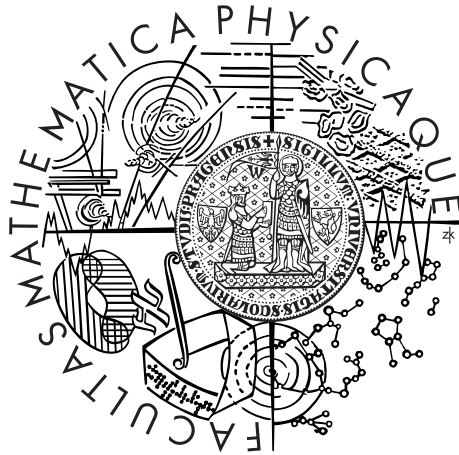


Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE



Marko Borovský

Intervalové odhady binomického podílu

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. RNDr. Karel Zvára, CSc.

Studijní program: Matematika

Studijní obor: Obecná matematika

Praha 2013

Na tomto mieste by som rád poďakoval vedúcemu práce docentovi Karlovi Zvárovi za odborný dozor a za venovaný čas. Taktiež ďakujem rodičom za podporu počas celej doby štúdia.

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona.

V Praze dne 24.05.2013

Podpis autora

Název práce: Intervalové odhady binomického podílu

Autor: Marko Borovský

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. RNDr. Karel Zvára, CSc.

Abstrakt: Předmětem této práce jsou bodový odhad a intervalové odhady binomického podílu. Konstrukce intervalových odhadů pro pravděpodobnost úspěchu v binomickém rozdělení je jedním z nejzákladnějších a klíčových problémů v statistické praxi. Práce je rozdělena do tří kapitol. První kapitola se zabývá maximálně věrohodným odhadem binomického podílu. Dále popíšeme několik způsobů konstrukce intervalů spolehlivosti. Nakonec porovnáme všechny intervaly z hlediska skutečné pravděpodobnosti pokrytí a střední hodnoty šířky.

Klíčová slova: binomické rozdělení, intervaly spolehlivosti, pravděpodobnost pokrytí

Title: Interval estimates for binomial proportion

Author: Marko Borovský

Department: Department of Probability and Mathematical Statistics

Supervisor: doc. RNDr. Karel Zvára, CSc.

Abstract: The subject of this thesis is the point estimate and interval estimates of the binomial proportion. Interval estimation of the probability of success in a binomial distribution is one of the most basic and crucial problems in statistical practice. The thesis is divided into three chapters. The first chapter is about maximum-likelihood estimation for a binomial proportion. Furthermore, we will describe several methods of the construction of confidence intervals. In the end, we will compare all intervals in term of the actual coverage probability and expected length.

Keywords: binomial distribution, confidence intervals, coverage probability

Názov práce: Intervalové odhady binomického podielu

Autor: Marko Borovský

Katedra: Katedra pravdepodobnosti a matematické statistiky

Vedúci bakalárskej práce: doc. RNDr. Karel Zvára, CSc.

Abstrakt: Predmetom tejto práce sú bodový odhad a intervalové odhady binomického podielu. Konštrukcia intervalových odhadov pre pravdepodobnosť úspechu v binomickom rozdelení je jedným z najzákladnejších a kľúčových problémov v štatistickej praxi. Práca je rozdelená do troch kapitol. Prvá kapitola sa zaoberá maximálne vierohodným odhadom binomického podielu. Ďalej popíšeme niekoľko spôsobov konštrukcie intervalov spoľahlivosti. Nakoniec porovnáme všetky intervaly z hľadiska skutočnej pravdepodobnosti pokrytia a strednej hodnoty šírky.

Kľúčové slová: binomické rozdelenie, intervaly spoľahlivosti, pravdepodobnosť pokrytia

Obsah

Úvod	1
1 Bodový odhad binomického podielu	2
1.1 Binomické rozdelenie	2
1.2 Bodový odhad parametra p	2
2 Intervaly spoľahlivosti	5
2.1 Waldov interval	5
2.2 Wilsonov interval	6
2.3 Agresti-Coullov interval	8
2.4 Arkussínusový interval	9
2.5 Clopper-Pearsonov interval	10
2.6 Interval z vierohodnostného pomeru	12
3 Porovnanie intervalov	14
3.1 Skutočná pravdepodobnosť pokrytia	14
3.2 Stredná hodnota šírky intervalu	19
3.3 Optimálna pravdepodobnosť pokrytia	21
Záver	23
A Použité vety	24
B Kód v programe R	26
Zoznam použitej literatúry	32
Zoznam obrázkov	33
Zoznam tabuliek	34

Úvod

Intervalové odhady binomického podielu sa používajú na odhadnutie pravdepodobnosti úspechu z výsledku n nezávislých pokusov. V štatistických textoch sa vyskytuje mnoho metód konštrukcie intervalov spoľahlivosti pre parameter binomického rozdelenia.

V tejto práci najskôr priblížime binomické rozdelenie a skonštruujeme bodový odhad parametra p , ktorý je „skoro“ maximálne vierohodným odhadom.

V ďalšej kapitole teoreticky popíšeme vybrané intervaly spoľahlivosti. Medzi inými sú to intervaly odvodené z troch testov založených na vierohodnostnej funkcii: Waldov interval, Wilsonov interval a interval z vierohodnostného pomeru. Ďalej si priblížime transformáciu stabilizujúcu rozptyl a z nej odvodený arkus-sínusový interval. Agresti a Coull navrhli v článku [1] interval, ktorý je zjednodušením Wilsonovho intervalu, a preto je podľa nich aj pomenovaný. Posledným zo skúmaných intervalov je Clopper-Pearsonov interval, známy aj pod názvom exaktný interval, pretože vychádza z exaktného testu pre parameter binomického rozdelenia.

V poslednej kapitole porovnáme jednotlivé metódy z hľadiska skutočnej pravdepodobnosti pokrytia a strednej hodnoty šírky intervalov. K tomu si napíšeme funkcie v štatistickom programe R, ktorý je pre daný účel viac než vhodný.

1. Bodový odhad binomického podielu

1.1 Binomické rozdelenie

Definícia. Nech $n \in \mathbb{N}$ a $p \in (0, 1)$. Ďalej nech náhodná veličina X nadobúda iba celočíselných hodnôt od 0 do n , a to s pravdepodobnosťami

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n.$$

Potom hovoríme, že X má *binomické rozdelenie* a píšeme $X \sim \text{Bi}(n, p)$.

Poznámka. V prípade $n = 1$ je zrejme $P(X = 1) = p$ a $P(X = 0) = (1 - p)$. Tento špeciálny prípad nazývame *alternatívne rozdelenie* a označujeme ho $\text{Alt}(p)$. Nech X_1, \dots, X_n sú nezávislé náhodné veličiny s alternatívnym rozdelením s parametrom p alebo inak povedané, je to náhodný výber z alternatívneho rozdelenia. Potom pre $X = \sum_{i=1}^n X_i$ platí $X \sim \text{Bi}(n, p)$. Náhodná veličina X teda predstavuje počet úspechov v n nezávislých pokusoch, pričom v každom pokuse úspech nastane s pravdepodobnosťou p .

Veta 1.1.1. Nech $X \sim \text{Bi}(n, p)$. Potom $E X = np$ a $\text{var } X = npq$, kde $q = (1 - p)$.

Dôkaz. Ak $X_1 \sim \text{Alt}(p)$, potom zrejme platí $E X_1 = p$ a $\text{var } X_1 = pq$. S využitím predchádzajúcej poznámky dostávame

$$\begin{aligned} E X &= E \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n E X_i = n E X_1 = np, \\ \text{var } X &= \text{var} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \text{var } X_i = n \text{var } X_1 = npq. \end{aligned}$$

□

1.2 Bodový odhad parametra p

Často môže nastať situácia, že $X \sim \text{Bi}(n, p)$, ale nepoznáme parameter p , ktorý nás zaujíma. Preto chceme odhadnúť p pomocou náhodnej veličiny X . Predpokladajme, že $X = x$, $X \neq 0$, $X \neq n$ a vytvoríme odhad \hat{p} metódou *maximálnej vierohodnosti*.

Odhad \hat{p} musí vyhovovať podmienke $P_{\hat{p}}(X = x) = \max_{p \in (0,1)} P_p(X = x)$. To znamená, že \hat{p} maximalizuje $f_p(x) = \binom{n}{x} p^x (1-p)^{n-x}$ ako funkciu premennej p , ktorá sa nazýva aj *vierohodnostná funkcia*. Potom maximum tejto funkcie dostaneme ako riešenie *vierohodnostnej rovnice*

$$\frac{\partial}{\partial p} \log f_x(p) = 0.$$

Maximum môžeme nájsť aj s využitím nasledujúceho triku (vid'. Anděl [3], Príklad 7.95). Pretože $\binom{n}{x}$ nezávisí na p , tak \hat{p} maximalizuje aj výraz $p^x(1-p)^{n-x}$. Ďalej definujeme

$$z_i = \begin{cases} \frac{np}{x}, & i = 1, \dots, x, \\ \frac{n(1-p)}{n-x}, & i = x+1, \dots, n \end{cases}$$

a vypočítame aritmetický a geometrický priemer čísel z_i . Platí

$$\bar{z} = \frac{1}{n} \left(x \frac{np}{x} + (n-x) \frac{n(1-p)}{n-x} \right) = p + (1-p) = 1,$$

$$\bar{z}_G = \sqrt[n]{\left(\frac{np}{x}\right)^x \left(\frac{n(1-p)}{n-x}\right)^{n-x}}$$

a z nerovnosti $\bar{z}_G \leq \bar{z}$ dostávame

$$\begin{aligned} \left(\frac{np}{x}\right)^x \left(\frac{n(1-p)}{n-x}\right)^{n-x} &\leq 1, \\ \left(\frac{n}{x}\right)^x \left(\frac{n}{n-x}\right)^{n-x} p^x (1-p)^{n-x} &\leq 1, \\ p^x (1-p)^{n-x} &\leq \left(\frac{x}{n}\right)^x \left(1 - \frac{x}{n}\right)^{n-x}. \end{aligned}$$

Takže $\hat{p} = \frac{X}{n}$ je maximálne vierohodný odhad parametra p .

V prípade, že $X = 0$ alebo $X = n$, maximálne vierohodný odhad nenájde, pretože výrazy $(1-p)^n$ a p^n nenadobúdajú na intervale $(0, 1)$ svojho maxima. Každopádne odhad \hat{p} používame aj pre tieto prípady, len si treba uvedomiť, že tento odhad už nie je vždy maximálne vierohodným odhadom.

Definícia. Nech X_1, \dots, X_n je náhodný výber z rozdelenia, ktoré je prvkom rodiny $\{F_\theta; \theta \in \Theta\}$. Odhad $\hat{\theta}(X_1, \dots, X_n)$ parametra θ je *nestranný*, ak

$$\mathbf{E} \left[\hat{\theta}(X_1, \dots, X_n) \right] = \theta.$$

Veta 1.2.1. Odhad \hat{p} je nestranný odhad parametra p a platí $\text{var } \hat{p} = \frac{pq}{n}$.

Dôkaz.

$$\begin{aligned} \mathbf{E} \hat{p} &= \mathbf{E} \left(\frac{X}{n} \right) = \frac{\mathbf{E} X}{n} = \frac{np}{n} = p. \\ \text{var } \hat{p} &= \text{var} \frac{X}{n} = \frac{1}{n^2} \text{var } X = \frac{npq}{n^2} = \frac{pq}{n}. \end{aligned}$$

□

Ďalej ukážeme, že \hat{p} je najlepší nestranný odhad parametra p , teda neexistuje iný nestranný odhad, ktorý by mal menší rozptyl ako \hat{p} .

Definícia. Rodina rozdelení $\{F_\theta; \theta \in \Theta\}$ spĺňa *podmienky regularity*, ak $\Theta \subseteq \mathbb{R}$ je neprázdna otvorená a existujú nezáporné funkcie f_θ , $\theta \in \Theta$, a σ -konečná miera ν spĺňajúce

- (a) $F_\theta(x) = \int_{-\infty}^x f_\theta(t) d\nu(t), \quad \forall x \in \mathbb{R}, \forall \theta \in \Theta,$
- (b) $M = \{x; f_\theta(x) > 0\}$ nezávisí na $\theta,$
- (c) existuje $\frac{\partial}{\partial \theta} f_\theta(x) = f'_\theta(x), \quad \forall \theta \in \Theta, \forall x \in M - N,$ kde $\nu(N) = 0,$
- (d) $\forall \theta \in \Theta$ je $\frac{\partial}{\partial \theta} \int_M f_\theta(t) d\nu(t) = \int_M \frac{\partial}{\partial \theta} f_\theta(t) d\nu(t),$
- (e) $\forall \theta \in \Theta$ je integrál

$$I(\theta) = \int_M \left(\frac{f'_\theta(x)}{f_\theta(x)} \right)^2 f_\theta(x) d\nu(x) = \mathbb{E} \left(\frac{f'_\theta(X)}{f_\theta(X)} \right)^2$$

konečný a nenulový.

Poznámka. Veličina $I(\theta)$ sa nazýva *Fisherova informácia*.

Ak pre $\forall p \in (0, 1)$ položíme $f_p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \nu(x) = 1,$ kde $x \in M = \{0, 1, \dots, n\}$ a $f_p(x) = 0, \nu(x) = 0,$ pre $x \in \mathbb{R} - M.$ Potom $F_p(x) = \int_{-\infty}^x f_p(t) d\nu(t)$ je distribučná funkcia binomického rozdelenia. Splnenie predpokladov (b)–(d) je zrejmé. Ďalej platí

$$\frac{f'_p(x)}{f_p(x)} = \frac{xp^{x-1}(1-p)^{n-x} - (n-x)p^x(1-p)^{n-x-1}}{p^x(1-p)^{n-x}} = \frac{x}{p} - \frac{n-x}{1-p} = \frac{x-np}{p(1-p)}$$

a teda

$$I(p) = \mathbb{E} \left(\frac{X-np}{p(1-p)} \right)^2 = \frac{\text{var } X}{p^2(1-p)^2} = \frac{n}{p(1-p)}.$$

To znamená, že binomické rozdelenie spĺňa podmienky regularity a podľa vety A.1 platí

$$\text{var } \tilde{p} \geq \frac{1}{I(p)} = \frac{p(1-p)}{n},$$

pre každý nestranný odhad $\tilde{p}.$ Z toho vyplýva, že \hat{p} je najlepší nestranný odhad parametra $p.$

2. Intervaly spoľahlivosti

Motivácia. Bodový odhad \hat{p} nám dáva len jednu jedinú hodnotu z intervalu $(0, 1)$, a preto je zrejmé, že takýto odhad nám toho veľa nepovie o neznámom parametri. Preto by sme radi zostrojili na základe náhodnej veličiny X nejaký interval, o ktorom môžeme prehlásiť, že hodnota parametra p leží s určitou pravdepodobnosťou vo vnútri tohoto intervalu. Lepšie je ale povedať, že interval pokrýva parameter p , pretože p aj keď je neznáme, tak je pevne dané a náhodný je práve ten interval.

V praxi najprv zvolíme malú hodnotu $\alpha \in (0, 1)$, nazývanú *hladina významnosti* a snažíme sa zostaviť *interval spoľahlivosti* $I(X) = (d(X), h(X))$, pre ktorý platí

$$P_p [d(X) < p < h(X)] = 1 - \alpha, \quad p \in (0, 1).$$

Hodnota $1 - \alpha$ sa nazýva *koeficient spoľahlivosti*, prípadne hovoríme o *nominálnej úrovni spoľahlivosti* intervalu $I(X)$.

Kvôli diskkrétnej povahe binomického rozdelenia nie je možné zostaviť interval, ktorý by vždy pokryl parameter p s požadovanou pravdepodobnosťou, dokonca pre skoro všetky p nebude mať žiadny interval dané pokrytie. Preto chceme zostrojiť interval spoľahlivosti, ktorý pokrýva parameter p s pravdepodobnosťou aspoň približne rovnajúcou sa nominálnej úrovni spoľahlivosti.

V tejto kapitole si predstavíme niekoľko možných spôsobov zostrojenia intervalu spoľahlivosti pre parameter p binomického rozdelenia.

Poznámka. V praxi najčastejšie volíme $\alpha = 0,05$ a konštruujeme 95% interval spoľahlivosti parametra p .

Označenie. V celej kapitole bude u_α značiť $(1 - \alpha)$ kvantil rozdelenia $N(0, 1)$, nazýva sa taktiež *kritická hodnota* rozdelenia $N(0, 1)$.

2.1 Waldov interval

Definujme náhodnú veličinu

$$\xi := \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{X - np}{\sqrt{\frac{pq}{n}}} = \frac{X - np}{\sqrt{npq}}.$$

Vieme, že $X = X_1 + \dots + X_n$, kde X_i sú nezávislé rovnako rozdelené veličiny také, že $E X_i = p$ a $\text{var } X_i = pq$. A teda podľa Lindebergovej CLV (Veta A.2) aplikovanej na X_1, X_2, \dots dostávame

$$\frac{X - np}{\sqrt{n}} \xrightarrow{d} N(0, pq), \quad n \rightarrow \infty.$$

Ďalej z princípu spojitosti pre konvergenciu podľa distribučnej funkcie (Veta A.3) plynie

$$\xi \xrightarrow{d} N(0, 1), \quad n \rightarrow \infty. \quad (2.1)$$

Rozptyl pq/n odhadu \hat{p} nahradíme odhadom tohoto rozptylu \widehat{pq}/n , kde $\widehat{q} = 1 - \widehat{p}$. Definujme náhodnú veličinu

$$Z := \frac{\hat{p} - p}{\sqrt{\frac{\widehat{pq}}{n}}} = \frac{\xi}{\sqrt{\frac{\widehat{pq}}{pq}}}$$

Ďalej podľa vety A.5 zrejme platí $\hat{p} \xrightarrow{P} p$ a z princípu spojitosti pre konvergenciu podľa pravdepodobnosti (Veta A.4) dostávame

$$\sqrt{\frac{\widehat{pq}}{pq}} \xrightarrow{P} 1, \quad n \rightarrow \infty.$$

Nakoniec z vety A.6 plynie asymptoticky normálne rozdelenie náhodnej veličiny Z

$$Z \xrightarrow{d} \mathbf{N}(0, 1), \quad n \rightarrow \infty.$$

Vieme, že ak $X \sim \mathbf{N}(0, 1)$, tak platí $\mathbf{P}[|X| \leq u_{\alpha/2}] = 1 - \alpha$, a teda pre $n \rightarrow \infty$ dostaneme

$$\begin{aligned} \mathbf{P}[|Z| \leq u_{\alpha/2}] &\rightarrow 1 - \alpha, \\ \mathbf{P}\left[\frac{|\hat{p} - p|}{\sqrt{\widehat{pq}/n}} \leq u_{\alpha/2}\right] &\rightarrow 1 - \alpha, \\ \mathbf{P}\left[\hat{p} - u_{\alpha/2}\sqrt{\widehat{pq}/n} \leq p \leq \hat{p} + u_{\alpha/2}\sqrt{\widehat{pq}/n}\right] &\rightarrow 1 - \alpha. \end{aligned} \quad (2.2)$$

Interval

$$I_S = \left[\hat{p} - u_{\alpha/2}\sqrt{\widehat{pq}/n}, \hat{p} + u_{\alpha/2}\sqrt{\widehat{pq}/n}\right]$$

sa nazýva *Waldov (štandardný) interval spoľahlivosti* pre parameter p a často je doporučovaný v základných kurzoch matematickej štatistiky. Jeho veľká výhoda je ľahko zapamätateľný tvar. Naopak nevýhodou je, že ho nemôžeme použiť, ak $X \in \{0, 1\}$, pretože potom dostávame triviálny interval.

Poznámka. Vzťah (2.2) zaručuje, že pre všetky $p \in (0, 1)$ platí $\mathbf{P}(I_S \ni p) \rightarrow 1 - \alpha$ ak $n \rightarrow \infty$.

Tento interval môžeme dostať aj invertovaním Waldovho štatistického testu, z čoho vznikol aj jeho názov. Je to jeden z parametrických štatistických testov založených na vierohodnostnej funkcii. Predpokladáme, že \hat{p} je maximálne vierohodný odhad parametra p . Ak platí nulová hypotéza $H_0 : p = p_0$, tak Waldova testovacia štatistika $W(p_0) = (\hat{p} - p_0)^2 / \text{var } \hat{p}$ má χ_1^2 -rozdelenie. Hypotézu H_0 zamietame, pokiaľ platí $W(p_0) \geq \chi_{1,\alpha}^2$, kde $\chi_{1,\alpha}^2$ je $1 - \alpha$ kvantil rozdelenia chí-kvadrát s jedným stupňom voľnosti. Teda do intervalového odhadu zahrnieme také $p \in (0, 1)$, ktoré spĺňajú $W(p) \leq \chi_{1,\alpha}^2$. S využitím $\chi_{1,\alpha}^2 = u_{\alpha/2}^2$ dostaneme očakávaný výsledok.

2.2 Wilsonov interval

Rovnako ako Waldov interval, vychádza z asymptoticky normálneho rozdelenia náhodnej veličiny ξ , ale s tým rozdielom, že už sa ďalej neodhaduje rozptyl odhadu

\hat{p} . A preto očakávame, že takto zkonštruovaný interval bude presnejší. Priamo zo vzťahu (2.1) plynie

$$P \left[\frac{|\hat{p} - p|}{\sqrt{pq/n}} \leq u_{\alpha/2} \right] \rightarrow 1 - \alpha, \quad \text{pre } n \rightarrow \infty.$$

Kvôli prehľadnosti označme $\mu = u_{\alpha/2}$.

Interval spoľahlivosti zrejme dostaneme ako riešenie nerovnice

$$\begin{aligned} \frac{|\hat{p} - p|}{\sqrt{pq/n}} &\leq \mu, \\ (\hat{p} - p)^2 &\leq \mu^2 p(1-p)/n, \\ \hat{p}^2 - 2\hat{p}p + p^2 &\leq (p - p^2) \frac{\mu^2}{n}, \\ \left(1 + \frac{\mu^2}{n}\right) p^2 - \left(2\hat{p} + \frac{\mu^2}{n}\right) p + \hat{p}^2 &\leq 0. \end{aligned} \quad (2.3)$$

Vyriešením kvadratickej rovnice dostávame koncové body intervalu

$$p_{1,2} = \frac{X + \mu^2/2}{n + \mu^2} \pm \frac{\mu}{n + \mu^2} \sqrt{n\hat{p}(1-\hat{p}) + \mu^2/4}.$$

Interval

$$I_W = \left[\frac{X + \mu^2/2}{n + \mu^2} - \frac{\mu\sqrt{n}}{n + \mu^2} \sqrt{\hat{p}(1-\hat{p}) + \frac{\mu^2}{4n}}, \frac{X + \mu^2/2}{n + \mu^2} + \frac{\mu\sqrt{n}}{n + \mu^2} \sqrt{\hat{p}(1-\hat{p}) + \frac{\mu^2}{4n}} \right]$$

sa nazýva *Wilsonov interval spoľahlivosti*.

Poznámka. Stred Wilsonova intervalu sa dá prepísať do tvaru

$$\hat{p} \frac{n}{n + \mu^2} + 1/2 \frac{\mu^2}{n + \mu^2},$$

čo je vážený priemer čísiel \hat{p} a $1/2$, a to znamená, že leží bližšie k $1/2$ ako stred štandardného intervalu, ale pretože $n/(n + \mu^2) \rightarrow 1$, tak zrejme stred Wilsonovho intervalu sa asymptoticky blíži k bodovému odhadu.

Wilsonov interval taktiež môžeme dostať invertovaním štatistického testu založeného na vierohodnostnej funkcii. Konkrétne ide o Raov test, ktorý sa nazýva aj skórový test, a preto sa tomuto intervalu hovorí tiež *skórový interval*. Aj v tomto teste vezmeme za nulovú hypotézu $H_0 : p = p_0$. Testovacia štatistika má v našom prípade tvar

$$LM(p_0) = \frac{[L'(p_0)]^2}{I(p_0)},$$

kde $L(p) = \log f_p(X)$ je logaritmická vierohodnostná funkcia a $I(p_0)$ je Fischerova informácia v bode p_0 . Ak platí hypotéza H_0 , potom $LM(p_0)$ má χ^2_1 -rozdelenie. Hypotézu H_0 zamietneme, ak platí $LM(p_0) \geq \chi^2_{1,\alpha}$ a do intervalového odhadu teda zahrnieme také $p \in (0, 1)$, ktoré spĺňajú $LM(p) \leq \chi^2_{1,\alpha} = u_{\alpha/2}^2$. Pretože

$$L'(p) = \frac{X - np}{p(1-p)} \quad \text{a} \quad I(p) = \frac{n}{p(1-p)},$$

dostaneme intervalový odhad ako riešenie nerovnice

$$\left[\frac{X - np}{p(1-p)} \right]^2 \leq u_{\alpha/2}^2 \frac{p(1-p)}{n},$$

ktorá sa už dá upraviť na tvar (2.3).

2.3 Agresti-Coullov interval

Motivácia. Chceli by sme dostať interval, ktorý ma podobne jednoduchý tvar ako štandardný interval pre parameter p , ale je spoľahlivejší. Toho môžeme dosiahnuť napríklad tým, že použijeme stred Wilsonovho intervalu namiesto bodového odhadu.

Položme

$$\tilde{X} = X + u_{\alpha/2}^2/2, \quad \tilde{n} = n + u_{\alpha/2}^2$$

a označme

$$\tilde{p} = \tilde{X}/\tilde{n}, \quad \tilde{q} = 1 - \tilde{p}.$$

Potom interval

$$I_{AC} = \left[\tilde{p} - u_{\alpha/2} \sqrt{\tilde{p}\tilde{q}/\tilde{n}}, \tilde{p} + u_{\alpha/2} \sqrt{\tilde{p}\tilde{q}/\tilde{n}} \right]$$

nazveme *Agresti-Coullov interval spoľahlivosti*.

Poznámka. Názov pochádza z článku [4] s odkazom na článok [1], v ktorom autori, pre špeciálny prípad $\alpha = 0.05$, navrhli používať pravidlo „pridať dva úspechy a dva neúspechy“ a potom použiť klasický interval spoľahlivosti. Pretože $u_{0.025}^2 \approx 1.96^2 \approx 4$, tak platí

$$\frac{X + u_{0.025}^2/2}{n + u_{0.025}^2} \approx \frac{X + 2}{n + 4},$$

čo odpovedá danému pravidlu.

Veta 2.3.1. *Agresti-Coullov interval nie je nikdy kratší ako Wilsonov interval.*

Dôkaz. Pretože stredy oboch intervalov splývajú stačí ukázať, že platí

$$\begin{aligned} \frac{\sqrt{n}}{n + \mu^2} \sqrt{\hat{p}(1-\hat{p}) + \frac{\mu^2}{4n}} &\leq \sqrt{\tilde{p}\tilde{q}/\tilde{n}}, \\ \frac{1}{n + \mu^2} \sqrt{X - \frac{X^2}{n} + \frac{\mu^2}{4}} &\leq \sqrt{\frac{X + \mu^2/2}{(n + \mu^2)^2} \left(1 - \frac{X + \mu^2/2}{n + \mu^2}\right)}, \\ \frac{1}{(n + \mu^2)^2} \left(X - \frac{X^2}{n} + \frac{\mu^2}{4}\right) &\leq \frac{X + \mu^2/2}{(n + \mu^2)^2} \left(\frac{n - X + \mu^2/2}{n + \mu^2}\right), \\ Xn + X\mu^2 - X^2 - \frac{X^2\mu^2}{n} + \frac{n\mu^2}{4} + \frac{\mu^4}{4} &\leq Xn + \frac{n\mu^2}{2} - X^2 - \frac{X\mu^2}{2} + \frac{X\mu^2}{2} + \frac{\mu^4}{4}, \\ -\frac{X^2\mu^2}{n} + X\mu^2 - \frac{n\mu^2}{4} &\leq 0, \\ -X^2 + nX - n^2/4 &\leq 0. \end{aligned}$$

Keďže $D = n^2 - 4(-1)(-n^2/4) = 0$ a koeficient u X^2 je záporný, tak daná nerovnica platí pre všetky $X \in \mathbb{R}$. \square

2.4 Arkussínusový interval

Transformácia stabilizujúca rozptyl

Nech náhodná veličina X má rozdelenie závislé na parametri θ a navyše nech platí $\mathbf{E} X = \theta$ a $\text{var } X = \sigma^2(\theta)$, kde $\sigma(\theta)$ je hladká funkcia premennej θ . Chceme nájsť netriviálnu funkciu g takú, aby rozptyl náhodnej veličiny $Y = g(X)$ nezávisel na parametri θ . Táto úloha všeobecne nemusí mať presné riešenie, ale môžeme použiť nasledujúcu aproximáciu. Podľa Taylorovho rozvoja v bode $\mathbf{E} X = \theta$ platí

$$g(X) \approx g(\theta) + (X - \theta)g'(\theta)$$

a s využitím tohoto výrazu dostávame

$$\mathbf{E} g(X) \approx g(\theta) + (\mathbf{E} X - \theta)g'(\theta) = g(\theta), \quad (2.4)$$

$$\text{var } g(X) \approx [g'(\theta)]^2 \text{var } (X - \theta) = [g'(\theta)]^2 \sigma^2(\theta). \quad (2.5)$$

Výraz $[g'(\theta)]^2 \sigma^2(\theta)$ zrejme nebude závisieť na θ , ak položíme

$$g'(\theta)\sigma(\theta) = c, \quad (2.6)$$

kde c je nejaká konštanta. Z tejto podmienky vieme vyjadriť hľadanú funkciu g v tvare

$$g(\theta) = c \int \frac{d\theta}{\sigma(\theta)}. \quad (2.7)$$

Konštantu c volíme tak, aby $g(\theta)$ mala čo najvýhodnejší tvar. Navyše zo vzťahov (2.5) a (2.6) je zrejmé, že $\text{var } g(X) \approx c^2$.

Poznámka. Rozdelenie náhodnej veličiny Y býva veľmi blízke normálnemu rozdeleniu (viď Anděl [2], str. 62).

Arkussínusová transformácia

Náhodná veličina \hat{p} zrejme spĺňa predpoklady predchádzajúcej úlohy. V tomto prípade je $\theta = p$ a $\sigma^2(\theta) = p(1-p)/n$. Dosadením do vzťahu (2.7) dostávame

$$g(p) = c \int \frac{dp}{\sqrt{p(1-p)/n}} = c\sqrt{n} \int \frac{dp}{\sqrt{p}\sqrt{1-p}}.$$

S využitím substitúcie $\sqrt{p} = x$ ďalej platí

$$g(x^2) = 2c\sqrt{n} \int \frac{dx}{\sqrt{1-x^2}} = 2c\sqrt{n} \arcsin x.$$

Voľbou $c = 1/(2\sqrt{n})$ nakoniec dostávame tvar

$$g(p) = \arcsin \sqrt{p}.$$

Náhodná veličina $Y = g(\hat{p}) = \arcsin \sqrt{\hat{p}}$ má približne normálne rozdelenie a ešte navyše platí

$$\mathbf{E} Y \approx g(p) = \arcsin \sqrt{p}, \quad \text{var } Y \approx c^2 = 1/4n.$$

Potom náhodná veličina

$$\frac{Y - \mathbf{E} Y}{\text{var } Y} = \frac{\arcsin \sqrt{\hat{p}} - \arcsin \sqrt{p}}{\sqrt{1/4n}}$$

má približne normované normálne rozdelenie $N(0, 1)$. A teda platí

$$P \left[\arcsin \sqrt{\widehat{p}} - u_{\alpha/2} \sqrt{1/4n} \leq \arcsin \sqrt{p} \leq \arcsin \sqrt{\widehat{p}} + u_{\alpha/2} \sqrt{1/4n} \right] \approx 1 - \alpha.$$

Pretože \sin^2 je na intervale $(0, \pi/2)$ rastúci, tak v prípade, že

$$\arcsin \sqrt{\widehat{p}} - u_{\alpha/2} \sqrt{1/4n} \geq 0, \quad (2.8)$$

$$\arcsin \sqrt{\widehat{p}} + u_{\alpha/2} \sqrt{1/4n} \leq \pi/2, \quad (2.9)$$

potom platí

$$P \left[\sin^2(\arcsin \sqrt{\widehat{p}} - u_{\alpha/2} \sqrt{1/4n}) \leq p \leq \sin^2(\arcsin \sqrt{\widehat{p}} + u_{\alpha/2} \sqrt{1/4n}) \right] \approx 1 - \alpha.$$

Ak nebude splnená podmienka (2.8), tak za dolnú hranicu intervalu vezmeme $\sin^2(0) = 0$ a ak neplatí (2.9), potom za hornú hranicu vezmeme $\sin^2(\pi/2) = 1$.

Interval $I_{Arc} = [d, h]$, kde

$$d = \sin^2 \left(\max\{0, \arcsin \sqrt{\widehat{p}} - u_{\alpha/2} \sqrt{1/4n}\} \right),$$

$$h = \sin^2 \left(\min\{\pi/2, \arcsin \sqrt{\widehat{p}} + u_{\alpha/2} \sqrt{1/4n}\} \right),$$

nazveme *arkussínusový interval spoľahlivosti*.

2.5 Clopper-Pearsonov interval

Uvažujme nasledujúci test v binomickom rozdelení.

Testujeme hypotézu $H_0 : p \geq p_0$ oproti $H_1 : p < p_0$. Hypotézu H_0 zamietneme, ak pozorovaná hodnota $X = x$ bude dostatočne malá. Presnejšie, najšš kritický obor na hladine α znamená najšš najväčšie $k \in \{0, 1, \dots, n-1\}$, ktoré vyhovuje podmienke

$$P_p[X \leq k] \leq \alpha, \quad \forall p \geq p_0.$$

Pretože výraz $P_p[X \leq k]$ je pre pevné k klesajúci v premennej p , stačí, aby platilo

$$P_{p_0}[X \leq k] \leq \alpha.$$

Z diskretnosti rozdelenia náhodnej veličiny X vyplýva, že presná hladina testu je zvyčajne menšia ako zvolené α .

Na základe tohoto testu odvodíme horný odhad pre parameter p . Pokiaľ $X = x$ je pozorovaná hodnota z binomického rozdelenia, tak zrejme p_0 vyhovujúce rovnici

$$P_{p_0}[X \leq x] = \alpha$$

je najmenšie p_0 , pre ktoré môžeme zamietnuť hypotézu H_0 na hladine α . To znamená, že p_0 je horný odhad parametra p o spoľahlivosti $1 - \alpha$. Analogicky môžeme odvodiť dolný odhad o spoľahlivosti $1 - \alpha$ ako riešenie rovnice

$$P_{p_0}[X \geq x] = \alpha.$$

Takže dolnú a hornú hranicu intervalového odhadu o spoľahlivosti na úrovni $1 - \alpha$ dostaneme ako riešenie nasledujúcich rovníc v tomto poradí

$$\sum_{k=x}^n \binom{n}{k} p_1^k (1 - p_1)^{n-k} = \alpha/2, \quad (2.10)$$

$$\sum_{k=0}^x \binom{n}{k} p_2^k (1 - p_2)^{n-k} = \alpha/2, \quad (2.11)$$

okrem prípadu $X = 0$, kedy vezmeme za dolnú hranicu 0 a pre $X = n$ položíme hornú hranicu priamo 1.

Označme $F_{m,n}^*(x)$ distribučnú funkciu F-rozdelenia o m a n stupňoch voľnosti. Potom pre $x = 0, 1, \dots, n - 1$ platí nasledujúci vzťah

$$\sum_{k=0}^x \binom{n}{k} p^k (1 - p)^{n-k} = F_{2(n-x), 2(x+1)}^* \left[\frac{(x+1)(1-p)}{p(n-x)} \right]. \quad (2.12)$$

Vid' Anděl [2], str. 91.

Rovnicu (2.10) môžeme prepísať do tvaru

$$\sum_{k=0}^{x-1} \binom{n}{k} p_1^k (1 - p_1)^{n-k} = 1 - \alpha/2$$

a s využitím vzťahu (2.12) dostávame

$$F_{2(n-x+1), 2x}^* \left(\frac{x(1-p_1)}{p_1(n-x+1)} \right) = 1 - \alpha/2.$$

Symbolom $F_{m,n}(\alpha)$ označujeme kritickú hodnotu rozdelenia $F_{m,n}$ na hladine α . To znamená, že $F_{m,n}^*(F_{m,n}(\alpha)) = 1 - \alpha$ a teda dostávame rovnicu v tvare

$$\frac{x(1-p_1)}{p_1(n-x+1)} = F_{2(n-x+1), 2x}(\alpha/2) = \frac{1}{F_{2x, 2(n-x+1)}(1 - \alpha/2)},$$

pričom druhá rovnosť plynie z vety A.7. Odtiaľ zrejme už môžeme vyjadriť p_1 ako

$$p_1 = \left(\frac{n-x+1}{xF_{2x, 2(n-x+1)}(1 - \alpha/2)} + 1 \right)^{-1}.$$

Analogickým postupom dostaneme z rovnice (2.11) riešenie

$$p_2 = \left(\frac{n-x}{(x+1)F_{2(x+1), 2(n-x)}(\alpha/2)} + 1 \right)^{-1}.$$

Interval

$$I_{CP} = \left[\left(\frac{n-x+1}{xF_{2x, 2(n-x+1)}(1 - \alpha/2)} + 1 \right)^{-1}, \left(\frac{n-x}{(x+1)F_{2(x+1), 2(n-x)}(\alpha/2)} + 1 \right)^{-1} \right],$$

kde $x = 1, 2, \dots, n - 1$, sa nazýva *Clopper-Pearsonov interval spoľahlivosti*.

Poznámka. Takto zkonštruovaný interval spoľahlivosti má vždy skutočnú pravdepodobnosť pokrytia parametra p aspoň $1 - \alpha$, každopádne väčšinou je toto pokrytie ostro väčšie ako daná hladina, čo je spôsobené tým, že binomické rozdelenie patrí medzi diskkrétne pravdepodobnostné rozdelenia. Interval pre parameter nejakého spojitého rozdelenia odvodený z analogického testu, by mal pravdepodobnosť pokrytia presne $(1 - \alpha)100\%$.

2.6 Interval z vierohodnostného pomeru

Ďalší interval spoľahlivosti dostaneme invertovaním štatistického testu založeného na vierohodnostnom pomere, v ktorom je oproti nulovej hypotéze $H_0 : p = p_0$ postavená obojstranná alternatíva $H_1 : p \neq p_0$. Testovacia štatistika

$$LR(p_0) = 2 [L(\hat{p}) - L(p_0)]$$

má v prípade platnosti nulovej hypotézy asymptoticky χ^2_1 -rozdelenie. Vid' Anděl [3], str. 179. To znamená, že nulovú hypotézu zamietame na hladine α v prípade, ak $LR(p_0) \geq \chi^2_{1,\alpha}$.

Intervalový odhad o spoľahlivosti $1 - \alpha$ bude pozostávať zo všetkých p_0 , pre ktoré sa nedá zamietnuť nulová hypotéza na hladine α , čo znamená, že spĺňajú podmienku

$$2 [L(\hat{p}) - L(p_0)] \leq \chi^2_{1,\alpha} = u^2_{\alpha/2}.$$

Logaritmická vierohodnostná funkcia $L(p) = \log \left[\binom{n}{X} p^X (1-p)^{n-X} \right]$ je definovaná na intervale $(0, 1)$, a teda ďalej predpokladáme, že $\hat{p} \neq 0$ a $\hat{p} \neq 1$, čo znamená, že $X \notin \{0, n\}$. $L(p)$ zrejme nadobúda maximum v bode \hat{p} , pretože sa jedná o maximálne vierohodný odhad a navyše zo vzťahu

$$L'(p) = \frac{X - np}{p(1-p)}$$

vyplýva, že je rastúca na intervale $(0, \hat{p})$ a klesajúca na intervale $(\hat{p}, 1)$. Takže do intervalového odhadu zahrnieme tie p_0 , pre ktoré $L(p_0)$ neklesne od maximálnej hodnoty v bode \hat{p} o viac ako $u^2_{\alpha/2}/2$. Krajný body intervalu spoľahlivosti teda nájdeme ako riešenia rovnice

$$\begin{aligned} L(\hat{p}) - L(p_0) &= u^2_{\alpha/2}/2, \\ \log \left[\frac{(X/n)^X (1 - X/n)^{n-X}}{p_0^X (1 - p_0)^{n-X}} \right] &= u^2_{\alpha/2}/2. \end{aligned} \quad (2.13)$$

Túto rovnicu vo väčšine prípadoch nedokážeme vyriešiť analyticky ale len pomocou počítača. Napríklad v štatistickom programe R môžeme použiť funkciu `uniroot()`, ktorá numericky hľadá koreň zadanej funkcie na danom intervale s požadovanou presnosťou. Spoľahlivo funguje v prípade, keď je daná funkcia na tomto intervale monotónna, pretože vie nájsť len jeden koreň a vyžaduje, aby funkčné hodnoty v krajných bodoch mali opačné znamienka, prípadne sa rovnali nule. Takže zvlášť nájdeme spodnú hranicu intervalu spoľahlivosti ako riešenie rovnice (2.13) na intervale $(0, X/n)$ a potom hornú hranicu na intervale $(X/n, 1)$.

V prípade $X = 0$ je jediný problém s výrazom $(X/n)^X$, ktorý nie je definovaný. Pretože ale existuje vlastná jednostranná limita $\lim_{x \rightarrow 0^+} x^x = 1$, môžeme nahradiť tento výraz v (2.13) jeho limitou a po dosadení $X = 0$, dostávame rovnosť v nasledujúcom tvare

$$\begin{aligned} \log \left[\frac{1}{(1-p_0)^n} \right] &= u^2_{\alpha/2}/2, \\ -n \log(1-p_0) &= u^2_{\alpha/2}/2, \\ p_0 &= 1 - \exp \left(-\frac{u^2_{\alpha/2}}{2n} \right). \end{aligned}$$

Zrejme platí $\exp(-u_{\alpha/2}^2/2n) < 1$, a teda $p_0 \in (0, 1)$. Interval spoľahlivosti bude mať v tomto špecifickom prípade tvar $(0, p_0)$.

Podobne dostaneme v prípade $X = n$ riešenie

$$p_0 = \exp\left(-\frac{u_{\alpha/2}^2}{2n}\right)$$

a zkonštruujeme interval v tvare $(p_0, 1)$. Takto zkonštruovaný interval označíme I_{VP} .

3. Porovnanie intervalov

V tejto časti sa budeme snažiť kvalitatívne zhodnotiť a porovnať jednotlivé intervaly spoľahlivosti (IS) teoreticky rozpracované v kapitole 2. Z diskkrétnej povahy binomického rozdelenia vieme, že pre väčšinu p nie je vôbec možné dosiahnuť presnú pravdepodobnosť pokyitia rovnú danej nominálnej úrovni. Preto základným kritériom pre porovnávanie intervalových odhadov binomického podieľu bude závislosť skutočnej pravdepodobnosti pokrytia od neznámeho parametra p . O nič menej dôležitým ukazovateľom nie je ani šírka intervalu, ktorá ale úzko súvisí s prvotným kritériom.

Nutnosťou k hodnoteniu jednotlivých metód je schopnosť vedieť skúmané intervaly vypočítať. Pre tento účel si napíšeme jednoduchú funkciu v štatistickom programe R, ktorý ďalej budeme využívať aj na samotné porovnávanie. Základná verzia softvéru R síce má nejaké funkcie na intervalové odhady, ale je to značne neprehľadné a pravdepodobne sa všetky intervaly vypočítať ani nedajú. Len pre zaujímavosť, napríklad Clopper-Pearsonov interval dostaneme z funkcie `binom.test(x,n)` a Wilsonov interval pomocou `prop.test(x,n,correct=F)`. Preto jednoduchšie je napísať si vlastnú funkciu pre všetky potrebné intervaly. Z repozitárov sa dá stiahnuť aj balíček `binom`, v ktorom sú až na arkussínusový interval všetky ostatné metódy. Zvedavého čitateľa by mohla z tohoto balíčka zaujať funkcia `binom.confint()`, ktorá dokáže vypočítať aj rôzne iné intervaly.

3.1 Skutočná pravdepodobnosť pokrytia

Nech $I(X) = (d(X), h(X))$ je nejaký interval spoľahlivosti. Pre pevnú hodnotu parametra p definujeme skutočnú pravdepodobnosť pokrytia intervalového odhadu ako pravdepodobnosť, že interval pokryje túto hodnotu. Takže

$$\begin{aligned}
 C_n(p) &= \mathbf{P}_p [d(X) < p < h(X)], \\
 C_n(p) &= \sum_{x=0}^n I_{\{d(x) < p < h(x)\}} \mathbf{P}_p [X = x], \\
 C_n(p) &= \sum_{x=0}^n I_{\{d(x) < p < h(x)\}} \binom{n}{x} p^x (1-p)^{n-x},
 \end{aligned}$$

pričom $I_{\{d(x) < p < h(x)\}}$ je rovné 1, ak platí $d(x) < p < h(x)$, v opačnom prípade je to 0.

Interval spoľahlivosti		$d(X)$	$h(X)$
Štandardný	I_S	0.000	0.829
Wilsonov	I_W	0.118	0.769
Agresti-Coullov	I_{AC}	0.116	0.771
Arkussínusový	I_{Arc}	0.060	0.813
Clopper-Pearsonov	I_{CP}	0.053	0.853
IS z vierohod. pomeru	I_{VP}	0.081	0.801

Tabuľka 3.1: 95% IS pre $X = 2, n = 5$.

Jednotlivé intervaly môžeme jednoducho porovnať pomocou grafov závislosti skutočnej pravdepodobnosti pokrytia $C_n(p)$ od parametra p na celom intervale $(0, 1)$. Prípadmi, keď n je veľmi malé, nemá význam sa zaoberať, pretože intervaly spoľahlivosti sú príliš široké, a teda konštrukcia takéhoto odhadu je skôr strata času ako získanie nejakej informácie o parametri p . V tabuľke 3.1 vidieť, že 95% intervaly spoľahlivosti pre $X = 2, n = 5$ zahŕňajú zhruba $3/4$ intervalu $(0, 1)$.

Z obrázka 3.1 vidíme, že pre malé n , sa v pravdepodobnosti pokrytia objavujú nemalé výkyvy od nominálnej úrovne. Nie je prekvapujúce, že Waldov (štandardný) IS je nedostatočný, keďže viackrát využíva asymptotický prístup. Prekvapivo celkom dobre vyzerá Wilsonov interval, a teda sa ukazuje, že rozdiel medzi použitím odhadu rozptylu alebo presného rozptylu je značný. Ešte lepším sa v tomto prípade zdá byť Agresti-Coullov IS, ktorý nie je pre krajné hodnoty parametra p taký „divoký“ ako Wilsonov. Clopper-Pearsonov interval musí mať v každom prípade skutočnú pravdepodobnosť pokrytia väčšiu ako je nominálna, no pre malé n je skoro na celom intervale $(0, 1)$ priveľmi nadhodnotený.

Závislosť pravdepodobnosti pokrytia pre väčšie n , konkrétne pre $n = 40$, vidíme na obrázku 3.2. Oproti predchádzajúcemu prípadu došlo k veľkému zlepšeniu, obzvlášť čo sa týka Waldovho intervalu, no v porovnaní s ďalšími je tento nekonkurencieschopný. Clopper-Pearsonov IS sa zlepšuje asi najpomalšie, no to je spôsobené tým, že sa drží stabilne nad nominálnou úrovňou, a teda nemá toľko priestoru na zlepšovanie. Ostatné intervaly sa už len na základe týchto grafov porovnávajú dosť zložito, pretože sú si veľmi podobné až na krajné hodnoty parametra p , kde sú dosť nevyspytateľné. Len Agresti-Coullov interval nemá také extrémne výkyvy.

Nakoniec na obrázku 3.3 vidíme ešte grafy pre $n = 75$. Situácia začína byť dosť neprehľadná, pretože v grafoch je už príliš mnoho skokov, ktoré pribúdajú s rastúcim n . Jednoznačne sa dá ale rozoznať nedostatočná pravdepodobnosť pokrytia štandardným IS a naopak zbytočne opatrný prístup v prípade Clopper-Pearsonovho IS. Celkovo môžeme usúdiť, že s rastúcim n nedochádza k nepredvídateľným zmenám, ale všetky IS sa pomaly vylepšujú v pokrytí. Za povšimnutie stojí aj určitá podoba medzi pokrytiami Wilsonovho a Agresti-Coullovho intervalu, čo je spôsobené tým, že druhý menovaný v podstate vychádza z dosť komplikovaného vzorca skórového intervalu, ktorý ale zjednodušuje.

Aby sme mohli jednotlivé IS medzi sebou jednoznačne porovnať, zadefinujeme ešte priemernú pravdepodobnosť pokrytia na intervale $(0, 1)$ ako

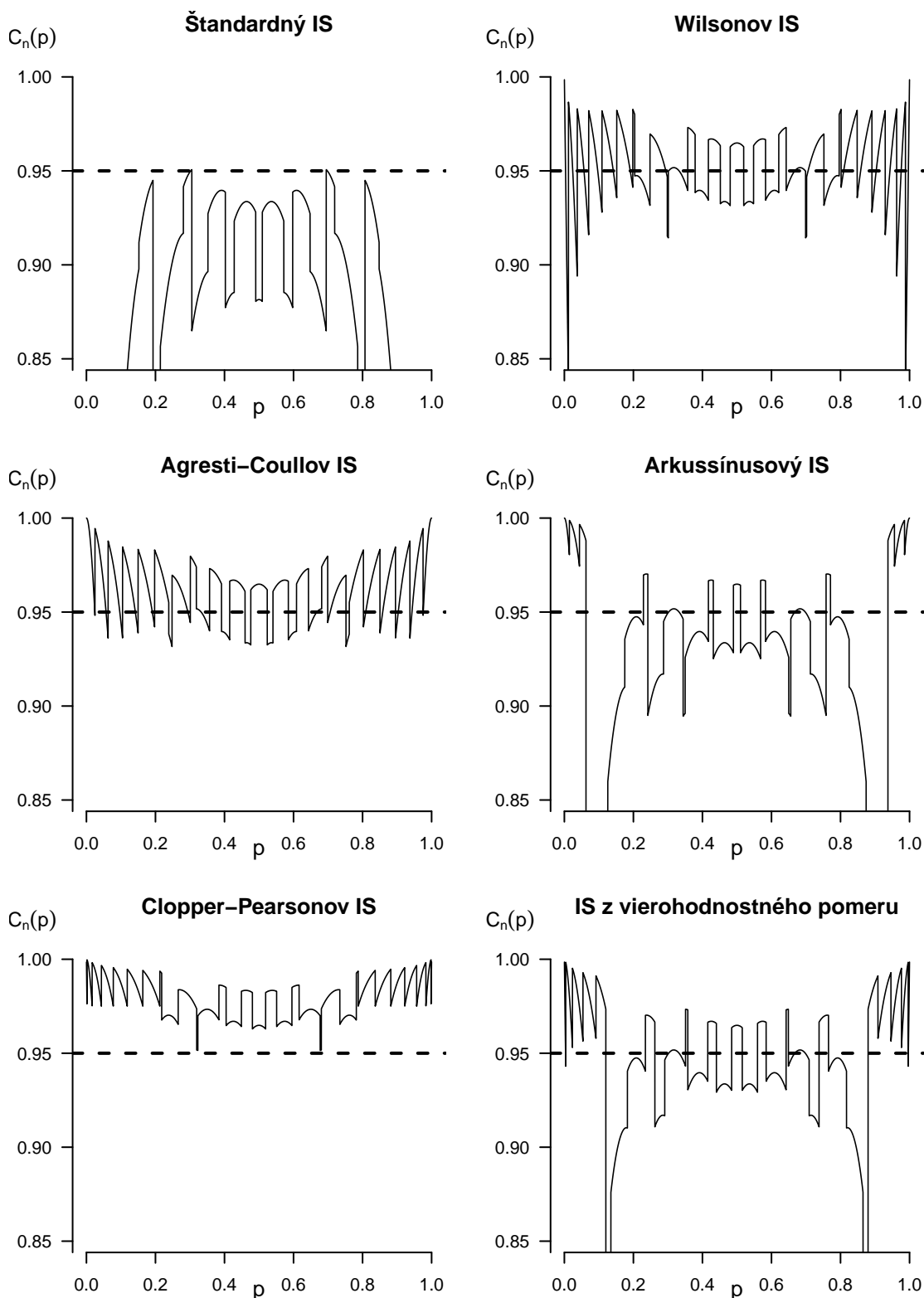
$$\bar{C}_n = \int_0^1 C_n(p) dp.$$

Tento integrál vypočítame numericky. Zvolíme si dostatočne veľký počet rovnomerne rozmiestnených hodnôt p z intervalu $(0, 1)$ a pre každé p z tejto množiny vypočítame skutočnú pravdepodobnosť pokrytia $C_n(p)$. Aritmetický priemer všetkých hodnôt $C_n(p)$ bude približná hodnota \bar{C}_n .

Priemerné pravdepodobnosti pokrytia pre prípady z obrázka 3.3 sú zhrnuté v tabuľke 3.1. Okrem Waldovho IS majú všetky ostatné intervaly priemerné pokrytie na prijateľnej úrovni. Najlepšie si ale vedú Wilsonov interval, ktorý je mierne nad nominálnou úrovňou a interval z vierohodnostného pomeru, ktorého pokrytie nepatrne zaostáva za nominálnym.

Pravdepodobnos pokrytia

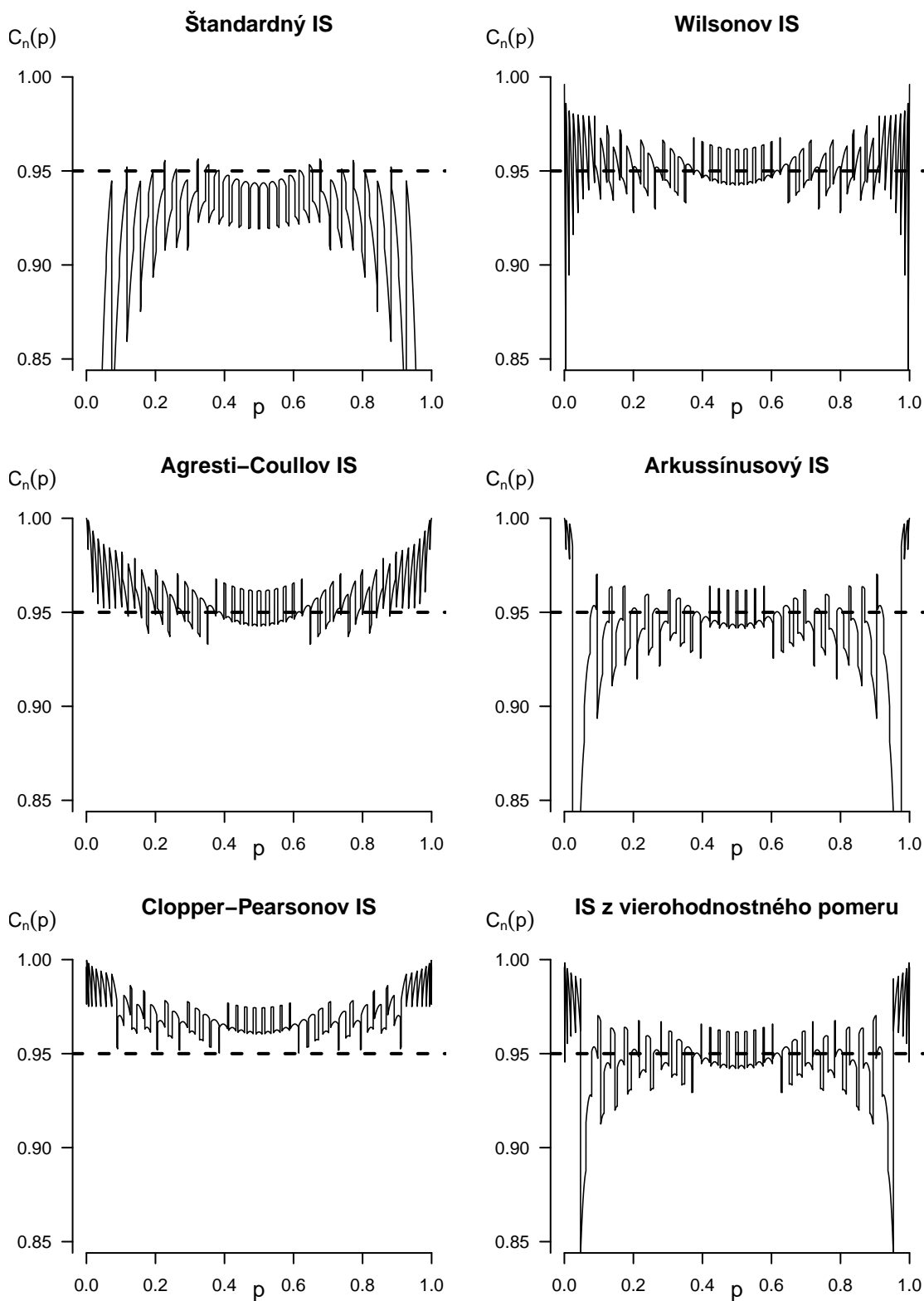
($n = 15$, $\alpha = 0.05$)



Obr. 3.1: Porovnanie skutočnej pravdepodobnosti pokrytia 95% intervalov spoľahlivosti pre $n = 15$.

Pravdepodobnos pokrytia

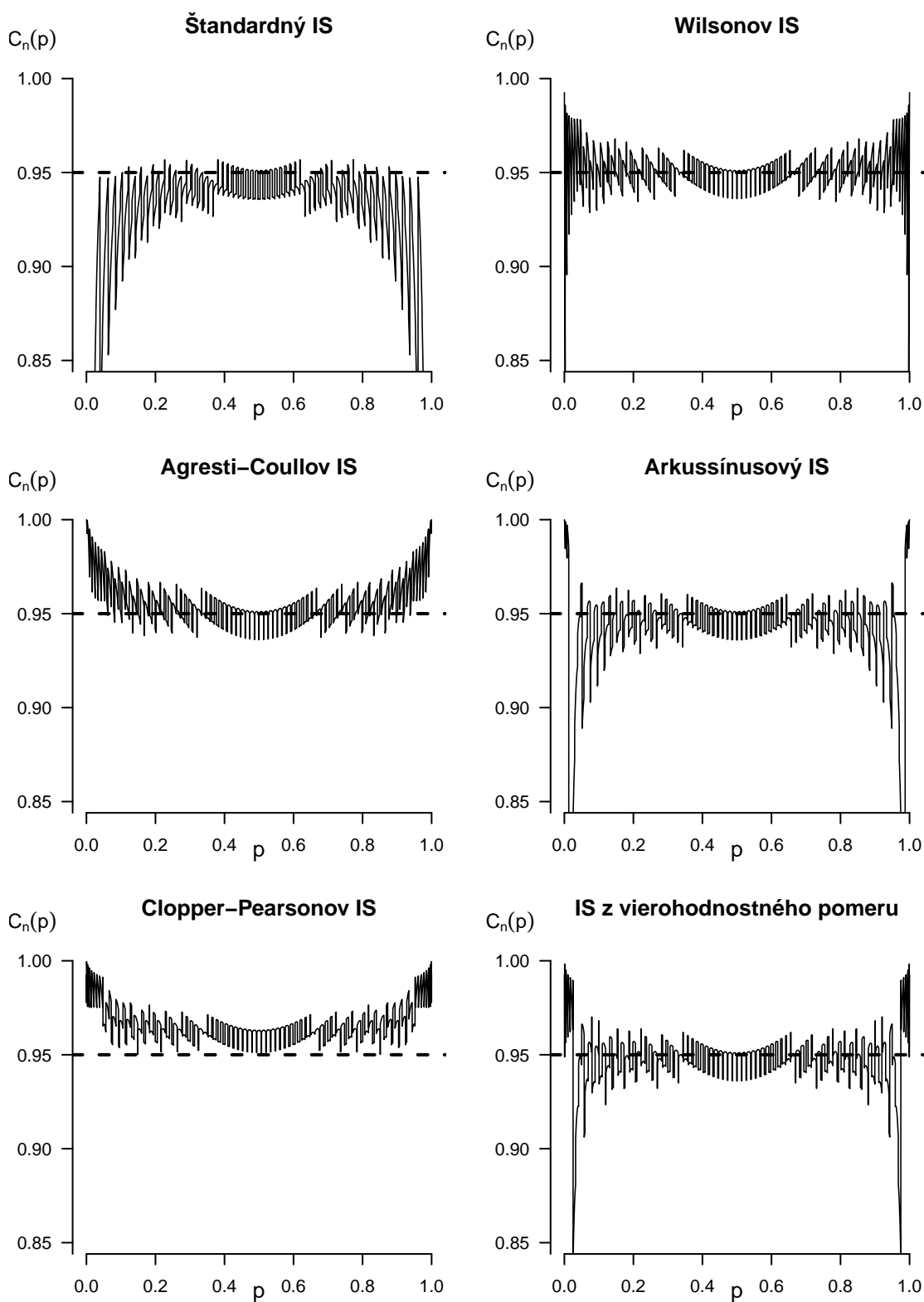
($n = 40$, $\alpha = 0.05$)



Obr. 3.2: Porovnanie skutočnej pravdepodobnosti pokrytia 95% intervalov spoľahlivosti pre $n = 40$.

Pravdepodobnos pokrytia

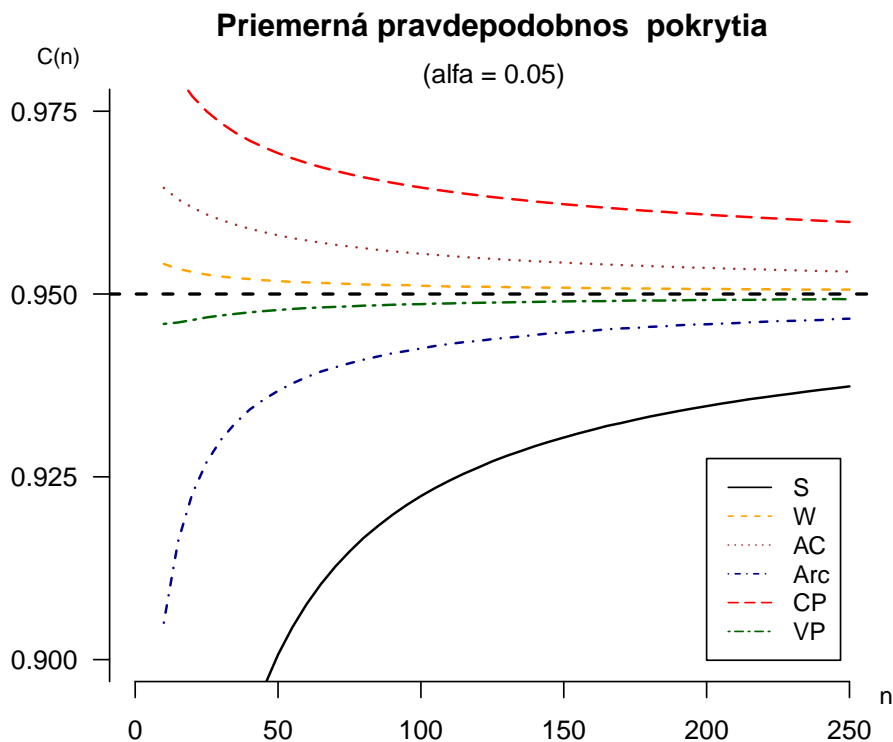
($n = 75$, $\alpha = 0.05$)



Obr. 3.3: Porovnanie skutočnej pravdepodobnosti pokrytia 95% intervalov spoľahlivosti pre $n = 75$.

IS	I_S	I_W	I_{AC}	I_{Arc}	I_{CP}	I_{VP}
\bar{C}_n	0.9148	0.9514	0.9565	0.9405	0.9664	0.9483

Tabuľka 3.2: Priemerná pravdepodobnosť pokrytia pre $n = 75$.



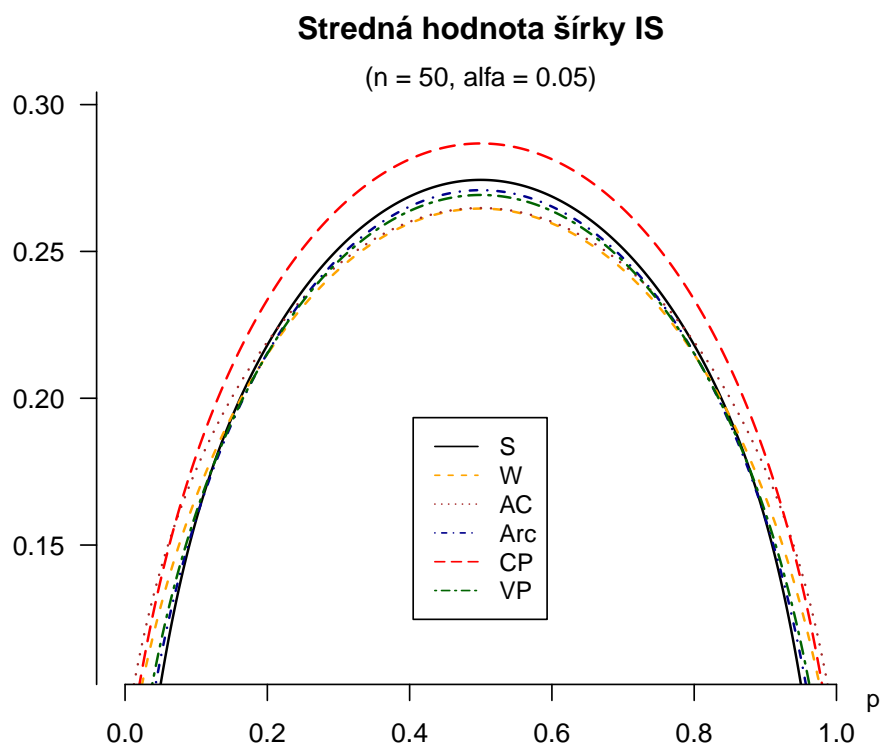
Obr. 3.4: Graf závislosti priemernej pravdepodobnosti pokrytia 95% intervalov spoľahlivosti od n .

Ďalej zostrojíme graf závislosti priemernej pravdepodobnosti pokrytia od n . Výpočet je ale už značne náročnejší a potrvá rádovo až niekoľko minút. Výsledok pre $n = 10$ až $n = 250$ vidíme na obrázku 3.4. Zaujímavé intervaly, aspoň čo sa týka priemernej pravdepodobnosti pokrytia, sú skórový a vierohodnostný interval, ktoré sa dostatočne približujú nominálnej úrovni už pre malé hodnoty n .

3.2 Stredná hodnota šírky intervalu

Označme $l_I(X) = h(X) - d(X)$ šírku intervalu spoľahlivosti $I(X) = (d(X), h(X))$. Potom pre pevné $p \in (0, 1)$ existuje stredná hodnota náhodnej veličiny $l_I(X)$, pre ktorú platí

$$\begin{aligned}
 E_p[l_I(X)] &= \sum_{x=0}^n l_I(x) P[X = x] \\
 &= \sum_{x=0}^n (h(X) - d(X)) \binom{n}{x} p^x (1-p)^{n-x}.
 \end{aligned}$$



Obr. 3.5: Graf závislosti strednej hodnoty šírky 95% intervalov spoľahlivosti od parametra p pre $n = 50$.

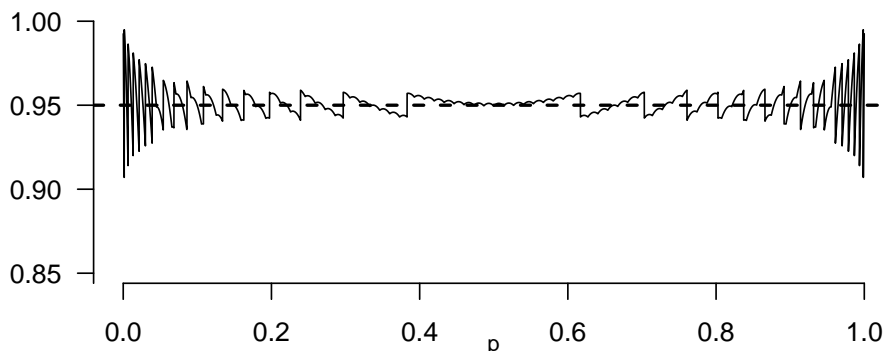
IS	$n = 15$	$n = 40$	$n = 75$	$n = 100$	$n = 150$	$n = 200$	$n = 250$
I_W	0.3687	0.2367	0.1752	0.1523	0.1248	0.1083	0.0970
I_{VP}	0.3642	0.2350	0.1744	0.1517	0.1245	0.1081	0.0968

Tabuľka 3.3: Priemerné šírky Wilsonovho IS a IS z vierohodnostného pomeru.

Obrázok 3.5 znázorňuje závislosť strednej hodnoty šírky intervalov v prípade $n = 50$. V súlade s pravdepodobnosťou pokrytia je skoro na celom intervale $(0, 1)$ najširší Clopper-Pearsonov interval. Celkom prekvapivá môže byť príliš veľká šírka Waldovho intervalu v porovnaní s jeho pokrytím, no to je spôsobené tým, že stred tohto intervalu tvorí práve bodový odhad parametra p . Toto je so vzrastajúcou vzdialenosťou parametra p od hodnoty $1/2$ viac a viac nevhodnejšie, pretože s touto vzdialenosťou narastá šikmosť binomického rozdelenia. Tento fakt takisto vysvetľuje absolútne nedostatočnú pravdepodobnosť pokrytia, keď sa p blíži ku krajným hodnotám. Inak sú všetky intervaly celkom vyrovnané, čo sa týka priemernej šírky, a preto je tento graf dosť neprehľadný.

Podobne ako v predchádzajúcej časti, môžeme ešte vypočítať priemerné stredné hodnoty šírky IS. Tieto hodnoty porovnáme už iba pre skórový a vierohodnostný interval, ktoré sa z hľadiska priemernej pravdepodobnosti pokrytia ukázali ako najvhodnejšie. Z tabuľky 3.2 vidíme, že tieto dva intervaly sú na tom skoro rovnako. Malý rozdiel v neprospech skórového intervalu znižujúci sa s rastúcim n odpovedá tomú, že priemerná pravdepodobnosť pokrytia sa v jeho prípade blíži k nominálnej úrovni zhora narozdiel od IS z vierohodnostného pomeru, ktorý sa k tejto úrovni blíži zospodu.

Optimálna pravdepodobnosť pokrytia
 $n = 75$



Obr. 3.6: Závislosť optimálnej pravdepodobnosti pokrytia od parametra p pre $n = 75$.

3.3 Optimálna pravdepodobnosť pokrytia

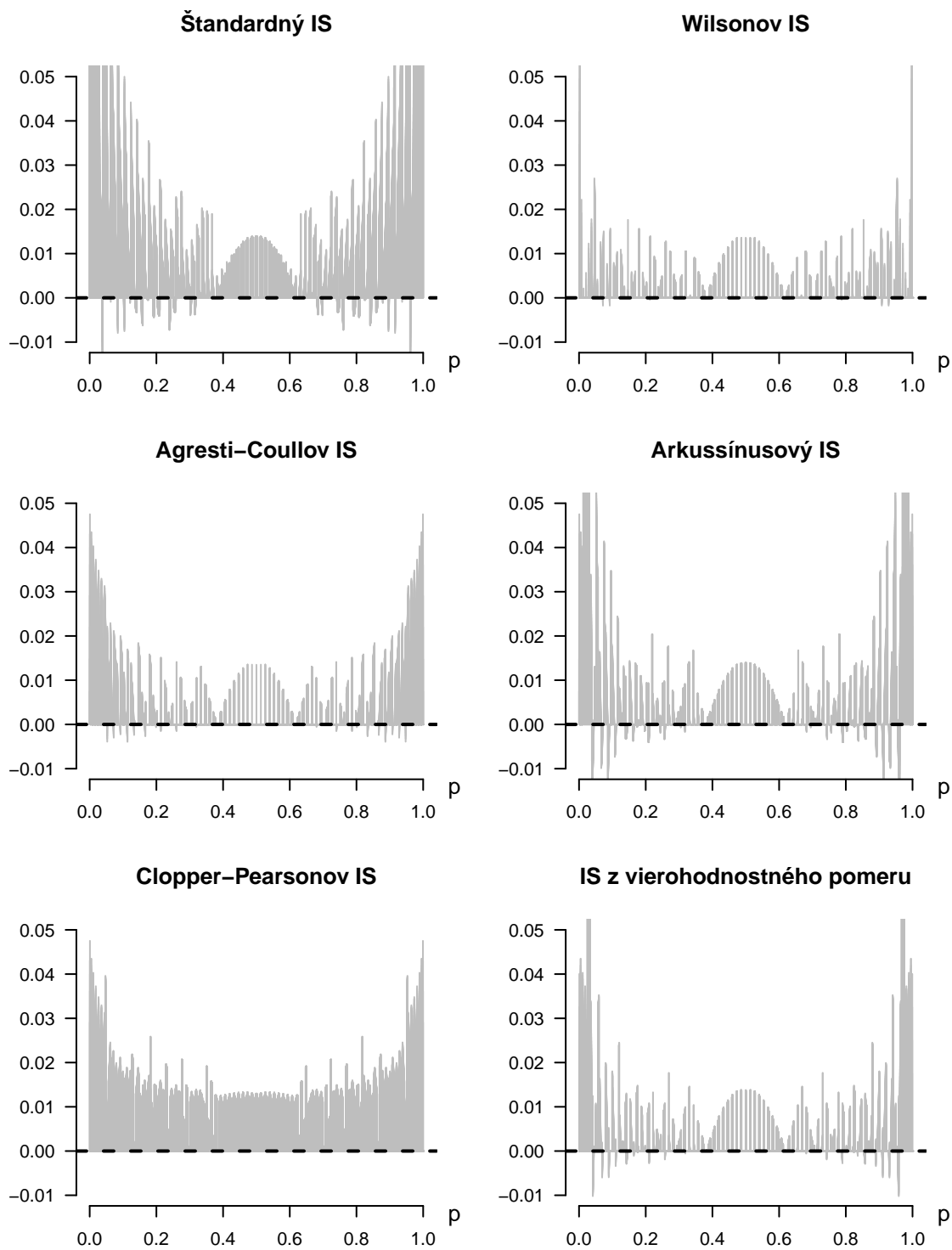
V tejto časti navrhujeme ešte jedno kritérium, podľa ktorého môžeme porovnať jednotlivé IS. Pretože X nadobúda len $n + 1$ hodnôt, tak existuje len konečný počet možných hodnôt skutočných pravdepodobností pokrytia. Konkrétne zrejme $(n + 1) + n + \dots + 1 = (1/2)(n + 1)(n + 2)$, ak nepočítame nulu. To je stále ešte dosť veľa, každopádne je rozumné požadovať, aby IS nepokryl parameter p pre pre čo najviac hodnôt X . Teda ak pre pevné p zoradíme podľa veľkosti pravdepodobnosti $P_p[X = x]$, kde $x \in \{0, 1, \dots, n\}$ a sčítame k najmenších tak, aby sme sa priblížili čo najbližšie k hladine významnosti α , dostaneme najbližšiu možnú hladinu významnosti a z tej vypočítame niečo ako optimálnu pravdepodobnosť pokrytia pre dané n a p , ktorú označíme $O_n(p)$.

Teraz už môžeme porovnať rozdiel medzi absolútnou chybou skutočnej pravdepodobnosti pokrytia IS a absolútnou chybou optimálnej pravdepodobnosti. Teda porovnáme $|C_n(p) - (1 - \alpha)| - |O_n(p) - (1 - \alpha)|$ v závislosti od p pre všetky intervaly spoľahlivosti. Grafy pre $n = 75$ vidíme na obrázku 3.7. Záporné hodnoty pri všetkých intervaloch okrem Clopper-Pearsonovho znamenajú, že tieto intervaly niekedy pokrývajú parameter p pre x , ktoré je menej pravdepodobné než iná hodnota x , pre ktorú p nepokrývajú. To by mohlo byť tiež považované za mierny nedostatok, každopádne záporné hodnoty sú veľmi sporadické a oproti kladným odchýlkam sú zrejme zanedbateľné. V porovnaní s obrázkom 3.3 sú tieto grafy prehľadnejšie a celkom dobre vidieť, že Wilsonov interval je na tom najlepšie v krajných oblastiach, čo predtým nebolo dobre porovnateľné. Každopádne to, že väčšina IS ($I_W, I_{AC}, I_{Arc}, I_{VP}$) je temer rovnaká v okolí bodu $p = 1/2$, bolo zrejme už z 3.3.

Určite by bolo zaujímavé mať IS so skutočnou pravdepodobnosťou pokrytia ako je na obrázku 3.6. Teoreticky je možné určiť pre čo najviac $p \in (0, 1)$ množinu M_p s prvkami x , pre ktoré ma platiť $I(x) \ni p$ a potom pre každé $x \in \{0, 1, \dots, n\}$ zahrnúť do IS tie p , pre ktoré $x \in M_p$. Prakticky však vychádza pre hodnoty blízke $X = 0$ a $X = n$ ako riešenie zjednotenie viacerých intervalov a to nie to, čo by sme očakávali od intervalového odhadu.

Porovnanie optimálnej a skutočnej pravdepodobnosti pokrytia

($n = 75$, $\alpha = 0.05$)



Obr. 3.7: Rozdiely chýb skutočnej a optimálnej pravdepodobnosti pokrytia od nominálnej úrovne pre $n = 75$.

Záver

Hlavnou témou tejto práce bolo teoreticky popísať a štatisticky porovnať niektoré metódy konštrukcie intervalových odhadov parametra binomického rozdelenia.

Úplne základným intervalom, ktorý sa objavuje v každej učebnici štatistiky je štandardný alebo aj Waldov interval spoľahlivosti, pretože sa dá odvodiť invertovaním Waldovho štatistického testu, patriaceho medzi testy založené na vierohodnostnej funkcii. V protiklade k jeho oľúbenosti stojí nedostatočná pravdepodobnosť pokrytia, ktorá sa s rastúcim n zlepšuje len veľmi pomaly oproti iným metódam. Na praktické použitie je veľmi nevhodný a ťažko odporúčateľný.

Z testu založeného na vierohodnostnej funkcii vychádza aj Wilsonov interval, ktorý dostaneme invertovaním skórového testu. Skórový interval dáva jednoznačne najlepšie výsledky spomedzi všetkých intervalov skúmaných v tejto práci. Priemerná pravdepodobnosť pokrytia rýchlo konverguje k nominálnej úrovni a drží sa mierne nad ňou. Avšak oproti jednoduchému štandardnému intervalu je o dosť komplikovanejší.

Za konkurenta Wilsonovho intervalu môžeme považovať posledný z trojice intervalov získaných z nejakého testu založeného na vierohodnostnej funkcii. Konkrétne ide o interval z vierohodnostného pomeru a oproti skórovému intervalu je mierne podhodnotený, ale rýchlo sa približuje k nominálnej úrovni. Vďaka tomu je priemerne aj o niečo kratší než Wilsonov interval. Pre väčšinu prípadov ho nie je možné vypočítať analyticky, čo je jeho veľkou nevýhodou.

Agresti-Coullov interval je podstatným zjednodušením Wilsonovho intervalu, ale napriek tomu dosahuje dosť dobré výsledky. Priemerná pravdepodobnosť pokrytia konverguje síce o dosť pomalšie k nominálnej úrovni, ale to je spôsobené tým, že jeho pokrytie v krajných oblastiach intervalu $(0, 1)$ je stabilne nad nominálnou úrovňou a celkovo pod túto úroveň príliš neklesá, čo sa nedá povedať o predchádzajúcich dvoch intervaloch. Je využiteľný v prípadoch, keď chceme mať istotu, že skutočná pravdepodobnosť nebude príliš nízka. Ďalšou veľkou výhodou je jeho jednoduchý tvar, dokonca v prípade 95% IS, je temer totožný s pravidlom „pridať 2 úspechy a 2 neúspechy“ a použiť štandardný prístup. Pre približné odhadnutie binomického podielu snáď nie je lepšej voľby.

Špeciálnym prípadom je Clopper-Pearsonov interval nazývaný aj exaktný interval, pretože jeho skutočná pravdepodobnosť je vždy najmenej $1 - \alpha$, ale vo väčšine prípadoch je príliš nadhodnotený. S tým nutne súvisí aj stredná hodnota jeho šírky, ktorá je zbytočne moc veľká. No môže sa vyskytnúť prípad, kedy požadujeme aby 95% IS mal aspoň 95% pravdepodobnosť pokrytia a vtedy sa hodí práve tento prístup.

Posledným skúmaným intervalom v tejto práci je arkussínusový interval, ktorý však nevyniká v žiadnom ohľade a medzi ostatnými sa stráca. Stojí za ním však zaujímavá teória, ktorá môže byť vhodná aj v inej situácii.

A. Použité vety

Veta A.1 (Cramérova-Raova nerovnosť). *Nech $\mathbf{X} = (X_1, \dots, X_n)^T$ je náhodný výber z rozdelenia, ktoré je prvkom rodiny $\{F_\theta; \theta \in \Theta\}$ spĺňajúcej podmienky regularity. Nech $\phi(\mathbf{X})$ je nestranný odhad parametrickej funkcie $g(\theta)$ taký, že $E \phi^2(\mathbf{X}) < \infty, \theta \in \Theta$.*

Ďalej nech pre všetky $\theta \in \Theta$

(i) existuje derivácia $g'(\theta)$,

(ii) a platí

$$\begin{aligned} & \frac{\partial}{\partial \theta} \int_M \dots \int_M \phi(x_1, \dots, x_n) \prod_{i=1}^n f_\theta(x_i) d\nu(x_1) \dots d\nu(x_n) \\ &= \int_M \dots \int_M \phi(x_1, \dots, x_n) \frac{\partial}{\partial \theta} \prod_{i=1}^n f_\theta(x_i) d\nu(x_1) \dots d\nu(x_n). \end{aligned}$$

Potom pre všetky $\theta \in \Theta$ platí

$$\text{var} [\phi(X_1, \dots, X_n)] \geq \frac{(g'(\theta))^2}{nI(\theta)}.$$

Vid' [5], str. 97, Veta 5.2.

Veta A.2 (Lindebergova centrálna limitná veta). *Nech X_1, X_2, \dots sú nezávislé rovnako rozdelené náhodné veličiny so strednou hodnotou μ a s konečným rozptylom σ^2 . Potom pre $n \rightarrow \infty$ platí*

$$\frac{X_1 + \dots + X_n - n\mu}{\sqrt{n}} \xrightarrow{d} \mathbf{N}(0, \sigma^2).$$

Vid' [3], str. 331, Veta B.5.

Veta A.3 (princíp spojitosti pre konvergenciu podľa distribučnej funkcie). *Ak g je spojitá reálna funkcia a $X_n \xrightarrow{d} X$, potom $g(X_n) \xrightarrow{d} g(X)$.*

Veta A.4 (princíp spojitosti pre konvergenciu podľa pravdepodobnosti). *Ak g je spojitá reálna funkcia a $X_n \xrightarrow{P} X$, potom $g(X_n) \xrightarrow{P} g(X)$.*

Obe vety nájdeme v [3], str. 332-333

Veta A.5 (Chinčinova). *Nech X_1, X_2, \dots je postupnosť nezávislých rovnako rozdelených náhodných veličín s konečnou strednou hodnotou μ . Potom platí*

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow{P} \mu.$$

Vid' [2], str. 76, Veta 6.3.

Veta A.6 (Cramér-Slucky). *Nech X_1, X_2, \dots je postupnosť náhodných veličín s distribučnými funkciami F_1, F_2, \dots . Nech F je distribučná funkcia a $c > 0$ konštanta. Ďalej nech F_n slabo konvergujú k F a Y_1, Y_2, \dots je postupnosť takých náhodných veličín, že $Y_n \xrightarrow{P} c$.*

Potom distribučné funkcie F_n^T náhodnej veličiny $T_n = X_n/Y_n$ konvergujú slabo k $F(cx)$.

Vid' [3], str. 333, Veta B.10.

Veta A.7. *Pre kritické hodnoty F-rozdelenia platí vzťah*

$$F_{m,n}(\alpha) = \frac{1}{F_{n,m}(1 - \alpha)}.$$

Vid' [2], str. 99, Veta 8.4.

B. Kód v programe R

Na tomto mieste sú uvedené zdrojové kódy z programu R. Okrem funkcie na výpočet všetkých intervalov spoľahlivosti sa tu nachádzajú taktiež funkcie na vykreslenie grafov použitých v kapitole 3, v ktorých sa dajú zadať aj iné vstupné parametre, a tak napríklad dostať všetky obrázky pre $\alpha = 0.01$.

```
#####  
#                               Intervaly spoľahlivosti                               #  
#####  
int <- function(x,n,alpha=0.05,typ='all')  
# typ - podľa typu intervalu môže mať nasledujúce hodnoty:  
# 'S' - štandardný interval  
# 'W' - Wilsonov interval  
# 'AC' - Agresti-Coullov interval  
# 'Arc' - arkussínusový interval  
# 'CP' - Clopper-Pearsonov interval  
# 'VP' - interval z vierohodnostného pomeru  
# 'all' - pre vypísanie všetkých intervalov  
{  
  # p.o - bodový odhad parametra p  
  p.o <- x/n  
  # u - kvantil normálneho rozdelenia  
  u <- qnorm(1-alpha/2)  
  if(typ=='S'){  
    pol.sirky <- u * sqrt(p.o*(1 - p.o)/n)  
    d <- p.o - pol.sirky  
    h <- p.o + pol.sirky  
  }  
  else if(typ=='W'){  
    stred <- (x + u^2/2)/(n + u^2)  
    pol.sirky <- (u*sqrt(n)/(n+u^2))*sqrt(p.o*(1-p.o) + u^2/(4*n))  
    d <- stred - pol.sirky  
    h <- stred + pol.sirky  
  }  
  else if(typ=='AC'){  
    p.o <- (x + u^2/2)/(n + u^2)  
    pol.sirky <- u*sqrt(p.o*(1 - p.o)/(n + u^2))  
    d <- p.o - pol.sirky  
    h <- p.o + pol.sirky  
  }  
  else if(typ=='Arc'){  
    d <- sin(max(0, asin(sqrt(p.o)) - u*sqrt(1/(4*n))))^2  
    h <- sin(min(pi/2, asin(sqrt(p.o)) + u*sqrt(1/(4*n))))^2  
  }  
  else if(typ=='CP'){  
    if(x>0) {Fd <- qf(1-alpha/2, 2*x, 2*(n-x+1), lower.tail = FALSE)  
      d <- 1/((n-x+1)/(x*Fd)+1)}
```



```

if(x<n) {Fh <- qf(alpha/2, 2*(x+1), 2*(n-x), lower.tail = FALSE)
  h <- 1/((n-x)/((x+1)*Fh)+1)}
}

else if(typ=='VP'){
LR <- function(p)
  x*log(x/n)+(n-x)*log(1-x/n)-x*log(p)-(n-x)*log(1-p)-u^2/2
if(x==0) h <- 1 - (exp(u^2/2))^(1/n)
else if(x==n) d <- (exp(u^2/2))^(1/n)
  else { d <- uniroot(LR, c(0,x/n), tol = 0.00001)$root
        h <- uniroot(LR, c(x/n,1), tol = 0.00001)$root }
}

# v prípade, že chceme všetky intervaly
if(typ=='all'){
nazvy <- c('S', 'W', 'AC', 'Arc', 'CP', 'VP')
int <- data.frame()
for(i in 1:6) int <- rbind(int,int(x,n,alpha,nazvy[i]))
rownames(int) <- nazvy
} else {
# V extrémnych prípadoch x = 0 resp. x = n,
# vezmeme za dolný odhad d = 0 resp. za horný odhad h = 1
if(x==0) d <- 0 else if(x==n) h <- n
# výsledok ešte upravíme tak, aby IS bol podmnožinou [0,1]
int <- data.frame('d' = max(0,d), 'h' = min(h,1))
}
int
}

#####
# Funkcia na vypočítanie skutočných pravdepodobností pokrytia #
# prípadne stredných hodnôt šírky intervalu pre "všetky" p. #
#####
hodnoty <- function(n,typ,co='prst',alpha=0.05,nkrok=10000)
# co='prst' pre pravdepodobnosti pokrytia
# co='sirka' pre stredné hodnoty šírky
{
# vypočítame IS pre všetky x = 0,1,...,n
intervaly <- data.frame()
for(i in 0:n) intervaly <- rbind(intervaly,int(i,n,alpha,typ))
# pre všetky intervaly dopočítame aj šírku
intervaly$sirka <- intervaly$h - intervaly$d
# výsledky budeme ukladať do vektora hodnoty
hodnoty <- rep(0,nkrok-1)

# prejdeme celý interval (0,1) s krokom 1/nkrok
for(i in 1:(nkrok-1)){
p <- i/nkrok
# vypočítame pravdepodobnosti P_p[X=0,1,...,n]
prsti <- dbinom(0:n,n,p)

```

```

if(co=='prst')
  # skutočná pravdepodobnosť pokrytia pre dané p
  hodnoty[i] <- sum((intervaly$d < p)*(intervaly$h > p)*prsti)
else if(co=='sirka')
  # stredná hodnota šírky intervalu spoľahlivosti pre dané p
  hodnoty[i] <- sum(intervaly$sirka*prsti)
}
hodnoty
}
#####
# Funkcia na zostrojenie grafov skutočnej pravdepodobnosti #
# pokrytia v závislosti od p pre všetky IS a dané n. #
#####
grafy.prst.p <- function(n,alpha=0.05,nkrok=10000)
{
  # vytvoríme tabuľku data so stĺpcom p obsahujúcim
  # hodnoty z intervalu (0,1) s krokom 1/nkrok
  data <- data.frame(p = 1:(nkrok-1)/nkrok)
  # pre každý interval pridáme stĺpec s pravdepodobnosťami pokrytia
  typ <- c('S','W','AC','Arc','CP','VP')
  for(i in 1:6) data[,i+1] <- hodnoty(n,typ[i],'prst',alpha,nkrok)
  # uložíme do súboru (grafy.co.n.alfa) vo formáte pdf
  pdf(paste0('grafy_prst_p_',n,'.pdf'),
      width=6, height=9, encoding='CP1250')
  # vykreslíme v 2 riadkoch po 3 grafy a navyše
  # pomocou malého triku vytvoríme prázdne miesto na nadpis :)
  layout(matrix(c(0,0:6), 4, 2, byrow=T),heights=c(1,4,4,4))
  # zmenšený spodný okraj
  par(mar=c(2.5, 3, 4, 2))
  nazov <- c('Štandardný IS','Wilsonov IS','Agresti-Coullov IS',
            'Arkussínusový IS','Clopper-Pearsonov IS',
'IS z vierohodnostného pomeru')
  for(i in 1:6){
    plot(data$p,data[,i+1], ylim=c(1-3*alpha,1), xlim=c(0,1),
         type='l', xlab='', ylab='', axes=F, lwd=0.8)
    title(main=paste(nazov[i]))
    # vodorovná čiara
    abline(h=1-alpha, col='black', lty=2, lwd=2)
    # osi zadefinujeme zvlášť
    axis(2,at=seq(0,1, by=alpha), las=2)
    axis(1,at=seq(0,1, by=0.2))
    # názov osi x
    mtext('p',side=1,adj=0.5,line=1.3,cex=0.8)
    # názov osi y
    mtext(expression('C' [n] (p)),side=3,adj=-0.2,line=0.6,cex=0.75)
  }
  nadpis <- 'Pravdepodobnosť pokrytia'
  mtext(nadpis, side=3, outer=T, line=-2)
}

```

```

podnadpis <- paste0('(n = ',n,', alfa = ',alpha,')')
mtext(podnadpis, side=3, outer=T, line=-4, cex=0.9)
dev.off()
}
#####
# Funkcia na zostrojenie grafov priemernej pravdepodobnosti #
# pokrytia v závislosti od n pre zadané IS. #
# !Výpočet môže trvať niekoľko minút! #
#####
grafy.prst.n <-
  function(min.n, max.n, typy='all', alpha=0.05, nkrok=10000)
# typy je vektor s typmi intervalov, pre ktoré chceme vykresliť
# grafy, prípadne typy='all' pre všetky IS.
{
  if(typy[1]=='all')
    typ <- c('S','W','AC','Arc','CP','VP') else typ <- typy
  n <- seq(min.n, max.n, by=max(round((max.n-min.n)/50),1))
  l.t <- length(typ)
  l.n <- length(n)
  data <- matrix(nrow=l.n,ncol=l.t)
  k <- 0
  for(i in 1:l.t) for(j in 1:l.n){
    # Približné počítadlo stavu výpočtu
    k <- k+1
    print(round(k*100/(l.t*l.n),2))
    data[j,i] <- mean(hodnoty(n[j], typ[i], 'prst', alpha, nkrok))
  }
  if(l.t <= 4) col <- rep('black',l.t) else
    col <- c('black','orange','brown','darkblue','red','darkgreen')
  pdf(paste0('grafy_prst_n_',min.n,'_',max.n,'.pdf'),
      width=6, height=5, encoding='CP1250')
  par(mar=c(2.5, 4, 4, 2))
  plot(1:max.n, 1:max.n, ylim=c(1-2*alpha,1-alpha/2),
      type='n', xlab='', ylab='', axes=F)
  title(main='Priemerná pravdepodobnosť pokrytia')
  mtext(paste0('(alfa = ',alpha,')'))
  for(i in 1:l.t)
    lines(n, data[,i], col=col[i], type='l', lty=i, lwd=1.5)
  abline(h=1-alpha, col='black', lty=2, lwd=2)
  axis(2,at=seq(0,1, by=alpha/2), las=2)
  axis(1,at=seq(0,max.n, by=max.n/5))
  mtext('n',side=1,adj=1.02,line=-0.1,cex=0.9)
  mtext('C(n)',side=3,adj=-0.1,line=0.6,cex=0.85)
  legend(4*max.n/5, 1-1.45*alpha, typ,cex=0.9,col=col,lty=1:l.t)
  dev.off()
}

```

```
#####
# Funkcia na zostrojenie grafov strednej hodnoty šírky intervalov #
# spoľahlivosti v závislosti od p pre zadané IS a konkrétne n. #
#####
grafy.sirka.p <- function(n,typy='all',alpha=0.05,nkrok=10000)
# typy je vektor s typmi intervalov, pre ktoré chceme vykresliť
# stredné hodnoty šírky, prípadne typy='all' pre všetky IS.
{
  data <- data.frame(p = 1:(nkrok-1)/nkrok)
  if(typy[1]=='all') typ <- c('S','W','AC','Arc','CP','VP')
  else typ <- typy
  l <- length(typ)
  for(i in 1:l) data[,i+1] <- hodnoty(n,typ[i],'sirka',alpha,nkrok)
  pdf(paste0('grafy_sirka_p_',n,'.pdf'),
      width=6, height=5, encoding='CP1250')
  par(mar=c(2.5, 4, 4, 2))
  if(l <= 4) col <- rep('black',l) else
    col <- c('black','orange','brown','darkblue','red','darkgreen')
  plot(data$p, data[,2], ylim=c(0.1,max(data[, -1])+0.01, xlim=c(0,1),
      type='l',lwd=1.5, col=col[1], xlab='', ylab='', axes=F)
  if(l>1) for(i in 2:l)
    lines(data$p,data[,i+1],col=col[i],type='l',lty=i,lwd=1.5)
  title(main='Stredná hodnota šírky IS')
  mtext(paste0('(n = ',n,', alfa = ',alpha,')'))
  axis(2,at=seq(0,1, by=0.05), las=2)
  axis(1,at=seq(0,1, by=0.2))
  mtext('p',side=1,adj=1.02,line=-0.1,cex=0.9)
  #mtext(expression('C'[n](p)),side=3,adj=-0.15,line=0.6,cex=0.75)
  legend(0.405,(max(data[, -1])+0.1)/2,typ,cex=0.9,col=col,lty=1:1)
  dev.off()
}
#####
# Optimálna pravdepodobnosť pokrytia v závislosti od p. #
#####
optimal <- function(n, alpha=0.05, nkrok=10000)
{
  optimal <- rep(0,nkrok-1)
  for(i in 1:(nkrok-1)){
    p <- i/nkrok
    # Zoradené pravdepodobnosti P[X=x]
    prsti <- sort(dbinom(0:n,n,p))
    sucty <- rep(0,n+1)
    # Čiastočné súčty j najmenších pravdepodobností
    for(j in 1:(n+1)) sucty[j] <- sum(prsti[1:j])
    optimal[i] <- 1-sucty[which.min(abs(sucty-alpha))]
  }
  optimal
}

```

```
#####
#      Funkcia na zostrojenie grafov závislosti rozdielu chýb      #
#      skutočnej a optimálnej pravdepodobnosti pokrytia.        #
#####
grafy.opti <- function(n, alpha=0.05, nkrok=10000)
{
  data <- data.frame(p = 1:(nkrok-1)/nkrok)
  typ <- c('S','W','AC','Arc','CP','VP')
  o <- optimal(n, alpha, nkrok)
  for(i in 1:6){
    data[,i+1] <- hodnoty(n,typ[i],'prst',alpha,nkrok)
    data[,i+1] <- abs(data[,i+1]-1+alpha)-abs(o-1+alpha)
  }
  nazov <- c('Štandardný IS','Wilsonov IS','Agresti-Coullov IS',
            'Arkussínusový IS','Clopper-Pearsonov IS',
            'IS z vierohodnostného pomeru')
  pdf(paste0('grafy_prst_opti_',n,'.pdf'), width=6, height=8.5,
      encoding='CP1250')
  layout(matrix(c(0,0:6), 4, 2, byrow=T),heights=c(1,4,4,4))
  par(mar=c(2.5, 4, 4, 2))
  for(i in 1:6){
    plot(data$p,data[,i+1], ylim=c(-alpha/5,alpha), xlim=c(0,1),
         type='h', col='gray', xlab='', ylab='', axes=F)
    title(main=paste(nazov[i]))
    abline(h=0, col='black', lty=2, lwd=2)
    axis(2,at=seq(-alpha/5,alpha, by=alpha/5), las=2)
    axis(1,at=seq(0,1, by=0.2))
    mtext('p',side=1,adj=1.07,line=-0.1,cex=0.8)
  }
  nad<-'Porovnanie optimálnej a skutočnej pravdepodobnosti pokrytia'
  mtext(nad, side=3, outer=T, line=-2)
  podnadpis <- paste0('(n = ',n,', alfa = ',alpha,')')
  mtext(podnadpis, side=3, outer=T, line=-4, cex=0.9)
  dev.off()
}

```

Zoznam použitej literatúry

- [1] AGRESTI, A., COULL, B. A. *Approximate is better than "exact" for interval estimation of binomial proportions*. The American Statistician 52, 119-126, 1998.
- [2] ANDĚL, J. *Statistické metody* 3. vydání. Praha: Matfyzpress, 2003. ISBN 80-86732-08-8
- [3] ANDĚL, J. *Základy matematické statistiky* 2. vydání. Praha: Matfyzpress, 2007. ISBN 80-7378-001-1
- [4] BROWN, L. D., CAI, T. T., DASGUPTA, A. *Interval estimation for a binomial proportion*. Statistical Science Vol. 16, No. 2, 101-133, 2001.
- [5] DUPAČ, V., HUŠKOVÁ, M. *Pravděpodobnost a matematická statistika* 1. vydání. Praha: Nakladatelství Karolinum, 2009. ISBN 978-80-246-0009-3.

Zoznam obrázkov

3.1	Porovnanie skutočnej pravdepodobnosti pokrytia 95% intervalov spoľahlivosti pre $n = 15$	16
3.2	Porovnanie skutočnej pravdepodobnosti pokrytia 95% intervalov spoľahlivosti pre $n = 40$	17
3.3	Porovnanie skutočnej pravdepodobnosti pokrytia 95% intervalov spoľahlivosti pre $n = 75$	18
3.4	Graf závislosti priemernej pravdepodobnosti pokrytia 95% intervalov spoľahlivosti od n	19
3.5	Graf závislosti strednej hodnoty šírky 95% intervalov spoľahlivosti od parametra p pre $n = 50$	20
3.6	Závislosť optimálnej pravdepodobnosti pokrytia od parametra p pre $n = 75$	21
3.7	Rozdiely chýb skutočnej a optimálnej pravdepodobnosti pokrytia od nominálnej úrovne pre $n = 75$	22

Zoznam tabuliek

3.1	95% IS pre $X = 2, n = 5$	14
3.2	Priemerná pravdepodobnosť pokrytia pre $n = 75$	19
3.3	Priemerné šírky Wilsonovho IS a IS z vierohodnostného pomeru.	20