Základy matematickej štatistiky

1. Náhodný výber, výberové momenty a odhad parametrov

Aleš Kozubík

Katedra Matematických metód
Fakulta Riadenia a Informatiky
Žilinská Univerzita v Žiline

6. mája 2015

Náhodný výber

V realite často potrebujeme analyzovať náhodné premenné, pričom ich rozdelenie nie je úplne známe.

Často napr. predpokladáme určitý typ rozdelenia, ale nepoznáme hodnoty parametrov.

Jediným spôsobom ako tieto informácie doplniť je vychádzať z výsledkov opakovania pokusov pri rovnakých podmienkach. Hľadané parametre sú potom funkciami týchto výsledkov.

Pre výsledky často používame termín *súbor hodnôt*. Na rozdiel od množiny hodnôt, v súbore sa môžu rovnaké hodnoty opakovať.

Náhodný výber

Teraz sformalizujeme proces získania súboru hodnôt, a to v podobe náhodného výberu.

Definícia (Náhodný výber)

Postupnosť nezávislých, rovnako rozdelených náhodných premenných $X_1, X_2, \ldots, X_n$ sa nazýva *náhodný výber*.
Číslo n nazývame *rozsah náhodného výberu*.

Niekedy sa pre pohodlnejšie vyjadrovanie hovorí o náhodnom výbere z určitého rozdelenia a náhodné premenné $X_1, X_2, \ldots, X_n$ sa potom nazývajú prvkami tohto náhodného výberu.

Štatistiky

Už sme spomenuli, že náhodný výber využívame na určenie neznámych parametrov rozdelenia. Tieto parametre sú obvykle nejakou funkciou tohto náhodného výberu. Preto zavádzame nasledujúci pojem.

Definícia (Štatistika)

Funkciu jednej alebo viac náhodných premenných ktorá nezávisí na hodnotách neznámych parametrov sa nazýva *štatistika*.

Ak $X_1, X_2, \ldots, X_n$ sú náhodné premenné, tak náhodná premenná $Y = \sum_{i=1}^{n} X_i$ je štatistikou v zmysle predchádzajúcej definície. Naproti tomu náhodná premenná $Y = \frac{X_1 - \mu}{\sigma}$ je štatistikou iba v prípade, že hodnoty μ a σ sú známe.

Usporiadaný náhodný výber

Niekedy je nutné namerané hodnoty zoradiť podľa veľkosti. Usporiadajme teda náhodný výber $X_1, \ldots, X_n$ podľa veľkosti do novej postupnosti

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}.$$

Definícia (usporiadaný náhodný výber)

Postupnosť $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ nazývame *usporiadaný náhodný výber*.

Usporiadaný náhodný výber

Veta (Rozdelenie usporiadaného náhodného výberu)

Nech $r \in \{1, 2, \ldots, n\}$. Potom distribučná funkcia $G_r(x)$ náhodnej premennej $X_{(r)}$ je daná vzťahom

$$G_r(x) = \sum_{i=r}^{n} \binom{n}{i} F^i(x)[1 - F(x)]^{n-i},$$

kde $F(x)$ je distribučná funkcia náhodných premenných $X_1, \ldots, X_n$.

Usporiadaný náhodný výber

Dôkaz

Pravdepodobnosť, že medzi hodnotami $X_1, \ldots, X_n$ bude práve i takých, že ich hodnota je menšia než dané x je daná Bernoulliho vzorcom ako

$$\binom{n}{i} F^i(x)[1 - F(x)]^{n-i},$$

nakoľko počet takýchto hodnôt sa riadi binomickým rozdelením. Hodnota $X_{(r)}$ bude menšia než dané x vtedy, ak medzi $X_1, \ldots, X_n$ bude buďto r alebo $r + 1$ atď. alebo n takých, ktoré sú menšie než x. Tieto prípady predstavujú nezlučiteľné udalosti, preto výsledná pravdepodobnosť je súčtom ich pravdepodobností.

Usporiadaný náhodný výber

Poznámka

Ako špeciálne prípad z vety vyplývajú vzorce pre distribučné funkcie najmenšieho a najväčšieho člena usporiadaného náhodného výberu. Dostávame tak

$$G_1(x) = 1 - [1 - F(x)]^n,$$

resp.

$$G_n(x) = F^n(x).$$

Všeobecný výberový moment

Definícia (Všeobecný výberový moment)

Nech $X_1, \ldots, X_n$ je náhodný výber zo základného súboru X. Pod *všeobecným výberovým momentom k-teho rádu v bode a* rozumieme hodnotu

$$M_k(a) = \frac{1}{n} \sum_{i=1}^{n} (x_i - a)^k.$$

Začiatočný výberový moment

Pre $a = 0$ získame definíciu začiatočných výberových momentov.

Definícia (Začiatočný výberový moment)

Nech $X_1, \ldots, X_n$ je náhodný výber zo základného súboru X. Pod *začiatočným výberovým momentom k-teho rádu* rozumieme hodnotu

$$v_k = \frac{1}{n} \sum_{i=1}^{n} x_i^k.$$

Aritmetický priemer

Pre hodnotu $k = 1$ dostávame začiatočný moment prvého rádu, ktorý nazývame *aritmetický priemer* a píšeme

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

Centrálny výberový moment

Pre $a = \overline{x}$ získame definíciu centrálnych výberových momentov.

Definícia (Centrálny výberový moment)

Nech $X_1, \ldots, X_n$ je náhodný výber zo základného súboru X. Pod *centrálnym výberovým momentom k-teho rádu* rozumieme hodnotu

$$m_k = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^k.$$

Výberový rozptyl

Pre hodnotu $k = 2$ dostávame centrálny moment druhého rádu, ktorý nazývame *výberový rozptyl* a píšeme

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

Charakteristiky polohy

Medzi charakteristiky polohy zaraďujeme

Aritmetický priemer

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

Charakteristiky polohy

Medzi charakteristiky polohy zaraďujeme

Výberový medián

$X_{(1)}, \ldots, X_{(n)}$ je usporiadaný náhodný výber. *Výberový medián* definujeme ako

$$\tilde{x} = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & n \text{ nepárne} \\ \dfrac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n+2}{2}\right)}}{2} & n \text{ párne} \end{cases}$$

Charakteristiky polohy

Medzi charakteristiky polohy zaraďujeme

Výberový módus

Nech $X_1, \ldots, X_n$ je náhodný výber. *Výberový módus* definujeme ako hodnotu s najväčšou početnosťou a označujeme ho $\hat{x}$.

Charakteristiky polohy

Medzi charakteristiky polohy zaraďujeme

Geometrický priemer

Nech $X_1, \ldots, X_n$ je náhodný výber. Ak sú všetky hodnoty $x_1, \ldots, x_n$ kladné, tak definujeme *geometrický priemer* ako hodnotu:

$$\overline{x_G} = \sqrt[n]{x_1 x_2 \cdots x_n}.$$

Charakteristiky polohy

Medzi charakteristiky polohy zaraďujeme

Harmonický priemer

Nech $X_1, \ldots, X_n$ je náhodný výber. Ak sú všetky hodnoty $x_1, \ldots, x_n$ kladné, tak definujeme *harmonický priemer* ako hodnotu:

$$\overline{x_H} = \frac{n}{x_1^{-1} + x_2^{-1} + \cdots + x_n^{-1}},$$

Charakteristiky rozptýlenia

Medzi charakteristiky rozptýlenia zaraďujeme

Výberová disperzia

Nech $X_1, \ldots, X_n$ je náhodný výber. Výberovú disperziu (rozptyl) definujeme ako

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

a výberová smerodajná odchýlka $s = \sqrt{s^2}$.

Charakteristiky rozptýlenia

Medzi charakteristiky rozptýlenia zaraďujeme

Výberová absolútna odchýlka

nech $X_1, \ldots, X_n$ je náhodný výber. *Výberovú absolútnu odchýlku* definujeme ako

$$A = \frac{1}{n} \sum_{i=1}^{n} |x_i - \overline{x}|.$$

Charakteristiky rozptýlenia

Medzi charakteristiky rozptýlenia zaraďujeme

Výberový variačný koeficient

Nech $X_1, \ldots, X_n$ je náhodný výber. *Výberový variačný koeficient* definujeme ako podiel

$$V = \frac{s}{\overline{x}}.$$

Hodnotu

$$R = \max_i X_i - \min_i X_i,$$

nazývame *výberové variačné rozpätie*.

Dvojrozmerný náhodný výber

Definícia (Výberová kovariancia)

Nech $(X_{11}, X_{21}) \ldots, (X_{1n}, X_{2n})$ je náhodný výber z dvojrozmerného rozdelenia. *Výberovú kovarianciu* definujeme ako

$$s_{12} = \frac{1}{n} \sum_{i=1}^{n} (x_{1i} - \overline{x_1})(x_{2i} - \overline{x_2}).$$

a *výberový korelačný koeficient* ako

$$r_{12} = \frac{s_{12}}{s_1 s_2}.$$

Empirická distribučná funkcia

Definícia (Empirická distribučná funkcia)

Formálne je možné *empirickú distribučnú funkciu* definovať ako

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \chi_{(-\infty;x)}.$$

kde

$$\chi_A(x) = \begin{cases} 1 & \text{pre } x \in A \\ 0 & \text{pre } x \notin A \end{cases}$$

je charakteristická funkcia množiny A.

Vlastnosti výberových charakteristík

Veta

Nech $X_1, \ldots, X_n$ je náhodný výber z rozdelenia, ktoré má strednú hodnotu μ a konečný rozptyl σ^2. Potom platí

a) $\mathbb{E}(\overline{x}) = \mu$,

b) $\mathbb{D}(\overline{x}) = \frac{\sigma^2}{n}$,

c) $\mathbb{E}(s^2) = \frac{n}{n-1}\sigma^2$, pre $n \geq 2$.

Vlastnosti výberových charakteristík

Dôkaz

a) Postupne dostávame

$$\mathbb{E}(\overline{x}) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i) = \frac{1}{n}n\mu = \mu.$$

Vlastnosti výberových charakteristík

Dôkaz

b) Postupne dostávame

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

Vlastnosti výberových charakteristík

Dôkaz

c) Využijeme identitu

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

Z nej dostávame

$$\begin{aligned}\mathbb{E}\left(\sum_{i=1}^{n}(X_i - \overline{X})^2\right) &= \mathbb{E}\left(\sum_{i=1}^{n}(X_i - \mu)^2 - n(\overline{X} - \mu)^2\right) \\ &= \sum_{i=1}^{n}\mathbb{D}\left(\overline{X_i}\right) - n\mathbb{D}\left(\overline{X}\right) \\ &= n\sigma^2 - \frac{n\sigma^2}{n} = (n-1)\sigma^2,\end{aligned}$$

čo je ekvivalentné s tvrdením c)

Odhady parametrov
Členenie

Predpokladajme, že máme k dispozícii náhodný výber $X_1, \ldots, X_n$ z rozdelenia s hustotou $f(x, \boldsymbol{\theta})$, pričom $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$ sú neznáme parametre.

Na základe tohto výberu je potrebné určiť čo najpresnejší odhad parametra $\boldsymbol{\theta}$, o ktorom je známe len to, že patrí do nejakého parametrického priestoru $\mathcal{P}$.

Rozoznávame *bodové* a *intervalové* odhady.

Odhady parametrov
Bodový odhad

Definícia (Bodový odhad)

Nech náhodný výber $\mathbf{X} = (X_1, \ldots, X_n)$ závisí od neznámeho parametra θ. *Odhadom* parametra rozumieme štatistiku $g(\mathbf{X})$, ktorá každej množine pozorovaných hodnôt náhodného výberu priradí hodnotu odhadu parametra, ktorú označíme $\hat{\theta}$.

Terminologická poznámka

Odhad teda predstavuje nejakú funkciu definovanú na $\mathbb{R}^n$,kde n je rozsah náhodného výberu, kým veličina $\hat{\theta}$ už je konkrétnou odhadnutou hodnotou parametra θ. V anglickej literatúre sa tieto dva pojmy rozlišujú ako *estimator*, čo je funkcia a *estimate*, ktorý predstavuje konkrétnu hodnotu odhadu.

Odhady parametrov
Intervalový odhad

Definícia (Intervalový odhad)

Intervalovým odhadom parametra θ rozumieme nejakú vhodnú množinu, ktorá s dostatočne veľkou pravdepodobnosťou pokrýva hodnotu parametra θ. Niekedy sa používa tiež termín *interval spoľahlivosti* pre parameter θ a zmienená pravdepodobnosť, s ktorou hodnoty parametra θ patria do tejto množiny sa označuje ako *spoľahlivosť* tohto odhadu.

Odhady parametrov
Vlastnosti odhadov

Jednou z častých požiadaviek, ktoré kladieme na odhady je ich nevychýlenosť v zmysle nasledujúcej definície.

Definícia (Nevychýlenosť odhadu)

Štatistiku $T = g(\mathbf{X})$ nazývame *nevychýleným* odhadom parametra $\boldsymbol{\theta}$, ak platí

$$\mathbb{E}(T) = \boldsymbol{\theta} \quad \forall \boldsymbol{\theta} \in \mathcal{P}. \tag{1}$$

Poznámka

Nevychýlenosť je síce dôležitou vlastnosťou odhadov, ale nie vždy je možné na tejto požiadavke trvať. Môže sa dokonca stať, že nevychýlený odhad parametra ani nemusí existovať.

Vlastnosti odhadov
Príklady

Už sme si ukázali, že pre aritmetický priemer platí $\mathbb{E}(\overline{x}) = \mu$, teda aritmetický priemer je nevychýleným odhadom strednej hodnoty základného súboru.

Pre výberový rozptyl naopak platilo $\mathbb{E}(s^2) = \frac{n}{n-1}\sigma^2$, teda výberový rozptyl nie je nevychýleným odhadom rozptylu základného súboru.

V ďalšom sa teda pokúsime skonštruovať nevychýlený odhad rozptylu.

Vlastnosti odhadov
Príklady

Už sme odvodili vzťah

$$\mathbb{E}\left(\sum_{i=1}^{n}(X_i - \overline{X})^2\right) = (n-1)\sigma^2,$$

a teda

$$\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2\right) = \frac{n-1}{n}\sigma^2.$$

Aby sme na pravej strane rovnice dostali hodnotu σ^2, je potrebné je vynásobiť zlomkom $\frac{n}{n-1}$. Pre nevychýlený odhad rozptylu tak dostávame

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

Vlastnosti odhadov

Poznámka k príkladu

Ak by sme porovnali strednú kvadratickú odchýlku nevychýleného odhadu rozptylu a výberového rozptylu od skutočnej hodnoty parametra σ^2, zistili by sme, že paltí

$$\mathbb{E}\left((s^2 - \sigma^2)^2\right) < \mathbb{E}\left((s_{n-1}^2 - \sigma^2)^2\right).$$

To je dôvodom, prečo sa možno stretnúť s tým, že niektorí štatistici uprednostňujú výberový rozptyl pred nevychýleným odhadom rozptylu. Zároveň tento príklad ukazuje, že nevychýlený odhad nemusí byť najlepší, pokiaľ sa týka strednej kvadratickej odchýlky od skutočnej hodnoty.

Vlastnosti odhadov

Definícia (Konzistentný odhad)

Nech θ je jednorozmerný parameter a $X_1, \ldots, X_n$ je náhodný výber z rozdelenia, ktoré je závislé od tohto parametra. Nech pre každé n je definovaný odhad $T_n = g(X_1, \ldots, X_n)$. Povieme, že odhad T_n je *konzistentný* ak pre $n \to \infty$ platí $T_n \xrightarrow{\mathbb{P}} \theta$.

Vlastnosti odhadov

Veta

Nech $\mathbb{E}\left(T_n^2\right) < \infty$ pre každé prirodzené n. Ak sú splnené predpoklady

1. $\mathbb{E}\left(T_n\right) \to \theta$,
2. $\mathbb{D}\left(T_n\right) \to 0$,

tak T_n je konzistentným odhadom parametra θ.

Dôkaz

Pre každé $\epsilon > 0$ platí

$$\mathbb{P}\left(|T_n - \theta| < \epsilon\right) \geq \mathbb{P}\left(|T_n - \mathbb{E}\left(T_n\right)| < \frac{\epsilon}{2}, |\mathbb{E}\left(T_n\right) - \theta| < \frac{\epsilon}{2}\right).$$

Podľa Čebyševovej nerovnosti je

$$\mathbb{P}\left(|T_n - \mathbb{E}\left(T_n\right)| < \frac{\epsilon}{2}\right) \geq 1 - \frac{\mathbb{D}\left(T_n\right)}{\left(\frac{\epsilon}{2}\right)^2}.$$

Predpoklad 1 zaručuje, že existuje číslo n_0 také, že pre každé $n \geq n_0$ platí $|\mathbb{E}\left(T_n\right) - \theta| < \frac{\epsilon}{2}$.

Dôkaz

Na pravej strane máme pravdepodobnosť prieniku dvoch udalostí.

Pravdepodobnosť prvej z nich konverguje k jednej, druhá z nich je pre $n \geq n_0$ istá udalosť.

Z toho vyplýva

$$\mathbb{P}\left(|T_n - \theta| < \epsilon\right) \to 1.$$

Momentová metóda

Nech $X_1, \ldots, X_n$ je náhodný výber z rozdelenia, ktoré závisí od parametra $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$. Predpokladajme, že pre každé $\boldsymbol{\theta} \in \mathcal{P}$ existujú počiatočné momenty $\nu_k = \mathbb{E}\left(X_i^k\right)$, pre $k = 1, \ldots, m$.

Výberové počiatočné momenty sú dané vzťahom $v_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ pre $k = 1, \ldots, m$.

Momentová metóda odhadu parametra $\boldsymbol{\theta}$ spočíva v tom, že za odhad berieme riešenie rovníc

$$\nu_k = v_k, \quad k = 1, \ldots, m.$$

Metóda maximálnej vierohodnosti

Nech $X_1, \ldots, X_n$ je náhodný výber z rozdelenia, ktoré závisí od parametra $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$. Ak má vektor $\mathbf{X}$ spojité rozdelenie, tak jeho hustotu označme symbolom $L(\mathbf{x}, \boldsymbol{\theta})$. Ak má $\mathbf{X}$ diskrétne rozdelenie, tak $L(\mathbf{x}, \boldsymbol{\theta}) = \mathbb{P}(\mathbf{X} = \mathbf{x})$.

Funkciu $L(\mathbf{x}, .)$ premennej $\boldsymbol{\theta}$ nazývame *vierohodnostná funkcia*[1].

Hovoríme, že sme určili odhad parametra $\boldsymbol{\theta}$ *metódou maximálnej vierohodnosti*, ak za odhad berieme takú hodnotu $\hat{\boldsymbol{\theta}}$, že pre každé $\boldsymbol{\theta} \in \mathcal{P}$ platí $L(\mathbf{x}, \hat{\boldsymbol{\theta}}) \geq L(\mathbf{x}, \boldsymbol{\theta})$.

[1] Preto označenie $L(\mathbf{x}, \boldsymbol{\theta})$ z anglického *likelyhood*

Metóda maximálnej vierohodnosti

Poznámka k výpočtom

Vierohodnostná funkcia je, v dôsledku nezávislosti prvkov náhodného výberu, súčinom hustôt príp. pravdepodobnostných funkcií. Určenie jej lokálneho maxima si preto vyžaduje pomerne náročné derivovanie rozsiahleho súčinu.

Pre uľahčenie výpočtu sa preto obvykle prechádza od hľadania extrémov samotnej vierohodnostnej funkcie ku hľadaniu extrémov logaritmu vierohodnostnej funkcie $\ln L(\mathbf{x}, \hat{\boldsymbol{\theta}})$.[2]

Prejdeme tak od derivovania súčinu ku derivovania súčtu, čo je z výpočtového hľadiska podstatne pohodlnejšie.

[2] Rozmyslite si, prečo sa logaritmovaním extrémy funkcie nezmenia.

Intervalové odhady

Definícia (Interval spoľahlivosti)

Nech $\mathbf{X}$ je náhodný výber, ktorého rozdelenie závisí od neznámeho parametra θ a nech sú dané dve štatistiky $A = g_1(\mathbf{X})$ a $B = g_2(\mathbf{X})$ také, že $A < B$. Ak a a b sú nejaké realizácie A a B a $\mathbb{P}(a < \theta < b) = \alpha$. Interval $(a; b)$ nazývame 100α *percentným intervalom spoľahlivosti* pre θ a hodnotu α nazývame *hladinou spoľahlivosti*. Ak je $A = -\infty$ alebo $B = \infty$, tak hovoríme o *jednostranných intervaloch spoľahlivosti*.

Príklad

Interval spoľahlivosti pre strednú hodnotu normálneho rozdelenia

Nech $X_i \sim \mathrm{N}\left(\mu, \sigma^2\right)$ je náhodný výber z normálneho rozdelenia so známym rozptylom σ^2. Potom $\overline{X_n} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$ a preto

$$\mathbb{P}\left(\left|\frac{\overline{X_n} - \mu}{\frac{\sigma}{\sqrt{n}}}\right| < c\right) = 2\Phi(c) - 1, \quad \forall c > 0,$$

čo je ekvivalentné s

$$\mathbb{P}\left(\overline{X_n} - c\frac{\sigma}{\sqrt{n}} < \mu < \overline{X_n} + c\frac{\sigma}{\sqrt{n}}\right) = 2\Phi(c) - 1.$$

Príklad

Interval spoľahlivosti pre strednú hodnotu normálneho rozdelenia

$\overline{X_n}$ je štatistika určená náhodným výberom.

Ak použijeme konkrétne realizácie, dostaneme hodnotu aritmetického priemeru $\overline{x_n}$ a interval spoľahlivosti v tvare

$$\left(\overline{x_n} - c\frac{\sigma}{\sqrt{n}}; \overline{x_n} + c\frac{\sigma}{\sqrt{n}}\right).$$

Toto je interval spoľahlivosti pre strednú hodnotu normálneho rozdelenia μ s hladinou spoľahlivosti $\alpha = 2\Phi(c) - 1$.

Príklad

Interval spoľahlivosti pre strednú hodnotu normálneho rozdelenia

Ak si uvedomíme, že platí

$$c = \Phi^{-1}\left(\frac{\alpha+1}{2}\right),$$

tak pre príslušný interval spoľahlivosti na hladine 100α percent môžeme písať

$$\left(\overline{x_n} - \Phi^{-1}\left(\frac{\alpha+1}{2}\right)\frac{\sigma}{\sqrt{n}}; \overline{x_n} + \Phi^{-1}\left(\frac{\alpha+1}{2}\right)\frac{\sigma}{\sqrt{n}}\right).$$

Ak konkrétne zvolíme spoľahlivosť 95%, tak pre $\alpha = 0,95$ určíme $\Phi^{-1}(0,975) \approx 1,96$ a interval spoľahlivosti dostávame v tvare

$$\left(\overline{x_n} - 1,96\left(\frac{\sigma}{\sqrt{n}}\right); \overline{x_n} + 1,96\left(\frac{\sigma}{\sqrt{n}}\right)\right).$$