Утв. 2. Центральная предельная теорема

$X_1, \dots, X_n, \dots$ — послед. незав. СВ, $X_i \sim X$,

$\exists\, EX_n = EX = a,\ DX_n = DX = \sigma^2$

$$Y_n = \frac{\overline{X_n} - E(\overline{X_n})}{\sigma(\overline{X_n})} \overset{\text{см. док-во в 3.5.2}}{=\!=} \frac{\overline{X_n} - a}{\sigma}\sqrt{n} \xrightarrow[n\to\infty]{d} Y$$

$$F_{Y_n}(t) \overset{R}{\underset{n\to\infty}{\rightrightarrows}} F_Y(t),\ Y \sim N(0,1)$$

Док-во 3.5.4 в форме Хинчина

I Достаточно док-ть, что $\varphi_{\overline{X}_n}(t) \xrightarrow[n\to\infty]{} \varphi_{\delta_a}(t)$

δ_a — распред., сосредот. в (·) a, т.е.

$P(\delta_a = a) = 1$

(*) по т. непр. $\Rightarrow F_{\overline{X}_n}(t) \xrightarrow[n\to\infty]{} F_{\delta_a}(t)\ \forall$ точке непр. $F_{\delta_a}(t)$

$\forall\ u = u(x)$ — огр. и непр.

$$E(u(\overline{X_n})) \to E(u(\delta_a)) = u(a) \overset{\text{по т. о сх-ти МО}}{\Longleftrightarrow}$$

$$\Leftrightarrow \overline{X_n} \xrightarrow[n\to\infty]{\text{п.н.}} a$$

II $\varphi_{\delta_a}(t) = E e^{it\delta_a} = e^{ita} \cdot 1 = e^{ita}$

III Нужно док-ть, что $\varphi_{\overline{X_n}}(t) \xrightarrow[n\to\infty]{} e^{ita}$

NB $\varphi_{X_i}(t) \equiv \varphi_X(t)\ \forall i$

$\varphi_X(0) = 1$, $\varphi_X(t)$ — непр. ф-ия (равн. непр.)

для $\varepsilon = \frac{1}{2}$ $\exists\, h$: $|\varphi_{x_n}(t) - \varphi''^{1}_{x_n}(0)| \le \frac{1}{2}$

при $|t| < h$ $\Rightarrow$ $\varphi''_x(t)$

$\Rightarrow$ опред. ф-ия $\ell(t) = \ln \varphi_{x_n}(t)$ – главное знач. log

19.05

$\overline{X}_n \xrightarrow[x\to\infty]{P} EX = a$

$\varphi_{\overline{X}_n}(t) \to \varphi_{\delta_a}(t) = e^{ita}$ $\forall t$

$\varphi(0) = 1$,

$\varepsilon = \frac{1}{2}$, $|\varphi_{x_n}(t) - \varphi_{x_n}(0)| \le \frac{1}{2}$ при $|t| < h$

При $|t| < h$ $\ell(t) = \ln \varphi_{x_n}(t)$

$\exists$ $EX_n = a \Rightarrow \exists$ $\ell'(0) = \frac{\varphi'_{x_n}(0)}{\varphi_{x_n}(0)} = \frac{iEX_n}{1} = ia$

$\ell(0) = \ln \varphi_{x_n}(0) = 0$

$\forall$ фикс. t_0 возьмем дост. большое N :

$\left|\frac{t_0}{N}\right| < h \Rightarrow \ell\left(\frac{t_0}{N}\right)$ – определено и

(n > N аналог.)

$\varphi_{\overline{X}_n}(t_0) = \varphi_{\frac{x_1+\ldots+x_N}{n}}(t_0) = \varphi_{x_1+\ldots+x_N}\left(\frac{t_0}{N}\right) \overset{\text{незав}}{=}$

$= \left(\varphi_x\left(\frac{t_0}{N}\right)\right)^N$

$$\varphi_X\left(\frac{t_0}{N}\right) = \varphi_{X_n}\left(\frac{t_0}{N}\right) \ \forall n$$

$$\varphi_{\overline{X}_n}(t_0) = e^{\left(\ln \varphi_X\left(\frac{t_0}{N}\right)\right) N} = e^{\ell\left(\frac{t_0}{N}\right)\cdot N}$$

$$\underset{N\to\infty}{\nearrow 0}$$

$$e^{\ell\left(\frac{t_0}{N}\right)\cdot N} = \exp\left(\left(\ell\left(\frac{t_0}{N}\right) - \ell(0)\right)\cdot t_0 \Big/ \frac{t_0}{N}\right) \xrightarrow[N\to\infty]{}$$

$$\xrightarrow[N\to\infty]{} \exp(t_0 \ell'(0)) = e^{at_0}$$

$$\forall t_0 : \varphi_{\overline{X}_n}(t_0) \to e^{at_0} = \varphi_{\delta_a}(t_0) \Rightarrow (\text{нпд})$$

$$\Rightarrow \overline{X}_n \xrightarrow[n\to\infty]{P} a$$

ЦПТ для незав., один. ~~незав~~ распр. сл. велич. (НОР)

$X_1, \dots, X_n, \dots$ – послед. незав., $X_n \sim X$ – ген. совокуп.

и $\exists$ $EX = a$, $DX = \sigma^2$, тогда

$$Y_n = \frac{\overline{X}_n - E\overline{X}_n}{\sigma(\overline{X}_n)} \overset{*}{=} \frac{\overline{X}_n - a}{\sigma}\sqrt{n} \xrightarrow[n\to\infty]{d} Y \sim N(0,1)$$

или $F_{Y_n}(t) \xrightarrow[n\to\infty]{R} F_{N(0,1)}(t)$

$*$ – показывает при док-ве сл. з.б.ч.

$\Rightarrow$ следует из непр. $F_{N(0,1)}(t)$

Док-во:

19.05

Док-во:

I. $a = 0$, иначе вместо X_n $\triangleleft$ $\tilde{X}_n = X_n - a$

$E\tilde{X}_n = 0$, $D\tilde{X}_n = DX_n$

$\cancel{\tilde{Y}_n = \frac{(X_1 - a + \dots + X_n - a)/n - E\tilde{X}_n}{\sigma X_n}}$

$$\tilde{Y}_n = \frac{((X_1 - a) + \dots + X_n - a))/n - \overset{=0}{E\bar{\tilde{X}}_n}}{\sigma X_n} = \frac{\bar{X}_n - a}{\sigma}\sqrt{n}$$

Считаем, что $a = 0 = EX$

Дост. $\varphi_{Y_n}(t) \xrightarrow[n\to\infty]{} \varphi_{N(0,1)}(t) = e^{-\frac{t^2}{2}} \quad \forall t$

$$Y_n = \frac{X_1 + \dots + X_n - \overset{=0}{an}}{n\sigma}\cdot\sqrt{n} = \frac{X_1 + \dots + X_n}{\sigma\sqrt{n}}$$

Х.ф. $\varphi_{Y_n}(t) = \varphi_{X_1 + \dots + X_n}\left(\frac{t}{\sigma\sqrt{n}}\right) \overset{\text{незав.}}{\underset{\text{одинак. распр.}}{=}} \varphi_X^n\left(\frac{t}{\sigma\sqrt{n}}\right) ==$

(малое) а то хер поймешь

$$\varphi_X\left(\frac{t}{\sigma\sqrt{n}}\right) = \varphi_X(0) + \varphi'(0)\frac{t}{\sigma\sqrt{n}} + \overset{=-\sigma^2}{\varphi''(0)}\frac{t^2}{2\sigma^2 n} + o\left(\frac{t^2}{n}\right) =$$

$$= 1 + i\overset{=0}{a}\frac{t}{\sigma\sqrt{n}} - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)$$

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

$$\ln\varphi_{Y_n}(t) = n\ln\left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right) \sim$$

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

$$\Rightarrow \varphi_{Y_n}(t) \xrightarrow[n\to\infty]{} e^{-\frac{t^2}{2}} \sim \varphi_{N(0,1)}(t)$$

Метод Монте-Карло (пример исп. пред. теоремы т.в. для точечного и доверительн. оценивания)

Из з.б.ч. для $\forall$ послед. $X_1, \dots, X_n, \dots$ (незав. одинак. н.о.р. распр.)

$\exists\ EX_n = a$

$\overline{X_n} = \frac{X_1 + \dots + X_n}{n} \xrightarrow[n\to\infty]{} a$, $\exists\ DX_n = \sigma^2$, то

$\frac{\overline{X_n} - a}{\sigma}\sqrt{n} \xrightarrow[n\to\infty]{d} Y \sim N(0,1)$

Находим „n" незав. знач. СВ X (рассм. n незав. $X_1, \dots, X_n, \dots$ $X_n \sim X$)

$\frac{x_1 + \dots + x_n}{n} \xrightarrow[n\to\infty]{} a$ и $\overline{X_n}$ — точечной оценкой $EX = a$

$F_n(t) = P\left(\frac{X_n - a}{\sigma}\sqrt{n} < t\right) \xrightarrow{d \to Y} F_{N(0,1)}(t)$

$P(|Y_n| < t) = P(-t < Y_n < t) =$

$= P\left(\overline{X_n} - \frac{t\sigma}{\sqrt{n}} < a < \frac{t\sigma}{\sqrt{n}} + \overline{X_n}\right) \xrightarrow[n\to\infty]{} F_{N(0,1)}(t) -$

$- \underset{\substack{\parallel \\ 1-\Phi(t)}}{F_{N(0,1)}(-t)} = 2\Phi(t) - 1 = \gamma$

$P(\overline{X_n} - \delta_n < a < \overline{X_n} + \delta_n) \xrightarrow[n\to\infty]{} \gamma$, $\delta_n : t - t_{\frac{1+\gamma}{2}}$ –

квантиль $\Phi(t)$

$\delta_n = \frac{t_{\frac{1+\gamma}{2}} \cdot \sigma}{\sqrt{n}} + \overline{X_n}$

1

$\overline{\sigma_n} =$

$E(X^2)$: $\dfrac{X_1^2 + \ldots + X_n^2}{n} = \overline{X_n^2}$ — оценка $E(X^2)$

$\Rightarrow S_n^2 = \dfrac{\sum_{i=1}^{n} X_i^2}{n} - (\overline{X_n})^2 = \dfrac{\sum X_i^2 - n \overline{X_n}^2}{n}$

$\hat{\sigma}_n^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2$ — несмещенная оценка дисперсии

$(E_X \hat{\sigma}_n^2 = DX)$

$\Rightarrow$ ЦПТ позволяет строить асимпт. доверит. интервал для $EX = a$

$\hat{a}_n = \overline{X}_n$; ас. доверит. интервал надежности γ : $(\overline{X}_n - \delta_n, \overline{X}_n + \delta_n)$

$\delta_n = \dfrac{t_{\frac{1+\gamma}{2}} \cdot \hat{\sigma}_n}{\sqrt{n}}$

$\overline{X}_n$ — mean, $\hat{\sigma}_n$ — std cond

1. Оценить объем тела методом монте-карло

Геом. вер. $\forall$: $D \in \Omega$: $p(D) = \dfrac{\mu(D)}{\mu(\Omega)}$

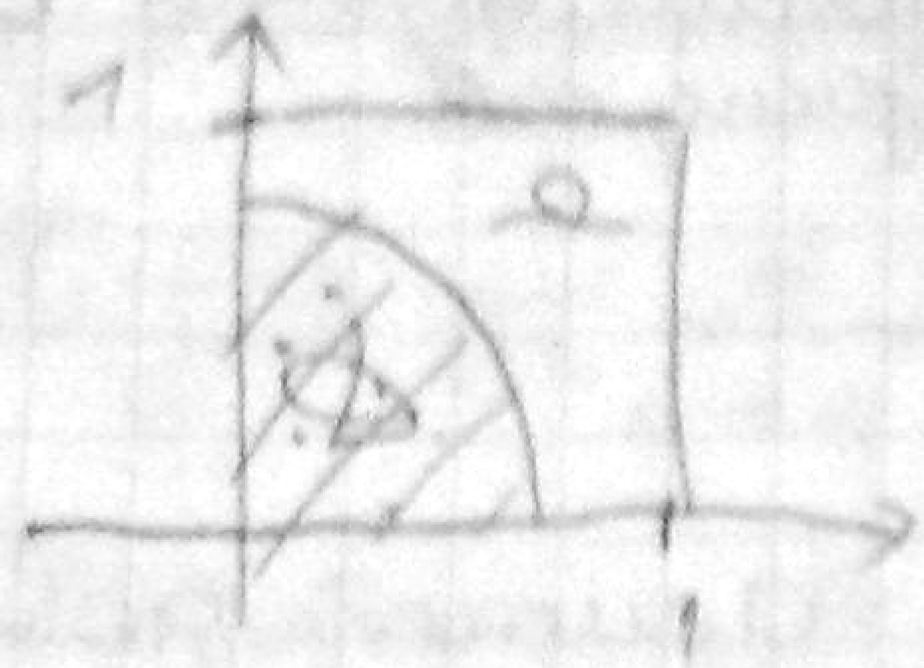
$\mu(D) = p(D) \cdot \mu(\Omega)$

$\uparrow$ объем

$\xi \sim B_{1,p}$, т.е. $\begin{array}{c|c|c} \xi_i & 0 & 1 \\ \hline p_i & 1-p & p \end{array}$, то $\qquad p = p(A)$

$\frac{\xi_1+\ldots+\xi_n}{n} \xrightarrow[n\to\infty]{} p(A)$, т.е. $\qquad E\xi = p$, $D\xi = pq$

$\frac{K_n}{n}$ (частота) асим. оценкой вер-ти событий $p(A)$



"n" — кол-во незав. точек из Ω

$\forall$ (·) имеет $B_{1,p}$

K_n — кол-во (·), попавших в обл

$\Rightarrow$ оценка $\underbrace{\mu(\delta) \Leftrightarrow \frac{K_n}{n}\mu(\Omega)}_{\hat{p}_n \text{ (оценка } p)}$

Тест: проверьте $x^2+y^2 \le 1$, $x \ge 0, y \ge 0$ (при $n >>$) → точечная оценка объема

$\sigma = \sqrt{pq} \Leftrightarrow \sqrt{\hat{p}_n(1-\hat{p}_n)}$ — оценка σ "$\hat{\sigma}_n$"

2. $I = \int_{-\infty}^{+\infty} g(x) \underbrace{e^{-\frac{(x-a)^2}{2\sigma^2}}}_{\text{ядро}} dx = \sigma\sqrt{2\pi}\left(\left(\frac{1}{\sigma\sqrt{2\pi}}\right)\int_R g(x) e^{-\frac{(x-a)^2}{2\sigma^2}} dx\right)$;

$e^{-\frac{(x-a)^2}{2\sigma^2}} \Rightarrow \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}$

$= \sigma\sqrt{2\pi}\int_R g(x) f_{N(a,\sigma^2)}(x)\, dx =$

$= \sigma\sqrt{2\pi}\, Eg(X)$, $X \sim N(a,\sigma^2)$

$I \Leftrightarrow \sigma\sqrt{2\pi} \cdot \overline{g(X_n)}$

$g(X_1) \ldots g(X_n)$

для $Eg(X)$ строить асимпт. доверит. интервал

ммт-для I: $\sigma\sqrt{2\pi}$ · мм-λ для $Eg(X)$
вол$\int$, аналог: $I \int_{-\infty}^{+\infty} g(x) e^{-\frac{x^2}{4}} \frac{1}{dx} \approx \frac{1}{4}$

Эл-ты мат. стат-ки. 26.05

Вариант: выборочная стат-ка, асимптот. вар-т ~~вар-ты~~

Лекции по мат стат. Нат. Чернова (сайт НГУ)

Задачи, обратные к $\overset{\text{ТВ}}{\cancel{\text{мат. ст-ке}}}$.
(по набору данных восст. распределение)

$\exists$ X – СВ с неизвестным распред.

$\forall$ $X_1, \dots, X_n$ – набор независимых одинак. распр. величин, $X_i \sim X$

Можно считать, что $X_1, \dots X_n$ – n копий СВ X

$\mathcal{X}_n = \{(x_1, \dots, x_n)\} \subseteq R_n$ – выборочн. пр-во, зад. Σ-ал.

Возникает т.н. статистич. эксперимент:

$\{\mathcal{X}_n, \Sigma_\sigma, P_\theta, \theta \in \Theta\}$ –

Θ – набор допустимых пар-ов, $P_x \Leftrightarrow \theta$

$\Theta = R_+ \times R$ – параметры $X \sim N(a, \sigma^2)$

$\Theta = \{F_x - \forall$ ф-ция распр-ия$\}$

Основ. задачи МСт.

I. оценивание пар-ра θ, $g(\theta)$ – хар-ки

$\theta = (a, \sigma^2)$, $g(\theta) = E_\theta(X^3)$

а) точечное оценивание

в). доверительное оценивание

II. Проверка гипотез (обнаружение)

Точечное оценивание пар-ра, хар-ки

g_n – оценки $g(\theta)$

$g_n : \mathcal{X}_n \longrightarrow \mathbb{R}$

g_n – $\forall$ измеримая ф-ия, отображ. стат. эксперимент в соотв. пр-во

g_n – γ статистикой, оценка $g(\theta)$ – $\forall$ статистика

Св-ва оценки:

1. несмещенность (не асимпт. св-во)

$E_X(g_n) = E_\theta(g_n) = g(\theta)$ – среднее по мере θ ~~...~~ P_θ совпадает с оцен. хар-ти

2. состоятельность (или стремление к нулю $R_\delta(g_n, g(\theta))$:

$g_n \xrightarrow[n \to \infty]{P} g(\theta)$ ($\to$ к 0 δ-риска ??)

3. Эффективность

g_n – эфф., если она не смещ. $E_X(g_n) = g(\theta)$

$f_x(x, \theta)$ – либо плотность (непр.), либо
набор вер-ей (дискрет.)

$l(x, \theta) = \ln f_x(x, \theta)$

$I(\theta)$ – информация Фишера (одномерная)

$I(\theta) = E_\theta\left[\left(l'_\theta(x, \theta)\right)^2\right]$

Нер-во Рау-Крамера:

Для несмещенной оценки

$$D_\theta(g_n) \geq \frac{1}{n \cdot I(\theta)}$$

Эфф. ɣ оценка, для которой в нер-ве Р.-К. достигается равенство.

Эфф. оценка – это несмещенная оценка с минимальной дисперсией.

4. Асимптотическая нормальность

– позволяет строить доверительные обл.

Точечное оценивание

$g_n; g_{n_1}, \ldots \cancel{g_n}$ – оценки хар-ки $g(\theta)$

Качество оценок — функции потерь и функции риска (среднее потери)

Обычно $\exists$ 2 функции потерь

1. функция потерь

$$W_\delta(g_n, \theta) = \begin{cases} 0, & \|g_n - g(\theta)\| < \delta \\ 1, & \|g_n - g(\theta)\| \geq \delta \end{cases}$$

$$R_\delta(g_n, \theta) = E_\theta(W_\delta(g_n, \theta)) = \int\limits_{\|g_n - g(\theta)\| \geq \delta} 1 \cdot P_\theta(x) =$$

$$= P_\theta(\|g_n - g(\theta)\| \geq \delta)$$

g_n — состоят., если $R_\delta(g_n, \theta) \xrightarrow[n\to\infty]{\forall \delta > 0} 0$ —

т.е. $g_n \xrightarrow[n\to\infty]{P} g(\theta)$

2. $W(g_n, g(\theta)) = \|g_n - g(\theta)\|^2$

$$R_2(g_n, g(\theta)) = E_\theta(\|g_n - g(\theta)\|^2)$$

Связь м/у R_δ и R_2

Утв. Для сост. оценки достаточно, что $R_2(g_n, g_n(\theta)) \xrightarrow[n\to\infty]{} 0$

$$R_\delta(g_n, g(\theta)) = \int\limits_{\|g_n - g(\theta)\|} P_\theta(dx) = P_\theta(\|g_n - g(\theta)\|^2 > \delta^2) \overset{\text{нер-во Маркова}}{\leqslant}$$

$$\leqslant E_\theta(\|g_n - g(\theta)\|^2)/\delta^2 = \frac{R_2(g_n, g(\theta))}{\delta^2}$$ квадратичный риск

Вид квадратичного риска для одномерного пар-ра.

$\Theta \subset \Omega \subseteq \mathbb{R}$, $\|\theta_n - \theta\| = |\theta_n - \theta|$ и

$$R_2(\theta_n, \theta) = E_\theta(|\theta_n - \theta|^2) =$$

$$= E_\theta\Big((\theta_n - E_\theta(\theta_n) + \underbrace{E_\theta(\theta_n) - \theta}_{\text{const}})^2\Big) =$$

$$= D_\theta(\theta_n) + E_\theta\underbrace{(E_\theta(\theta_n) - \theta)^2}_{(E_\theta(\theta_n) - \theta)^2} + 0 =$$

$$= D_\theta(\theta_n) + b_n^2$$

смещение оценки

$$b_n = E_\theta(\theta_n) - \theta$$

θ_n – несмещ. одномерная оценка, то

$$R_2(\theta_n, \theta) = D_\theta(\theta_n)$$

Вторая задача мат. ст.

Проверка гипотез.

тета с 0 ↓ тета ↓

H_0 – основная гипотеза : $\theta \subset \Theta_0 \subset \Theta$

H_1 – альтернатив. гип. : $\theta \in \Theta_1 \subset \Theta$

$$\Theta_0 \cap \Theta_1 \neq \varnothing, \quad \Theta_0 \cup \Theta_1 = \Theta \Rightarrow$$

$\Rightarrow$ проверка согласия с H_0.

Если Θ_i сост. из единств. эл-та, тогда ГН H_i простой, иначе сложной

Проверка гипотез осуществ. при помощи критерия (теста)

$$\text{тест} - \varphi_n(x_1, \dots, x_n) = \begin{cases} 0, & H_0 \\ 1, & H_1 \end{cases}$$

$$\varphi_n : X_n \to \{0, 1\}$$

X_0 – мн-во : $\varphi_n(\overset{\overline{x}}{\overbrace{x_1, \dots, x_n}}) = 0 \;\; \forall \overline{x} \in X_0$ – допустимое мн-во

$\mathcal{X}_1 : \varphi_n(\bar{x}) = 1 \in \mathcal{X}_1$ — критич. мн-во

$\mathcal{X}_0 \cup \mathcal{X}_1 = \mathcal{X}_n$

Качество теста хар-ся при помощи ошибок

Опр. Ошибка I рода: H_0 - верна, но $\overset{(\text{тест}=1)}{\text{отвергается}}$

Ошибка II рода: H_0 - неверна, но принимается)

Числен. хар-ки теста - вер-ти ошибок I и II родов

$\alpha(\varphi_n, \theta) = P_{H_0}(\mathcal{X}_1), \ \theta \in \mathcal{X}_0$ - ош. I р.

$\beta(\varphi_n, \theta) = P_{H_1}(\mathcal{X}_0), \ \theta \in \mathcal{X}_1$ - ош. II р.

Мощность теста: $\gamma(\varphi_n, \theta) = 1 - \beta(\varphi_n(\theta))$

Обычно тест базируется на тестов. стат-ке L

$L(x_1, \dots, x_n) : \mathcal{X}_n \to \mathbb{R}$ и порог $T \in \mathbb{R}$ (число)

$$\varphi_n(x_1, \dots x_n) = \begin{cases} 0, & L(\bar{x}) < T \\ 1, & L(\bar{x}) \geq T \end{cases}$$

$0 \leq \alpha(\varphi_n, \theta) + \beta(\varphi_n, \theta) \leq 1$

] $\exists \ F_L(x)$, тогда $\cancel{\alpha(\varphi_n, \theta) = \{ 0, P_{\theta \in \Theta_0}}$

$\alpha(\varphi_n, \theta) = P_{\theta \in \Theta_0}(L(\bar{x}) \geq T) = 1 - F_L(T), \searrow, T \uparrow \infty$

$\beta(\varphi_n, \theta) = P_{\theta \in \Theta_1}(L(\bar{x}) \leq T) = F_L(T) \uparrow, T \to \infty$

(от того, какой порог, зависит велич. ошибок)

Подход Неймана – Пирсона

фикс. α, $\measuredangle$ критерий $\Phi_\alpha = \{\varphi_n, \alpha(\varphi_n, \theta) \leq \alpha\}$

Среди всех $\varphi_n \in \Phi_\alpha$ находят $\min \beta(\varphi_n)$ или

$\beta(\varphi_n) \xrightarrow[n\to\infty]{} 0$ — АС. вар.

$\varlimsup\limits_{n\to\infty} \alpha(\varphi_n, \theta) \leq \alpha$, $\Phi_\alpha = \{\varphi_n, \varlimsup\limits_{n\to\infty} \alpha(\varphi_n, \theta) \leq \alpha\}$

Критерий $\varphi_n : \varlimsup\limits_{n\to\infty} \alpha(\varphi_n, \theta) \leq \alpha$ γ критерием асимпт. ур-ня значимости α.

Если $\beta(\varphi_n, \theta) \xrightarrow[n\to\infty]{} 0$ ($\varlimsup \varphi_n \leq \alpha$) γ сост. критерием асимпт. уровня значим. α

Задача оценивания сост. из построения оценки и изучения ее св-в.
Основ. методы постр.:

1. Выбороч. метод – общий, базируется на св-х $F_n(t)$ – эмпирич. ф-ии распред.
2. Метод моментов
3. Метод макс. правдоподобия — нужно знать вид F_X (н-р, $F_X \sim N(a, \sigma^2)$)

$X_1, \dots, X_n$ – выборка из ген. сов-ти X, возник.

вариац. ряд: $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$

Y_n	$X_{(1)}$	$X_{(2)}$	$\dots$	$X_{(n)}$
p_i	$\frac{1}{n}$	$\frac{1}{n}$	$\dots$	$\frac{1}{n}$

– вспомогат. дискр. СВ

Выб. метод постр.: $g(\theta) = g(X) \Rightarrow$ соотв. хар-ки СВ Y

Выб. оценка EX: $\overline{X_n} = EY = \frac{1}{n}\sum\limits_{i=1}^{n} X_{(i)} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ – выб. среднее

2.06. Точечное оценивание ✓

Задача довер оценивания.

$\hat{\theta} \in \Theta$ $\quad \hat{g} \in G$ – точечная оценка

Θ_γ, G_γ – обл, в которой с вер-ю (надежностью) γ лежит реальное значение пар-ра θ или характеристики $g(\theta)$

При $\Theta \in R_1 \Rightarrow \Theta_\gamma$ – доверит. интервал надежности γ

$\Theta_{n,\gamma}$, $G_{n,\gamma}$, $I_{n,\gamma}$ – области: при $n \to \infty$

$P_{n,\theta}(\theta \in \Theta_{n,\gamma}) \xrightarrow[n\to\infty]{} \gamma \Rightarrow \Theta_{n,\gamma}$ – асимп. дов. мн-во надежности γ

$g(\theta) \in G_{n,\gamma}$

Для постр. $I_{n,\gamma}$ – ас. дов. инт. используют ас. нормальность (см пример – метод М-К)

ЭФР (~~эпк~~ эмпир. ф-ия распр.)

Оценка $F_X(t)$ – незав. функц. распр-ия сл. вел-ны X у ЭФР и $F_n(t) = \frac{1}{n}\sum_{i=1}^{n} \Delta_{(-\infty,t)}(X_i)$

$\Delta_{(-\infty,t)}(X_i) = \begin{cases} 1, & X_i < t \\ 0, & X_i \geq t \end{cases}$; $F_n(t)$ – частота попадания в интервал $(-\infty, t)$

ЭФР – кусочно-непр. ф-ия (непр. слева)

Y	$X_{(1)}$	$X_{(2)}$	...	$X_{(n)}$
p	$1/n$	$1/n$	...	$1/n$

$F_n(t) = F_Y(t)$

Св-ва ЭФР

I. В точке $F_n(t_0) = \frac{1}{n}\sum_{i=1}^{n} \Delta_{(-\infty,t_0)}(X_i)$ – кол-во эл-ов выборки в $(-\infty, t_0)/n$ – – частота

$F_n(t_0) = \frac{K_n}{n} = \overline{X}_n$, $\quad X = \xi \sim B_{1,p}$

ξ_i	0	1
p_i	$1-p$	$F_X(t_0) = p$

CB
CV-

1. ~~$\overline{X}_n \xrightarrow[n\to\infty]{P}$~~ 1 $F_n(t_0) = \overline{X}_n \xrightarrow[n\to\infty]{P} E\xi = p = F_x(t_0)$ ~ $F_n(t_0)$ – сост. оценка $F_x(t_0)$

2. $E_\xi \ \overline{\xi}_n = E_\xi \frac{\xi_1 + \ldots + \xi_n}{n} = \frac{1}{n}\sum_{i=1}^{n} E\xi_i = \frac{n}{n} F_x(t_0)$ –
– несмещенность $F_n(t_0)$ как оценки $F_x(t_0)$

3. Из ЦПТ $\Rightarrow$ ас. нормальность: $\frac{\overline{\xi}_n - E_x\overline{\xi}_n}{\sigma_x \overline{\xi}_n} =$
$= \frac{[\overline{\xi}_n - F_x(t_0)]\sqrt{n}}{\sqrt{F_x(t_0)(1-F_x(t_0))}} \xrightarrow[n\to\infty]{d} Y \sim N(0,1)$

$F_x(t_0)$ для дисперсии оцениваем $F_n(t_0)$ – сост.

$P\Big(F_n(t_0) - \delta_n < F_x(t_0) < \underbrace{\frac{T_\gamma\sqrt{F_n(t_0)(1-F_n(t_0))}}{\sqrt{n}}}_{=\delta_n} + F_n(t_0)\Big) \xrightarrow[n\to\infty]{} \gamma$

$T_\gamma = \frac{1+\gamma}{2}$ – квантиль $N(0,1)$

$P\left(|F_x(t_0) - F_n(t_0)| \cdot \frac{\sqrt{n}}{\sqrt{F_n(t_0)(1-F_n(t_0))}} < t\right) \longrightarrow \underbrace{P(|Y| < t)}_{=\gamma}$

$= F_{N(0,1)}(t) - \underbrace{F_{N(0,1)}(-t)}_{=1-F(t)}$

Для $F_n(t_0)$ ас. дов. инт. уровня доверия γ:
$(F_n(t_0) - \delta_n,\ F_n(t_0) + \delta_n)$

2. Оценивание $F_x(t)$ в целом
Св-ва $F_n(t)$
Непарам. задача
Опр. Расстояние Колмогорова
$\rho_\infty(F_n(t), F_x(t)) = \sup_t |F_x(t) - F_n(t)|$

Расстояние Смирнова (Мизеса)
Пример несим. расст.
$\rho_2(F_n, F_x) = \left(\int_R (F_n(t) - F_x(t))^2 \, dF_x(t)\right)^{1/2}$

$\rho_2(F_n, F_x) \leq \left(\sup|F_n(t) - F_x(t)|^2\right)^{1/2} \cdot 1 = \rho_\infty(F_n, F_x)$

$\cancel{\rho_\infty(F_n(t)}$,

Теорема 1. (Гливенко – Кантелли)

] $F_n(t)$ ЭФР постр. по выборке из ген. сов-ти X и $F_x(t)$ – ф-ия распр. (неизвестная), тогда

$\rho_\infty(F_n, F_x) \xrightarrow[n\to\infty]{p=1} 0$ $\quad \rho_\infty(F_1, F_x), \rho_\infty(F_2, F_x), \ldots$ – числов. посл-ть

Следствие 1. Расст. Смирнова, т.к. как оно меньше с $p = 1 \longrightarrow 0$ $\left(\rho_2(F_n, F_x) \xrightarrow[n\to\infty]{p=1} 0\right)$

Следствие 2. Из ~~вер.~~ сх-ти с вер. 1 $\Rightarrow$ сх-ть по вер-ти

$\cancel{P\{x\ F_n(x)}$ $P\{x \ |F_n(x) - F_x(x)| > \varepsilon\} \xrightarrow[n\to\infty]{\forall \varepsilon > 0} 0$, т.е.
$F_n(t)$ – сост. оценка $F_x(t)$ по отн. к ρ_∞ и ρ_2

Теорема 2. (Колмогорова)] $F_x(t) \in C(R)$

$d_n = \sqrt{n}\,\rho_\infty(F_n, F_x)$ – дост. стат.

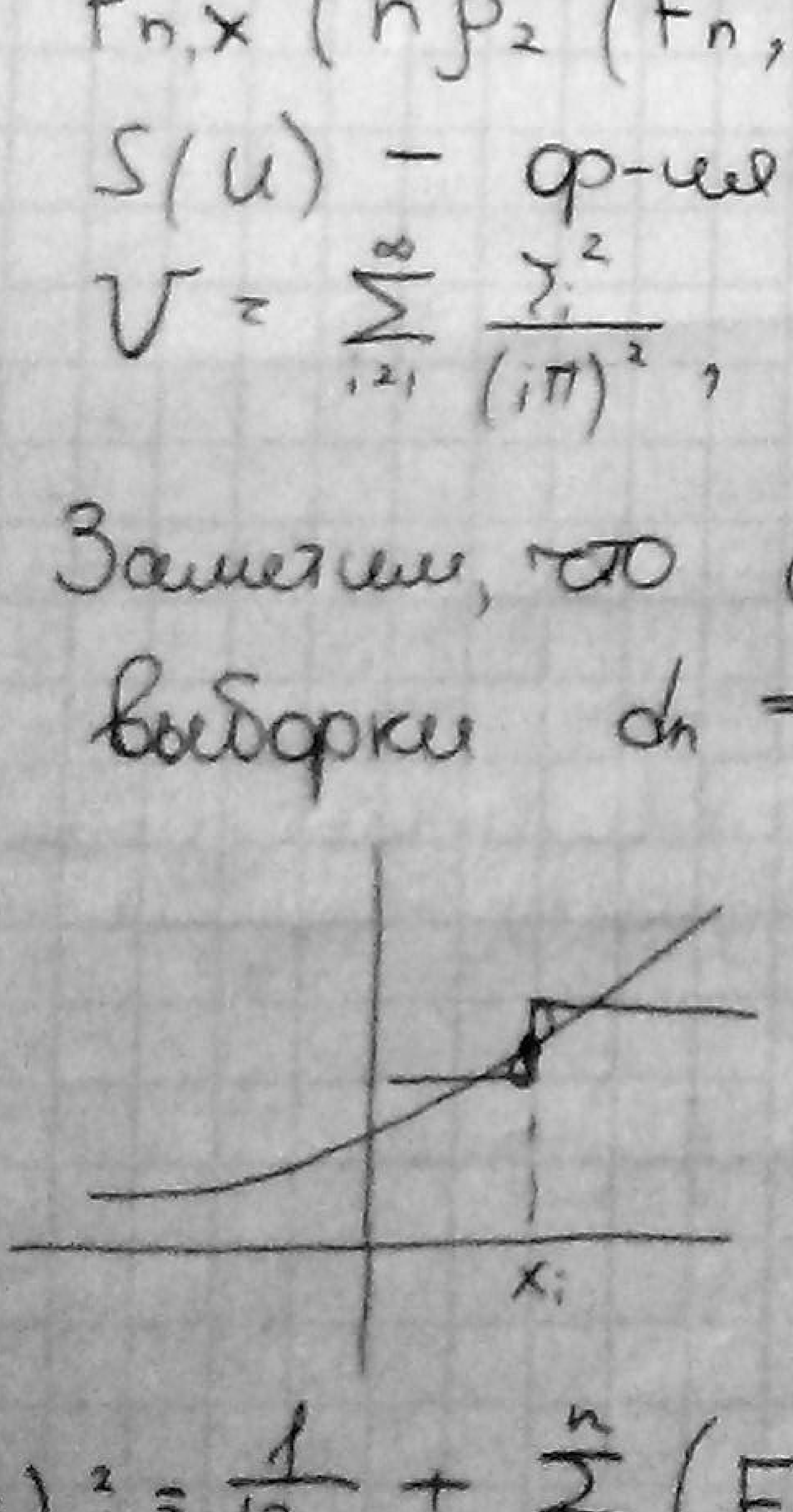
$$P(d_n < t) = F_{d_n}(t) \xrightarrow[n\to\infty]{} K(t), \quad K(t) = \begin{cases} 0, & t \leq 0 \\ \sum\limits_{j=-\infty}^{+\infty} (-1)^j e^{-2(jt)^2}, & t > 0 \end{cases}$$

$$P_{n,x}\big(\overbrace{n\rho_2^2(F_n, F_x)}^{=\omega_n^2} < t\big) = F_{n\rho_2^2(F_n, F_x)}(t) \xrightarrow[n\to\infty]{} S(u)$$

$S(u)$ – ф-ия распред. СВ U

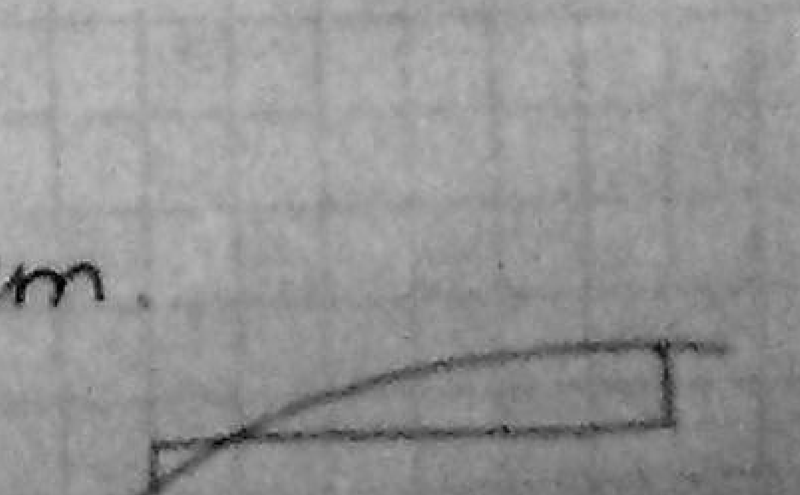
$U = \sum\limits_{i=1}^{\infty} \frac{\xi_i^2}{(i\pi)^2}$, $\xi_i \sim N(0,1)$

Заметим, что статистика $d_n = \sqrt{n}\,\rho(F_n, F_x)$ для $\forall$ выборки $d_n = \max\limits_{i=1,\ldots,n}\left[\max\left(\left|F_x(x_i) - \frac{i-1}{n}\right|, \left|F_x(x_i) - \frac{i}{n}\right|\right)\right]$



$F_x(t)$ – непр. и не убывает

$F_n(t)$ – не убыв., кус. пост.



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

Используя т. К. постр. довер. полосу

$p\left(F_x(t) \in \mathcal{D}_n\right) \xrightarrow[n\to\infty]{} \gamma$

$F^-_{n,\gamma}(t) \overset{\max}{=} \left(F_n(t) - \frac{u_\gamma}{\sqrt{n}}, 0\right)$ $\qquad u_\gamma$ – γ-кв. ф К

$\qquad K(u_\gamma) = \gamma$

$F^+_{n,\gamma}(t) = \min\left(F_n(t) + \frac{u_\gamma}{\sqrt{n}}, 1\right)$

$P_x\left(|F_x(t) - F_n(t)| \overset{\forall t}{<} t\right) =$

$= P_x\left(\sup_t \underbrace{|F_x(t) - F_n(t)|}_{=\rho_\infty(F_x, F_n)} < T\right) \to \cancel{K(t)} \; K(T)$

т.е. $K(T) = \gamma \Rightarrow T = u_\gamma$ – кв. $(K(u))$

$P(0 \leqslant F_x(t) \leqslant 1) = 1$

критерий пр-ки гипотезу Колмогорова

(критерий согласия)

$H_0: F_x(t) \equiv F_0(t)$, где $F_0(t)$ – фикс. ф. распред.

$H_1: F_x(t) \neq F_0(t)$

$F_0(t)$ и по предпол. $F_x(t)$ – непр. ф-ии

$$\varphi_n(X_1, \ldots, X_n) = \begin{cases} 0\,(H_0), & \sqrt{n}\,\rho_\infty(F_n(t), F_0(t)) < \delta \\ 1\,(H_1), & \sqrt{n}\,\rho_\infty(F_n, F_0) \geqslant \delta \end{cases}$$

δ выбираем так, чтобы $\overline{\alpha(\varphi_n)} \leqslant \alpha$

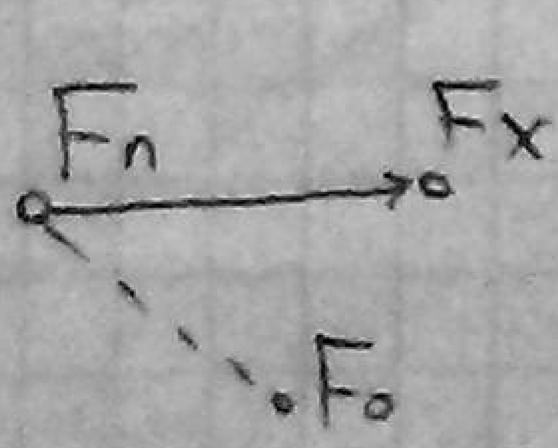
$\alpha(\varphi_n) = P_{F_x \equiv F_0}\left(\sqrt{n}\,\rho_\infty(F_n, F_x) \geqslant \delta\right) =$

$= 1 - p(\mathcal{D}_n < \delta) \to 1 - K(\delta) = \alpha \Rightarrow \delta = \cancel{1-\alpha}\; u_{1-\alpha}$

$\lim\limits_{n\to\infty} \alpha(\varphi_n) = \alpha$, если $\delta = u_{1-\alpha}$ – $(1-\alpha)$ – квантиль функ распр. Колмог.

Крит. Колмог. экв. сост. оценкой уровня

$\beta(\varphi_n, F_x = F_1 \neq F_0) = P_{F_x \neq F_0}\left(\sqrt{n}\,\rho_\infty(F_n, F_x) < u_{1-\alpha}\right)$



$F_n \xrightarrow[n\to\infty]{} F_x$, $\rho_\infty(F_x, F_0) = a > 0$

$\exists N: n > N \quad \rho_\infty(F_n, F_0) > \frac{a}{2} \Rightarrow$

$\Rightarrow \sqrt{n}\,\rho_\infty(F_n, F_0) \xrightarrow[n\to\infty]{} \infty \Rightarrow \beta(\varphi_n, F_1) \xrightarrow[n\to\infty]{} 0$