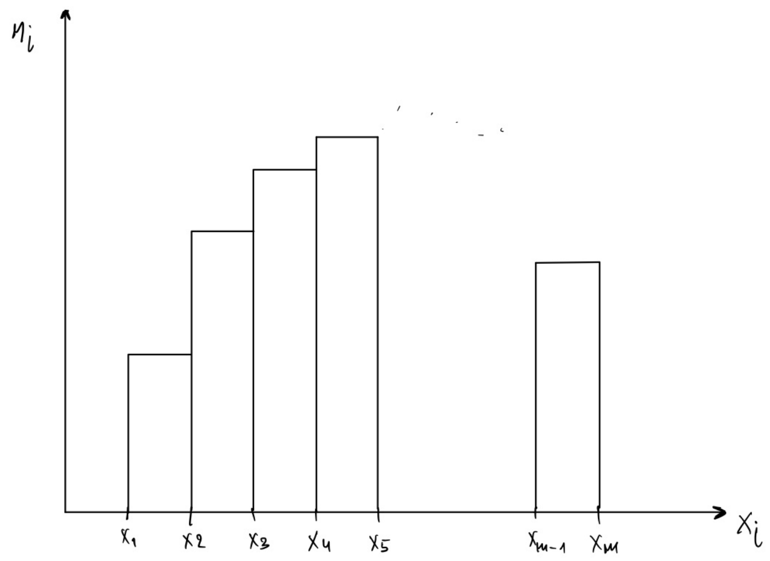
x_i	$[x_1, x_2]$	$(x_2, x_3]$	$\ldots$	$(x_m, x_{m+1}]$
n_i	n_1	n_2	$\ldots$	n_m

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 Эмперическая функция распределения ЭФР. Свойства ЭФР.

Эмперическая функция распределения ЭФР

Эмперической 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 $x \le 0$.

Таким образом, если $X_i < x$, то $e(x) = 1$, если $X_i \ge x$, 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 $x_1, x_2, \ldots, x_n$ — реализация выборки $X_1, X_2, \ldots, X_n$, 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. $0 \le \tilde{F}_n(x) \le 1$, т.к. $0 \le \mu(x) \le n$;
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 $g(t_1, \ldots, t_n)$ — непрерывная функция. Оценкой θ назовём $\tilde{\theta} = g(X_1, \ldots, X_2)$. Если $g(X_1, \ldots, X_n) = T$ 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.

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

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

Теорема:

Пусть $X_1, \ldots, X_n \sim L_\xi(x,\theta)$, 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:

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

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

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

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:

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

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

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

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

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

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

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

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

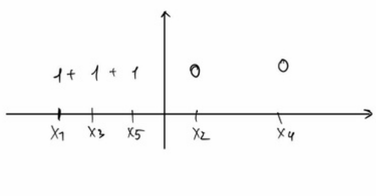
Состоятельность не доказывается.

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 $X_1, \dots, X_n \sim L_\xi(x, \theta)$, 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 $\left(M\left(\tilde{F}_n(x)\right) = F_\xi(x) = P\{\xi < x\}\right)$:

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

Построим ряд р-я сл. в. $e(x - X_i)$:

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

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

$$P\{x - X_i > 0\} = P\{x > X_i\} = P\{X_i < x\} = F_{X_i}(x)$$

$$M\,e(x - X_i) = 0\cdot\left(1 - F_{X_i}(x)\right) + 1\cdot F_{X_i}(x) = F_{X_i}(x) \quad \forall i$$

$$M e^2(x - X_i) = 0^2\left(1 - F_{X_i}(x)\right) + 1^2\cdot F_{X_i}(x) = F_{X_i}(x)$$

$$D e(x - X_i) = F_{X_i}(x) - F^2_{X_i}(x) = \underbrace{F_{X_i}(x)}_{0 \le \cdot \le 1}\underbrace{\left(1 - F_{X_i}(x)\right)}_{0 \le \cdot \le 1}$$

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

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

По неравенству Чебышёва:

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

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

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 $X_1, \ldots, X_n \sim L_\xi(x, \theta)$, 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 $\widetilde{\Theta}_n = \sum_{i=1}^{n} \alpha_i X_i$:

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

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

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

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

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.

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

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

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

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

Следовательно, эффективной оценкой является $\widetilde{\Theta}_n = \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{X}$.

3.8 Теорема о единственности эффективной оценки

Пусть $X_1, \ldots, X_n \sim L_\xi(x, \theta)$, где θ – параметр рас-ия, и пусть $\tilde{\theta}_1(\overline{X_n})$ и $\tilde{\theta}_2(\overline{X_n})$ – две эффективные оценки θ. Тогда $\tilde{\theta}_1(\overline{X_n}) = \tilde{\theta}_2(\overline{X_n})$.

То есть $P\{(X_1, \ldots, X_n) : \tilde{\theta}_1(\overline{X_n}) \neq \tilde{\theta}_2(\overline{X_n})\} = 0$.

Док-во:

Так как $\tilde{\theta}_1$ и $\tilde{\theta}_2$ – эффективные, то:

$$M\tilde{\theta}_1 = M\tilde{\theta}_2 = \theta.$$

$$D\tilde{\theta}_1 = D\tilde{\theta}_2 = \inf D\tilde{\theta}_n.$$

Рассмотрим $\hat{\theta} = \frac{\tilde{\theta}_1 + \tilde{\theta}_2}{2}$:

$$M\hat{\theta} = \frac{1}{2}\left(M\tilde{\theta}_1 + M\tilde{\theta}_2\right) = \theta \Rightarrow \hat{\theta} \text{ – несмещённая}$$

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

$$\left|cov(\tilde{\theta}_1, \tilde{\theta}_2)\right| = \left|M(\tilde{\theta}_1 - M\tilde{\theta}_1)(\tilde{\theta}_2 - M\tilde{\theta}_2)\right| \leq \sqrt{M(\theta_1 - M\theta_1)^2} \cdot \sqrt{M(\theta_2 - M\theta_2)^2} =$$

$$= \sqrt{D\tilde{\theta}_1} \cdot \sqrt{D\tilde{\theta}_2} = D\tilde{\theta}_1 \Rightarrow \left|cov(\tilde{\theta}_1, \tilde{\theta}_2)\right| \leq D\tilde{\theta}_1$$

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

$$\leq \frac{1}{2}\left(|D\tilde{\theta}_1| + |D\tilde{\theta}_1|\right) = D\tilde{\theta}_1 = \inf D\tilde{\theta}_n \Rightarrow D\hat{\theta} = D\tilde{\theta}_1$$

Решим (*):

$$D\tilde{\theta}_1 = \frac{1}{2}\left(D\tilde{\theta}_1 - cov(\tilde{\theta}_1, \tilde{\theta}_2)\right) \Rightarrow D\tilde{\theta}_1 = cov(\tilde{\theta}_1, \tilde{\theta}_2)$$

Коэффициент корреляции $r(\tilde{\theta}_1, \tilde{\theta}_2) = \frac{cov(\tilde{\theta}_1, \tilde{\theta}_2)}{\sqrt{D\tilde{\theta}_1}\sqrt{D\tilde{\theta}_2}} = \frac{D\tilde{\theta}_1}{D\tilde{\theta}_1} = 1 \Rightarrow \tilde{\theta}_2 = a\theta_1 + b$

$M\tilde{\theta}_2 = aM\tilde{\theta}_1 + b$ – несмещ.

$\theta = a\theta + b \Rightarrow a = 1 \quad b = 0 \Rightarrow \tilde{\theta}_1 = \tilde{\theta}_2$ □

3.8.1 Определение регулярной параметрической модели. Информация Фишера.

Определение:

Пусть ξ — наблюдаемая случайная величина.
Выборочным пространством $\mathbf{X}_n$ называется $\{\overline{X}_n = (X_1, \dots, X_n)\}$.
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:

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

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

$MX_1 = \lambda$, $DX_1 = \lambda^*$ (из ТВ) $\quad MX_1^2 = \lambda^2 + \lambda$

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

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

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

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

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
$\tilde{m}_k = \frac{1}{n}\sum\limits_{i=1}^{n} X_i^k$.

Определение:

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

Метод моментов:

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

$$\begin{cases} m_k = \tilde{m}_k \\ \mu_k = \tilde{\mu}_k \end{cases} \tag{1}$$

Пример (биномиальное распределение):

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

m – кол-во независ. испытаний Бернулли

p – веро-ть успеха

Получить $\tilde{p}_{мм}$

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

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

Пример (нормальное распределение):

Пусть $X_1, \ldots, X_n \sim N(a, \sigma^2)$

ММ:

$\mu_1 = M\xi = a$

$\mu_2 = M(\xi - M\xi)^2 = D\xi = \sigma^2$

$\tilde{\mu}_1 = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$\tilde{\mu}_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$

Следовательно:

$$\begin{cases} \mu_1 = \tilde{\mu}_1 \\ \mu_2 = \tilde{\mu}_2 \end{cases} \Rightarrow \begin{cases} a = \frac{1}{n}\sum_{i=1}^{n} X_i \\ \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 \end{cases}$$

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

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

Метод максимального правдоподобия:

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

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.

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

Пример (биномиальное распределение):

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

1. $L(\overline{X}_n; \theta) = C_m^{x_1} p^{x_1} q^{m-x_1} \cdot C_m^{x_2} p^{x_2} q^{m-x_2} \cdot \ldots \cdot C_m^{x_n} p^{x_n} q^{m-x_n} =$

$$= \left(\prod_{i=1}^{n} C_m^{x_i}\right) \cdot p^{n\bar{x}} q^{nm-n\bar{x}} \qquad \text{т.к.} \quad x_1 + \ldots + x_n = \sum_{i=1}^{n} x_i = n \cdot \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^{\nearrow \bar{x}} = n\bar{x}$$

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

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

$$\bar{x} - p\bar{x} - (m - \bar{x})p = 0$$

$$p_0 = \frac{\bar{x}}{m}.$$

4. $\frac{\partial^2 L(\overline{X}_n; p)}{\partial p^2} = -\frac{n\bar{x}}{p^2} - \frac{nm - n\bar{x}}{1-p^2}\Big|_{p_0} = -n\left(\frac{\bar{x}}{\bar{x}^2/m^2} + \frac{(m-\bar{x})}{(m-\bar{x})^2/m^2}\right) = -m^2 n \left(\frac{1}{\bar{x}} + \frac{1}{m-\bar{x}}\right) =$

$$= -m^2 n \left(\frac{m - \bar{x} + \bar{x}}{\bar{x}(m-\bar{x})}\right) = -\frac{m^2 n^2}{\bar{x}\underbrace{(m-\bar{x})}_{>0}} < 0 \Rightarrow$$

$$\Rightarrow \tilde{p}_{ммп} = p_0 = \frac{\bar{x}}{m}.$$

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

Пример (нормальное распределение):

Пусть $X_1, \dots, X_n \sim N(a, \sigma^2)$

ММП:

1. $L(X_n, a, \sigma) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x_i - a)^2}{2\sigma^2}} = \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right) e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - a)^2}$

2. $\ln L(X_n, a, \sigma) = -n \ln\sqrt{2\pi} - n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - a)^2$

$$\begin{cases} \frac{\partial \ln L(X_n, a, \sigma)}{\partial a} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - a) = 0 \\ \frac{\partial \ln L(X_n, a, \sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - a)^2 = 0 \end{cases} \Rightarrow \begin{cases} a_0 = \bar{x} \\ \sigma_0 = S^2 \end{cases}$$

$$\frac{\partial^2 \ln L(X_n, a, \sigma)}{\partial a^2} = -\frac{n}{\sigma^2} < 0$$

$$\frac{\partial^2 \ln L(\bar{X}_n, a, \sigma)}{\partial \sigma^2} < 0 \Rightarrow \hat{a} = \bar{x} \quad и \quad \hat{\sigma}^2 = S^2$$

5 Довриттельное оценивание

5.1 Поставнока задачи, определение доверительного интервала, доверительная вероятность

Пусть $X_1, \dots, X_n \sim L_\xi(x, \theta)$, где распределение случайной величины ξ известно, а $f_\xi(x, \theta)$ и θ неизвестны.

Пусть n - объём выборки и $n < 30 \Rightarrow$ состоятельность и несмещённость оценок $\tilde{\theta}_n$ не выполняется.

В этом случае строят интервальную оценку. То есть $(\tilde{\theta}_1; \tilde{\theta}_2)$, в котором предполагают наличие θ. $\tilde{\theta}_1 = \theta_1(x_1, \dots, x_n)$ $\tilde{\theta}_2 = \theta_2(x_1, \dots, x_2)$.

Доверительным интервалом называется интервал с границами $(\tilde{\theta}_1; \tilde{\theta}_2)$, который с вероятностью не меньше γ покрывает неизвестный параметр θ, т.е.

$$P\{\theta \in (\tilde{\theta}_1; \tilde{\theta}_2)\} \geq \gamma,$$ где γ - доверительная вероятность.

5.2 Лемма Фишера

Лемма Фишера(следствие)

Пусть $X_1, X_2, \dots, X_n \sim \mathcal{N}(a, \sigma^2)$ выборка из нормального распределения математическим ожиданием a и дисперсией σ^2.

Тогда сл. величины $\overline{x}$ и S^2 независимы,

а сл.в

$\frac{\overline{x}-a}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$ распределена по стандартному нормальному закону,

$\frac{\overline{x}-a}{\tilde{\sigma}/\sqrt{n}} \sim T(n-1)$ распределена по закону Стьюдента,

$\frac{\sqrt{n}\,\sigma^2}{\tilde{\sigma}^2} \sim \chi^2(n-1)$ распределена по закону хи-квадрат.

5.3 Распределения Стьюдента, хи-квадрат, формулы построения этих случайных величин

Определение:

Рассмотрим n н.о.р. случайных величин $\xi_1, \dots, \xi_n \sim N(0;1)$.

Случайной величиной с распределением $\chi^2(n)$ с n степенями свободы наз-ся величина $\chi^2(n) = \xi_1^2 + \dots + \xi_n^2$

$$\text{При } x \geq 0 \quad f(x) = \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} \qquad \text{при } x < 0 \quad f(x) = 0$$

$$\Gamma(\lambda) = \int_0^\infty t^{\lambda-1} e^{-t} dt, \quad \lambda > 0 \text{ – гамма функция}$$

$$M\xi = n, \quad D\xi = \sigma^2 = 2n$$

Определение:

Рассмотрим n н.о.р. случайных величин $\xi_1, \dots, \xi_n \sim N(0;1)$.

Случайной величиной с распределением Стьюдента или t-распределением с n степенями свободы наз-ся величина

$$t(n) = \frac{\eta}{\sqrt{\chi^2(n)/n}} = \frac{\eta}{\sqrt{\xi_1^2 + \dots + \xi_n^2 / n}}$$

$$f(x) = \frac{\Gamma\left((n+1)/2\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{(n+1)}{2}}$$

$$\text{при } n > 2 \quad M\xi = 0 \quad D\xi = \sigma^2 = \frac{n}{n} - 2$$

5.4 Построение доверительных интервалов для параметра a нормального распределения при известном σ

Постановка: Пусть $X_1, \ldots, X_n \sim N(a, \sigma^2)$, a — неизвестный параметр, σ^2 — известно.
Требуется построить интервальную оценку, то есть построить интервал $(\tilde{a}_1; \tilde{a}_2)$, т. что $P\{a \in (\tilde{a}_1; \tilde{a}_2)\} \geq \gamma$, γ — задано.

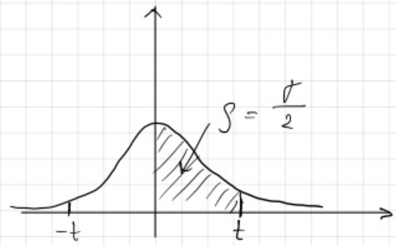
Решение: из леммы Фишера имеем $\frac{\bar{x} - a}{\sigma/\sqrt{n}} \sim N(0;1)$, следовательно, известна вероятность сл. соб. $P\{\xi < t\}$ $\forall t$.

Найдём: $P\left\{\left|\frac{\bar{x} - a}{\sigma/\sqrt{n}}\right| < t\right\} = P\left\{-t < \frac{\bar{x} - a}{\sigma/\sqrt{n}} < t\right\} = F_{\xi_0}(t) - F_{\xi_0}(-t) =$

$= 0.5 - \Phi(t) - 0.5 - \Phi(-t) = 2\Phi(t)$

$\boxed{\begin{array}{l} \Phi(t) \text{ - нечёт} \\ \xi_n \sim N(0;1) \end{array}}$

Пусть γ — доверительная вероятность, подберём значение t так, чтобы $2\Phi(t) = \gamma$



Тогда для найденного t получим: $P\left\{-t < \frac{\bar{x} - a}{\sigma/\sqrt{n}} < t\right\} =$

$= P\left\{-t\frac{\sigma}{\sqrt{n}} < \bar{x} - a < t\frac{\sigma}{\sqrt{n}}\right\} =$

$= P\left\{\bar{x} - t\frac{\sigma}{\sqrt{n}} < a < \bar{x} + t\frac{\sigma}{\sqrt{n}}\right\} =$

$= P\{\tilde{a}_1 < a < \tilde{a}_2\} \geq \gamma \Rightarrow$

$\Rightarrow$ ДИ для a есть $(\tilde{a}_1; \tilde{a}_2)$

5.5 Построение доверительных интервалов для параметра a нормального распределения при неизвестном σ

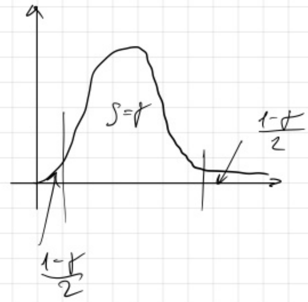
б) Постановка: Построить оценку $(\tilde{\sigma}_1; \tilde{\sigma}_2)$ для неизвестного параметра σ нормального рас-ия; γ – доверительная вероятность задана.

Решение: По лемме Фишера известно, что $\frac{\tilde{\sigma}^2}{nS^2} \sim \chi^2(n-1)$,

где $\chi^2(n) = \xi_1^2 + \ldots + \xi_n^2$; $\xi_i \sim$ н.о.р. сл. в. $\xi_i \sim N(0;1)$

Пусть γ такая, что вероятность того, что $P\{\chi^2(n-1) < q_1\} = \frac{1-\gamma}{2}$

$P\{\chi^2(n-1) > q_2\} = \frac{1-\gamma}{2}$



Обозначим через $F_{\chi^2}(q)$ – функция рас-ия сл. в. $\chi^2(n-1)$.

Тогда подберём q_1 и q_2 т.ч. $P\left\{q_1 < \frac{\sigma^2}{nS^2} < q_2\right\} = F_{\chi^2}(q_2) - F_{\chi^2}(q_1) = \gamma$

Тогда ДИ для параметра σ будет иметь вид:

$$(\tilde{\sigma}_1; \tilde{\sigma}_2) = \left(S\sqrt{nq_1} < \sigma < S\sqrt{nq_2}\right), \text{ где } S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

6 Проверка статистических гипотез

6.1 Определение статистической гипотезы, критерия, ошибок первого и второго рода, уровня значимости, мощности критерия, критической области.

Статистической гипотез называется предположение о виде наблюдаемого неизвестного распределения либо о неизвестных параметрах известного распределения.

Выдвинутую гипотезу называют основной или нулевой гипотезой. Обозначают H_0.
Предположения, противоречащие нулевой гипотезе, включают в конкурирующую или противоположную гипотезу. Обозначают H_1.

Различают простые гипотезы, состоящие из одного предположения, и сложные гипотезы.
Примеры простых гипотез H_0 :" $M\xi = 0$ ", H_0 :" $\xi \sim \mathcal{N}(0,1)$ ".
Примеры сложных гипотез H_1 :" $M\xi > 0$ ", H_1 :" $\xi \sim \mathcal{N}(a,\sigma)$ ".

Критерием проверки гипотез называют правило, согласно которому выдвинутую гипотезу принимают или отвергают.
Критерием назовем случайную величину $K = f(X_1, X_2, \dots, X_n)$--функцию от элементов выборки, закон распределения которой известен, и которую используют для проверки гипотез.

Прежде чем переходить к построению критериев, введем некоторые понятия.

Очевидно, что выдвинутая гипотеза H_0может быть как ложной так и истинной. Если гипотеза H_0 истинна, и ее приняли, либо гипотеза H_0 ложная и ее отвергли, то ошибки не произошло. В оставшихся случаях возникает одна из двух ошибок.

Ошибка первого рода при принятии гипотезы заключается в том, что нулевая гипотеза H_0 была верна, но ее отвергли.
Вероятность ошибки первого рода обозначают α, и называют **уровнем значимости критерия**.

Итак, уровень значимости критерия означает вероятность отвергнуть верную нулевую гипотезу.

Уровень значимости выбирают близким к нулю, как правило 0.05 или 0.01.

Ошибка второго рода при принятии гипотезы состоит в том, что нулевая гипотеза H_0 была ложной, но ее приняли.

Вероятность ошибки второго рода обозначается β.

Мощностью критерия называют величину β, то есть вероятность отвергнуть ложную нулевую гипотезу.

Рассмотрим выборочное пространство $\mathcal{X}_n = \{\overline{X}_n = (X_1, X_2, \dots, X_n)\}$ множество выборок объема n из распределения наблюдаемой случайной величины.Разобьем множество $\mathcal{X}_n$ на два подмножества W и $\overline{W}$,
$W \cup \overline{W} = \mathcal{X}_n$.

Если наблюдалась выборка $\overline{X}_n = (X_1, X_2, \dots, X_n) \in W$, то гипотезу H_0 отвергаем, если наблюдалась выборка $\overline{X}_n = (X_1, X_2, \dots, X_n) \in \overline{W}$, гипотезу H_0 принимаем. Множество W назовем **критическим множеством**, множество $\overline{W}$-- **областью принятия гипотезы**.

Основная задача при решении задачи принятия гипотезы, таким образом, сводится к тому, чтобы построить критическую область. Другими словами, указать правило, согласно которому наблюдавшиеся выборки относят к критическому множеству.

6.2 Критерий Неймана-Пирсона

Пусть $\overline{X_n} = (X_1, \ldots, X_n)$ – выборка из некоторого распределения с функцией плотности $f(x, \theta)$.

Функция правдоподобия: $L(X_n, \theta) = \prod\limits_{i=1}^{n} f(x_i, \theta)$.

Пусть выдвинуты 2 гипотезы:

$$H_0: \text{"}\theta = \theta_0\text{"} \quad \text{и} \quad H_1: \text{"}\theta = \theta_1\text{"}.$$

Определение:

Отношением правдоподобия наз-ся $\varphi(\overline{X_n}) = \dfrac{L(\overline{X_n}, \theta_1)}{L(\overline{X_n}, \theta_0)}$

Теорема (Критерий Неймана-Пирсона):

$$P\{\varphi(\overline{X_n}) \geq C_\varphi \mid H_0\} = \alpha,$$

где α – заданный уровень значимости,

C_φ – граница критической области.

6.3 Построение оптимального критерия Неймана-Пирсона для параметра a нормального распределения для двух простых гипотез

Пусть $X_1, \dots, X_n \sim N(a; \sigma^2)$; σ^2 - известно

Гипотеза: H_0: "$a = a_0$" ; H_1: "$a = a_1$" $\qquad a_0 < a_1$

Решение:

Функция правдоподобия для $N(a;\sigma^2)$ имеет вид:

$$L(\overline{X}_n, a, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sqrt{\sigma}} \exp\left(-\frac{(x_i-a)^2}{2\sigma}\right) = (2\pi)^{-\frac{n}{2}} \cdot \sigma^{-n} \cdot \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a)^2\right)$$

Отношение правдоподобия представляет собой:

$$\varphi(\overline{X}_n) = \frac{L(\overline{X}_n, a_1, \sigma^2)}{L(\overline{X}_n, a_0, \sigma^2)} = \frac{(2\pi)^{-\frac{n}{2}} \cdot \sigma^{-n} \cdot \exp\left(\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a_1)^2\right)}{(2\pi)^{-\frac{n}{2}} \cdot \sigma^{-n} \cdot \exp\left(\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a_0)^2\right)} =$$

$$= \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[(x_i-a_1)^2 - (x_i-a_0)^2\right]\right) = \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[(a_1-a_0)^2 - 2(a_1-a_0)x_i\right]\right) =$$

$$= \exp\left(-\frac{1}{2\sigma^2}\left[n(a_1^2-a_0^2) - 2(a_1-a_0)\sum_{i=1}^{n}x_i\right]\right)$$

По критерию $\varphi(\overline{X}_n) \geq C_\varphi$:

$$\exp\left(-\frac{1}{2\sigma^2}\left[n(a_1^2-a_0^2) - 2(a_1-a_0)\sum_{i=1}^{n}x_i\right]\right) \geq C_\varphi$$

Следовательно:

$$\sum_{i=1}^{n} x_i \geq \frac{\sigma^2}{a_1-a_0}\ln C_\varphi + \frac{n(a_1+a_0)}{2}$$

Рассмотрим случайную величину $Y = \sum_{i=1}^{n} x_i \sim N(a, \sigma^2)$.

Найдём МО и D Y:

$$\left.\begin{aligned} MY &= M\left(\sum_{i=1}^{n}x_i\right) = \sum_{i=1}^{n}Mx_i = na \\ DY &= D\left(\sum_{i=1}^{n}x_i\right) = \sum_{i=1}^{n}Dx_i = n\sigma^2 \end{aligned}\right\} \Rightarrow Y \sim N(na; n\sigma^2) \Rightarrow$$

$$\Rightarrow \frac{Y-na}{\sqrt{n}\cdot\sigma} \sim N(0,1) \text{ с } F(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt$$

Таким образом, если верна H_0, то $\frac{Y-na_0}{\sqrt{n}\sigma} \sim N(0,1)$,

если - H_1, то $\frac{Y-na_1}{\sqrt{n}\sigma} \sim N(0,1)$.

Положим, что H_0 верна:

Пусть α - задано и H_0 "+". Тогда:

$$P\{\varphi(\bar{X}_n) \geq C_\varphi\} = P\{\sum_{i=1}^{n} X_i \geq C_\varphi\} = P\left\{\frac{\sum_{i=1}^{n} X_i - na}{\sigma\sqrt{n}} - \frac{C - na}{\sigma\sqrt{n}}\right\} = \alpha$$

квантиль рас-ия уровня $1-\alpha$

Пусть $u_{1-\alpha}$ $\quad P\{\xi_0 \geq u_{1-\alpha}\} = 1-\alpha$.

Пусть для заданного α найден $u_{1-\alpha}$. Тогда $\frac{C - na}{\sigma\sqrt{n}} = u_{1-\alpha}$, следовательно,

$$C_\varphi = u_{1-\alpha} \cdot \sigma\sqrt{n} + na_0 \text{ — граница крит. области.}$$

Критерий: Если $\sum_{i=1}^{n} X_i \geq na_0 + \sigma\sqrt{n}\, u_{1-\alpha}$, то H_0 "−"

Если $\sum_{i=1}^{n} X_i < na_0 + \sigma\sqrt{n}\, u_{1-\alpha}$, то H_0 "+".

6.4 Построение оптимального критерия Неймана-Пирсона для двух простых гипотез для параметра показательного распределения

6.5 Проверка сложных гипотез

6.6 Проверка простой гипотезы H_0 против сложной гипотезы H_1

6.7 Критерий согласия

Пусть $X_1, \ldots, X_n \sim L(x, \theta)$, причём закон $L(x, \theta)$ - неизвестен.

Выдвинем гипотезу H_0: "$\xi \sim N(a, \sigma^2)$" и H_1: "$\xi \nsim N(a, \sigma^2)$"

Критерий согласия — критерий проверки гипотез о согласованности поведения наблюдаемых данных с предполагаемым распределением.

6.8 Критерий χ^2 и критерий Колмогорова

Пусть $\overline{X_n} = (X_1, \ldots, X_n) \sim F_0(x)$, где $F_0(x)$ – гипотетическое предполагаемое неизвестное распределение.

$$H_0: \text{“}\xi \sim N(a, \sigma^2)\text{”} = \text{“}F_\xi(x) = F_0(x)\text{”}$$

ЭФР: $\widehat{F_n}(x) = \frac{1}{n} \sum_{i=1}^{n} e(x - x_i) = \frac{\mu(x)}{n}$

Критерий Колмогорова:

$$D_{\text{набл}} = \sup_{n} |\widehat{F_n}(x) - F_0(x)|$$

Если $D_{\text{набл}} < D_{\text{крит}}$ (квантиль рас-ия Колмогорова), то гипотеза H_0 “+”, иначе “–”.

Критерий χ^2:

$$\chi^2_{\text{набл}} = \sum_{i=1}^{m} \frac{(n_i - n_i')^2}{n_i'},$$

где m – кол-во интервалов (точек),
n_i – наблюдаемая частота
n_i' – теоретическая частота (вычисленная в соответствии с H_0)

$\chi^2_{\text{крит}} = \chi^2(\alpha, m-1-k)$, где k – кол-во параметров распределения $F_0(x)$, которые оцениваются по выборке.

При $\chi^2_{\text{набл}} < \chi^2_{\text{кр}}$ H_0 “+”, иначе “–”

7 Регрессионный анализ

7.1 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.

7.2 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 $\hat{a}$

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

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

Отсюда получаем оценку параметра b:

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

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 $(\bar{x}, \bar{y})$

7.3 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 $\iff |\bar{r}| = 1$.

7.4 Проверка гипотезы о значимости выборочного коэффициента корреляции

Рассм. выборочный коэфф-т корреляции $r = \frac{\overline{xy} - \overline{x}\cdot\overline{y}}{S_x S_y}$, 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 $\alpha = (0{,}05; 0{,}01)$.
$t_{набл.} = r\sqrt{\frac{n-2}{1-r^2}}$ $t_{кр.} \sim T(\frac{\alpha}{2}; n-2)$ – 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.

7.5 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) $M\varepsilon_t = 0$ $M\varepsilon_t^2 = D\varepsilon_t = \sigma^2$ 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.

7.6 Статические свойства МНК оценок

Статистические свойства МНК – оценок.

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

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

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

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

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..

7.7 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

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

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

Total Sum of Squares Error Sum of Squares 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 $F(n,m) = \frac{\frac{1}{n}(\xi_1^2 + \ldots + \xi_n^2)}{\frac{1}{m}(\eta_1^2 + \ldots + \eta_m^2)}$