

Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE



Alena Babiaková

Limitní chování mohutnosti průniků nezávislých výběrů z konečné populace

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: Prof. RNDr. Josef Štěpán, DrSc.

Studijní program: Matematika

Studijní obor: Pravděpodobnost, matematická
statistika a ekonometrie

Studijní plán: Ekonometrie

2008

Na tomto místě bych chtěla upřímně poděkovat svému vedoucímu, Prof. RNDr. Josefu Štěpánovi, DrSc., za věnovaný čas, ochotu, odborné rady a příjemnou spolupráci. Také bych chtěla vyjádřit svůj vděk rodině a blízkým za podporu a zázemí.

Prohlašuji, že jsem svou diplomovou práci napsala samostatně a výhradně s použitím citovaných pramenů.

Souhlasím se zapůjčováním práce.

V Praze dne 1. srpna 2008

Alena Babiaková

Obsah

1	ZÁKLADNÉ POJMY A VLASTNOSTI	5
2	PRAVDEPODOBNOSTNÉ ROZDELENIE	8
2.1	EXAKTNÉ ROZDELENIE	8
2.2	MOMENTY A ICH KONVERGENCIA	15
3	ASYMPTOTICKÉ ROZDELENIE	20
3.1	KONVERGENCIA K POISSONOVEMU ROZDELENIU	20
3.2	KONVERGENCIA K NORMÁLNEMU ROZDELENIU	22
3.3	SIMULÁCIA	29
A	APPENDIX	37
A.1	POUŽITÉ TVRDENIA	37
A.2	ZDROJOVÝ KÓD	44
	LITERATÚRA	46

NÁZEV PRÁCE: Limitní chování mohutnosti průniků nezávislých výběrů z konečné populace

AUTOR: Alena Babiaková

KATEDRA (ÚSTAV): Katedra pravděpodobnosti a matematické statistiky

VEDOUcí DIPLOMOVÉ PRÁCE: Prof. RNDr. Josef Štěpán, DrSc.

E-MAIL VEDOUCÍHO: Josef.Stepan@mff.cuni.cz

ABSTRAKT: Hlavním cílem předložené práce je odvození vlastností náhodné veličiny, která představuje mohutnost průniku nezávislých výběrů (bez vracení) z konečné populace. Kromě základních vlastností, jako je například exaktní pravděpodobnostní rozdělení, centrální a faktoriální momenty, také studujeme konvergenci rozdělení (za daných podmínek) k Poissonovu a normálnímu rozdělení. Asymptotické vlastnosti se ukazují být užitečné, protože s exaktním rozdělením se dosti obtížně pracuje. Uvádíme také simulaci, která má za cíl vyšetřit vhodnost aproximace Poissonovým a normálním rozdělením.

KLÍČOVÁ SLOVA: průniky nezávislých náhodných výběrů, konečná populace, asymptotické rozdělení

TITLE: The asymptotic behaviour of the cardinality of intersections of independent samples from a finite population

AUTHOR: Alena Babiaková

DEPARTMENT: Department of Probability and Mathematical Statistics

SUPERVISOR: Prof. RNDr. Josef Štěpán, DrSc.

SUPERVISOR'S E-MAIL ADDRESS: Josef.Stepan@mff.cuni.cz

ABSTRACT: The main aim of the presented thesis is to derive properties of the random variable representing the cardinality of intersection of independent random samples (without replacement) from a finite population. Besides basic properties, such as exact probability distribution, central and factorial moments, we also study convergence of the moments and convergence of the probability distribution (under certain conditions) to Poisson and normal distribution. The asymptotic properties appear to be useful because the exact distribution is rather difficult to deal with. A simulation study designed to investigate the accuracy of Poisson and normal approximations to the exact probability distribution is also presented.

KEYWORDS: intersections of independent random samples, finite population, asymptotic distribution

1

ZÁKLADNÉ POJMY A VLASTNOSTI

Uvažujme konečnú populáciu pozostávajúcu z N prvkov, ktoré očísľujeme $1, 2, \dots, N$. Ide teda o konečnú množinu

$$\{1, 2, \dots, N\}.$$

Z tejto populácie robíme výbery bez vracania o rozsahu n , $0 < n < N$, pričom výberovú procedúru opakujeme k -krát, $k \in \mathbb{N}$. Podmnožinu prvkov množiny $\{1, 2, \dots, N\}$ vybranú v j -tom kroku výberovej procedúry označíme M_j , $j = 1, 2, \dots, k$. Výbery M_1, M_2, \dots, M_k sú nezávislé v zmysle nezávislosti indikátorov $(I_{M_1}, I_{M_2}, \dots, I_{M_k})$. Nás bude zaujímať mohutnosť množiny

$$\bigcap_{j=1}^k M_j,$$

tj. počet prvkov populácie, ktoré boli vytiahnuté v každom z k krokov. Definujme pre každé $k \in \mathbb{N}$ náhodnú veličinu

$$C_k = \text{card} \left\{ \bigcap_{j=1}^k M_j \right\}.$$

Ak sa na $k \in \mathbb{N}$ pozeráme ako na časový index, potom

$$C_1, C_2, \dots, C_k, \dots$$

je náhodná postupnosť s množinou stavov

$$S = \{0, 1, 2, \dots, n\}.$$

Zrejme je

$$C_1 = n. \quad (1.1)$$

Taktiež je zrejmé, že pre všetky $c_2, \dots, c_{k-1}, c_k \in S$ platí (ak je nasledujúca podmienená pravdepodobnosť definovaná)

$$\begin{aligned} & \text{P} [C_k = c_k | C_{k-1} = c_{k-1}, C_{k-2} = c_{k-2}, \dots, C_2 = c_2, C_1 = n] \\ &= \text{P} [C_k = c_k | C_{k-1} = c_{k-1}] \\ &= \binom{c_{k-1}}{c_k} \binom{N - c_{k-1}}{n - c_k} / \binom{N}{n}. \end{aligned}$$

Vidíme, že náhodný proces $\{C_k, k \in \mathbb{N}\}$ spĺňa markovskú vlastnosť a podmienené rozdelenie $\mathcal{L}(C_k | C_{k-1} = c_{k-1})$ je rovné hypergeometrickému rozdeleniu $HG(N, n, c_{k-1})$. Je to situácia, kedy uskutočňujeme výber o rozsahu n z populácie o veľkosti N , v ktorej má c_{k-1} prvkov tú vlastnosť, že je prvkom prieniku $\bigcap_{j=1}^{k-1} M_j$. Hypergeometrické rozdelenie popisuje pravdepodobnosť, že práve c_k z vybraných prvkov má danú vlastnosť, tj. c_k prvkov patrí zároveň do M_k aj do $\bigcap_{j=1}^{k-1} M_j$

$$c_k = \text{card} \left\{ \bigcap_{j=1}^{k-1} M_j \cap M_k \right\} = \text{card} \left\{ \bigcap_{j=1}^k M_j \right\} = C_k.$$

Z predchádzajúceho vieme, že $\{C_k, k \in \mathbb{N}\}$ je časovo homogénny Markovský reťazec s maticou pravdepodobností prechodu $P = (p_{ij})_{i,j \in S}$, kde

$$p_{ij} = \begin{cases} \text{P} [C_k = j | C_{k-1} = i] = \binom{i}{j} \binom{N-i}{n-j} / \binom{N}{n} & j = 0, 1, 2, \dots, i, \\ 0 & j = i + 1, \dots, n, \end{cases}$$

a počiatočným rozdelením $\{p_i\}_{i \in S}$

$$\begin{aligned} p_n &= 1 \\ p_i &= 0 \quad \text{pre } i = 0, 1, \dots, n-1. \end{aligned}$$

Ďalej nás zaujíma transformácia C_k taká, aby sme dostali martingal. Nech T_k je nejaká transformácia C_k , ktorá je martingal. Uvažujme mocninnú transformáciu $T_k = d^k C_k$, $k \in \mathbb{N}$, pre nejaké $d \in \mathbb{R}$. Potom ako dôsledok markovskej vlastnosti dostaneme

$$\text{E} [T_k | \mathcal{F}_{k-1}] = \text{E} [T_k | T_0, T_1, \dots, T_{k-1}] = \text{E} [T_k | T_{k-1}].$$

Na to, aby proces $\{T_k, k \in \mathbb{N}\}$ bol martingal, musí byť splnené

$$\text{E} [T_k | T_{k-1} = c] = c, \quad k = 2, 3, \dots$$

Postupnými úpravami a s využitím toho, že pre strednú hodnotu náhodnej veličiny X s hypergeometrickým rozdelením $HG(N, n, c)$ platí $E X = c \frac{n}{N}$, dostaneme

$$\begin{aligned} E [d^k C_k | d^{k-1} C_{k-1} = c] &= c \\ d^k E [C_k | C_{k-1} = c d^{-(k-1)}] &= c \\ d^k c d^{-(k-1)} \frac{n}{N} &= c \\ d &= \frac{N}{n}. \end{aligned}$$

Potom

$$T_k = \left(\frac{N}{n}\right)^k C_k \tag{1.2}$$

a náhodný proces $\{T_k, k \in \mathbb{N}\}$ je martingal.

2

PRAVDEPODOBNOSTNÉ ROZDELENIE

2.1 EXAKTNÉ ROZDELENIE

Od tejto chvíle budeme v tejto práci náhodnú veličinu C_k podrobnejšie značiť $C_{N,n,k}$. Exaktné rozdelenie $C_{N,n,k}$ je priamočiarym dôsledkom nasledujúcej vety:

VETA 2.1 (JORDANOVA IDENTITA). *Nech A_1, A_2, \dots, A_N sú náhodné javy na rovnakom pravdepodobnostnom priestore. Potom pravdepodobnosť toho, že nastalo práve r javov spomedzi A_1, A_2, \dots, A_N , je rovná*

$$W_{N,r} = \sum_{j=0}^{N-r} (-1)^j \binom{r+j}{j} S_{r+j},$$

kde

$$\begin{aligned} S_k &= \sum_{1 \leq i_1 < \dots < i_k \leq N} \mathbb{P} [A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}] \quad \text{pre } k = 1, 2, \dots, N, \\ S_0 &= 1. \end{aligned}$$

Dôkaz. Označme pravdepodobnostný priestor, na ktorom uvažujeme náhodné javy A_1, A_2, \dots, A_N , ako $(\Omega, \mathcal{A}, \mathbb{P})$. Ďalej definujme indexovú množinu

$$U = \{1, 2, \dots, 2^N - 1, 2^N\}.$$

Uvažujme disjunktný rozklad $\{E_1, E_2, \dots, E_{2^N}\}$ priestoru elementárnych javov Ω , tj.

$$\Omega = \bigcup_{j \in U} E_j,$$

kde E_j sú náhodné javy tvaru

$$A_1^{\delta_1} \cap A_2^{\delta_2} \cap \dots \cap A_N^{\delta_N} \quad \text{pričom} \quad A_s^{\delta_s} \text{ je buď } A_s \text{ alebo } A_s^c.$$

Poradia hodnôt, ktoré nadobúdajú indexy $\delta_1, \delta_2, \dots, \delta_N$, uvažujeme navzájom rôzne, vyčerpávajú teda všetkých 2^N možností. Nech r_i označuje počet javov A_1, A_2, \dots, A_N , ktoré nastanú, ak nastane E_i . Pre každé $i \in U$ definujeme nejakú permutáciu ℓ_i množiny $\{1, 2, \dots, N\}$ takú, že

$$E_i = A_{\ell_i(1)} \cap A_{\ell_i(2)} \cap \dots \cap A_{\ell_i(r_i)} \cap A_{\ell_i(r_i+1)}^c \cap \dots \cap A_{\ell_i(N)}^c,$$

pričom $\ell_i(1), \dots, \ell_i(r_i)$ sú indexy tých javov z A_1, \dots, A_N , ktoré nastali, keď nastal jav E_i . Zrejme náhodný jav, že nastalo práve r javov z A_1, A_2, \dots, A_N , je disjunktným zjednotením práve $\binom{N}{r}$ náhodných javov E_i takých, že $r = r_i$, a jeho pravdepodobnosť je rovná

$$W_{N,r} = \sum_{i \in \Lambda} P[E_i], \quad \text{kde } \Lambda = \{i \in U : r = r_i\}. \quad (2.1)$$

Teraz nás budú zaujímať pravdepodobnosti typu

$$P[A_{j_1} \cap A_{j_2} \cap \dots \cap A_{j_k}].$$

Keďže náhodný jav

$$A_{j_1} \cap A_{j_2} \cap \dots \cap A_{j_k}$$

je disjunktným zjednotením 2^{N-k} javov E_i takých, že $k \leq r_i$ a

$$\{j_1, j_2, \dots, j_k\} \subseteq \{\ell_i(1), \ell_i(2), \dots, \ell_i(r_i)\}, \quad (2.2)$$

pravdepodobnosť

$$P[A_{j_1} \cap A_{j_2} \cap \dots \cap A_{j_k}]$$

je súčtom pravdepodobností týchto 2^{N-k} javov E_i . Pre S_k , $k = 1, 2, \dots, N$, platí

$$\begin{aligned} S_k &= \sum_{1 \leq i_1 < \dots < i_k \leq N} P[A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}] \\ &= \sum_{i \in U} \binom{r_i}{k} P[E_i] \end{aligned}$$

vzhľadom k tomu, že inklúzia (2.2) je pre E_i splnená v práve $\binom{r_i}{k}$ prípadoch. Takže

$$\begin{aligned} \sum_{j=0}^{N-r} (-1)^j \binom{r+j}{j} S_{r+j} &= \sum_{i \in U} \left\{ \sum_{j=0}^{N-r} (-1)^j \binom{r+j}{j} \binom{r_i}{r+j} \right\} P[E_i] \\ &= \sum_{i \in U} c_i P[E_i]. \end{aligned}$$

Zrejme je $c_i = 0$ pre $r_i < r$ a pre $r_i = r$ (tj. $i \in \Lambda$) je $c_i = 1$. S využitím triviálnej kombinatorickej identity

$$\binom{r+j}{j} \binom{r_i}{r+j} = \binom{r_i}{r} \binom{r_i-r}{j} \quad (2.3)$$

a binomickej vety je pre $r_i > r$

$$\begin{aligned} c_i &= \sum_{j=0}^{N-r} (-1)^j \binom{r+j}{j} \binom{r_i}{r+j} \\ &= \binom{r_i}{r} \sum_{j=0}^{r_i-r} (-1)^j \binom{r_i-r}{j} = \binom{r_i}{r} (1-1)^{r_i-r} = 0. \end{aligned}$$

Teda na koniec podľa (2.1) dostávame

$$\sum_{j=0}^{N-r} (-1)^j \binom{r+j}{j} S_{r+j} = \sum_{i \in \Lambda} P[E_i] = W_{N,r}.$$

Nami uvedený dôkaz je založený na dôkaze z (Johnson, 1980). Q.E.D.

VETA 2.2 (EXAKTNÉ ROZDELENIE). *Pre pravdepodobnostné rozdelenie náhodnej veličiny $C_{N,n,k}$ platí*

$$P(C_{N,n,k} = r) = \begin{cases} \sum_{j=0}^{n-r} (-1)^j \binom{r+j}{j} \binom{N}{r+j} \left[\frac{\binom{N-r-j}{n-r-j}}{\binom{N}{n}} \right]^k & r = 0, 1, \dots, n, \\ 0 & \text{inak.} \end{cases}$$

Dôkaz. Náhodnú veličinu $C_{N,n,k}$ možno zapísať v tvare

$$\begin{aligned} C_{N,n,k} &= \sum_{i=1}^N I_i, \quad \text{kde} \\ I_i &= I_{\bigcap_{j=1}^k M_j}(i) = \begin{cases} 1 & , \text{ ak } i \in \bigcap_{j=1}^k M_j, \\ 0 & , \text{ ak } i \notin \bigcap_{j=1}^k M_j. \end{cases} \end{aligned} \quad (2.4)$$

Teda náhodný jav $[C_{N,n,k} = r]$ je ekvivalentný s tým, že nastalo práve r z náhodných javov $[I_1 = 1], [I_2 = 1], \dots, [I_N = 1]$. Ďalej pre pravdepodobnosť toho, že ľubovoľné rôzne prvky $1 \leq i_1, i_2, \dots, i_j \leq N$ patria do každého z k výberov, platí (vzhľadom k tomu, že výbery sú navzájom nezávislé)

$$P [[I_{i_1} = 1] \cap [I_{i_2} = 1] \cap \dots \cap [I_{i_j} = 1]] = \begin{cases} \left[\frac{\binom{N-j}{n-j}}{\binom{N}{n}} \right]^k & j = 1, 2, \dots, n, \\ 0 & \text{inak.} \end{cases}$$

A teda pre $S_j, j = 1, \dots, N$, určené rovnosťou

$$S_j = \sum_{1 \leq i_1 < \dots < i_j \leq N} P [[I_{i_1} = 1] \cap [I_{i_2} = 1] \cap \dots \cap [I_{i_j} = 1]]$$

platí

$$S_j = \begin{cases} \binom{N}{j} \left[\frac{\binom{N-j}{n-j}}{\binom{N}{n}} \right]^k & j = 1, 2, \dots, n, \\ 0 & \text{inak.} \end{cases}$$

Na koniec aplikovaním vety 2.1 a využitím faktu, že $S_{r+j} = 0$ pre $j > n - r$, dostaneme tvrdenie dokazovanej vety. Q.E.D.

POZNÁMKA. Ak opäť použijeme rovnosť (2.3) z dôkazu vety 2.1 v tvare

$$\binom{r+j}{j} \binom{N}{r+j} = \binom{N}{r} \binom{N-r}{j},$$

môžeme exaktné rozdelenie $C_{N,n,k}$ prepísať

$$P (C_{N,n,k} = r) = \begin{cases} \binom{N}{r} \binom{N}{n}^{-k} \sum_{j=0}^{n-r} (-1)^j \binom{N-r}{j} \binom{N-r-j}{n-r-j}^k & r = 0, 1, \dots, n, \\ 0 & \text{inak.} \end{cases}$$

PRÍKLAD 2.3 (“SPRAVODLIVÉ” LOSOVANIE). Bolo vypísané výberové konanie na práce v štyroch lokalitách. Výberového konania sa zúčastnilo 16 firiem. Pre každú lokalitu bolo spomedzi týchto 16 firiem vylosovaných 5, ktoré postúpili do ďalšieho kola výberového konania a ich ponuky boli ďalej uvažované. Výsledky losovania sú uvedené v tabuľke 2.1, hodnota 1 predstavuje úspech a hodnota 0 predstavuje neúspech pri losovaní.

Už pri prvom pohľade na tabuľku s výsledkami losovania sa objavujú pochybnosti o jeho spravodlivosti, keďže 9 firiem nebolo vytiahnutých ani v jednom zo 4 kôl, naproti čomu 2 firmy boli vytiahnuté v každom kole a 3

firmy boli vytiahnuté v 3 kolách.

V tomto výberovom konaní ide vlastne o postupnosť $k = 4$ výberov o rozsahu $n = 5$ spomedzi $N = 16$ firiem. S použitím znalosti exaktného rozdelenia náhodnej veličiny $C_{N,n,k}$ sme schopní vypočítať pravdepodobnosti, že práve r firiem, $r = 0, 1, \dots, 5$, bolo úspešných v každom kole losovania. Pravdepodobnosti sú uvedené v tabuľke 2.2. Z tejto tabuľky vidíme, že výrazne najpravdepodobnejší je výsledok, že žiadna z firiem nebude vytiahnutá v každom kole, tento výsledok má pravdepodobnosť 85%. Taktiež vidíme, že výsledok, ktorý podľa tabuľky 2.1 naozaj nastal, má pravdepodobnosť fakticky nulovú.

lokalita		1	2	3	4
firma	STAEG, s.r.o.	0	0	0	0
	ASP služby, s.r.o.	0	0	0	0
	AVE CZ odpad. hosp. s.r.o	1	0	1	1
	PARK, v.o.s	1	1	0	1
	Ing. M. Krčál	0	1	1	0
	TS Hustopeče, s.r.o.	1	0	1	1
	PET group, a.s.	0	0	0	0
	DVOŘÁK comte, a.s.	0	0	0	0
	FALCON, s.r.o.	1	1	1	1
	TS Brno, s.r.o.	0	0	0	0
	A.S.A. Žabrovresky, s.r.o.	0	0	0	0
	A.S.A., s.r.o.	0	0	0	0
	A.S.A. EKO Znojmo, s.r.o.	0	0	0	0
	TS A.S.A., s.r.o.	0	0	0	0
	INGE BRNO, s.r.o.	0	1	0	0
	ŠIMEK 96, s.r.o.	1	1	1	1

Tabuľka 2.1: Výsledky výberového konania.

Ďalej nás môže zaujímať počet firiem, ktoré sa nedostali ani do jedného zo 4 výberov. Ide o podobnú úlohu, ako bola predchádzajúca, avšak s tým rozdielom, že jednotlivé výbery reprezentujú firmy, ktoré v danom regióne (v danom kole) vybrané neboli, teda $n = 11$. Opäť použijeme exaktné rozdelenie a takto určíme pravdepodobnosti, že práve r firiem sa nedostane do žiadneho z výberov pre $r = 0, 1, \dots, 11$. Výsledky uvádzame v tabuľke 2.3. Tieto výsledky nás opäť utvrďujú v podozrení, že výberové konanie nebolo

uskutočňované deklarovanou metódou: náhodný jav, že 9 firiem nebolo vytiahnutých v ani jednom z kôl, je prakticky nemožný, jeho pravdepodobnosť je v ráde milióntín. Vidíme, že najpravdepodobnejšie je to, že do žiadneho z výberov by sa nedostali 4 firmy.

O spravodlivosti losovania by sa dalo rozhodnúť pomocou exaktného štatistického testu. Test hypotézy o korektnosti losovania je možné odvodiť analogickým postupom ako je uvedený v bakalárskej práci (Valášková, 2007).

r	pravdepodobnosť
0	0.853142307653257
1	0.141184229178218
2	0.005616831695838
3	0.000056527847638
4	0.000000103613049
5	0.000000000011999

Tabuľka 2.2: Tabuľka pravdepodobností toho, že práve r firiem sa vyskytne v každom výbere.

r	pravdepodobnosť
0	0.002102587260818
1	0.030262865297274
2	0.144048681206758
3	0.302513290072231
4	0.310913549712386
5	0.162007098085593
6	0.042496105812935
7	0.005355996074879
8	0.000293970677982
9	0.000005827409049
10	0.000000028378096
11	0.000000000011999

Tabuľka 2.3: Tabuľka pravdepodobností toho, že práve r firiem sa nedostane do žiadneho z výberov.

Pozerajme sa zase na chvíľu na $C_{N,n,k}$ ako na náhodnú postupnosť indexovanú k . Keď vezmeme do úvahy postup, akým sú hodnoty $C_{N,n,k}$ generované, je zrejmé, že tieto hodnoty tvoria (v čase) nerastúcu postupnosť. Uvažujme prvý časový okamih, kedy je prienikom výberov prázdna množina. V tomto momente sa pre nás proces zastaví (stav 0 je absorbčný stav). Definujme markovský čas prvého vstupu $C_{N,n,k}$ do nuly

$$\tau = \min \{s \in \mathbb{N} : C_{N,n,s} = 0\}.$$

VETA 2.4. *Pre rozdelenie náhodnej veličiny τ platí*

$$P[\tau = s] = \sum_{j=1}^n (-1)^j \binom{N}{j} \left(\frac{\binom{N-j}{n-j}}{\binom{N}{n}} \right)^{s-1} \left(\frac{\binom{N-j}{n-j}}{\binom{N}{n}} - 1 \right), \text{ kde } s = 2, 3, \dots,$$

a stredná hodnota τ je rovná

$$E\tau = 1 + \sum_{j=1}^n (-1)^{j+1} \frac{\binom{N}{j} \binom{N-j}{n-j}}{\binom{N}{n} - \binom{N-j}{n-j}}. \quad (2.5)$$

Dôkaz. Najprv odvodíme rozdelenie τ pre $s = 2, 3, \dots$ s využitím vety 2.2

$$\begin{aligned} P[\tau = s] &= P[\tau \leq s, \tau > s - 1] = P[\tau \leq s] - P[\tau \leq s - 1] \\ &= P[C_{N,n,s} = 0] - P[C_{N,n,s-1} = 0] \\ &= \sum_{j=1}^n (-1)^j \binom{N}{j} \left(\frac{\binom{N-j}{n-j}}{\binom{N}{n}} \right)^{s-1} \left(\frac{\binom{N-j}{n-j}}{\binom{N}{n}} - 1 \right). \end{aligned}$$

Počítajme strednú hodnotu

$$\begin{aligned} E\tau &= \sum_{s=1}^{\infty} s P[\tau = s] = \sum_{s=1}^{\infty} s (P[\tau \geq s] - P[\tau \geq s + 1]) \\ &= \sum_{s=1}^{\infty} s P[\tau \geq s] - \sum_{s=2}^{\infty} (s - 1) (P[\tau \geq s]) \\ &= \sum_{s=1}^{\infty} P[\tau \geq s] = \sum_{s=1}^{\infty} (1 - P[\tau < s]) \\ &= \sum_{s=1}^{\infty} (1 - P[\tau \leq s]) + \sum_{s=1}^{\infty} P[\tau = s] \\ &= \sum_{s=1}^{\infty} (1 - P[C_{N,n,s} = 0]) + 1 \\ &= 1 + \sum_{s=1}^{\infty} \left(1 - \sum_{j=0}^n (-1)^j \binom{N}{j} \left(\frac{\binom{N-j}{n-j}}{\binom{N}{n}} \right)^s \right) \\ &= 1 + \sum_{s=1}^{\infty} \left(\sum_{j=1}^n (-1)^{j+1} \binom{N}{j} \left(\frac{\binom{N-j}{n-j}}{\binom{N}{n}} \right)^s \right) \\ &= 1 + \sum_{j=1}^n (-1)^{j+1} \binom{N}{j} \frac{\binom{N-j}{n-j}}{\binom{N}{n} - \binom{N-j}{n-j}} \quad \text{Q.E.D.} \end{aligned}$$

2.2 MOMENTY A ICH KONVERGENCIA

Na začiatok zavedme značenie

$$p = \frac{n}{N}, \quad (2.6)$$

$$p_j = \left(\frac{n-j}{N-j} \right)^k. \quad (2.7)$$

Potom pre indikátory I_i definované pomocou (2.4) a pre $r, s \in \mathbb{N}$ platí

$$E I_i^r = P [I_i = 1] = \left(\frac{n}{N} \right)^k = p_0 \text{ pre } 1 \leq i \leq N,$$

$$E I_i^r I_j^s = P [I_i = 1, I_j = 1] = \left(\frac{n}{N} \frac{n-1}{N-1} \right)^k = p_0 p_1 \text{ pre } 1 \leq i, j \leq N, i \neq j.$$

VETA 2.5. *Pre strednú hodnotu a rozptyl náhodnej veličiny $C_{N,n,k}$ platí*

$$\begin{aligned} E C_{N,n,k} &= N p^k, \\ \text{var } C_{N,n,k} &= N^2 p_0 (p_1 - p_0) + N p_0 (1 - p_1). \end{aligned}$$

Dôkaz. Priamym výpočtom dostaneme

$$E C_{N,n,k} = \sum_{i=1}^N E I_i = N p_0 = N p^k.$$

Keďže je

$$E C_{N,n,k}^2 = \sum_{i=1}^N E I_i^2 + \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N E I_i I_j = N p_0 + N(N-1) p_0 p_1,$$

potom

$$\begin{aligned} \text{var } C_{N,n,k} &= N p_0 + N^2 p_0 p_1 - N p_0 p_1 - N^2 p_0^2 \\ &= N^2 p_0 (p_1 - p_0) + N p_0 (1 - p_1). \quad \text{Q.E.D.} \end{aligned}$$

VETA 2.6. *Pre faktoriálne momenty rádu r veličiny $C_{N,n,k}$ platí*

$$\begin{aligned} E \left[C_{N,n,k}^{\{r\}} \right] &= E [C_{N,n,k} (C_{N,n,k} - 1) \cdots (C_{N,n,k} - (r-1))] \\ &= \left(\frac{n^{\{r\}}}{N^{\{r\}}} \right)^{k-1} n^{\{r\}} = \frac{(n(n-1) \cdots (n-(r-1)))^k}{(N(N-1) \cdots (N-(r-1)))^{k-1}} \end{aligned}$$

Dôkaz. Z predchádzajúceho už vieme, že podmienené rozdelenie $C_{N,n,k}$ pri danom $C_{N,n,k-1}$ je hypergeometrické. Vytvorujúca funkcia náhodnej veličiny majúcej hypergeometrické rozdelenie $HG(N, n, c)$ má tvar

$$G(s) = \sum_{j=0}^{\min(n,c)} \left(\binom{c}{j} \binom{N-c}{n-j} / \binom{N}{n} \right) s^j.$$

Konkrétne v našom prípade označme

$$\begin{aligned} G(s) &= G(s, n) = F(s, n) / \binom{N}{n} \\ F(s, n) &= \sum_{j=0}^c \binom{c}{j} \binom{N-c}{n-j} s^j. \end{aligned}$$

Definujme

$$Q(s, x) = \sum_{l=0}^{\infty} F(s, l) x^l$$

a počítajme

$$\begin{aligned} Q(s, x) &= \sum_{l=0}^{\infty} \sum_{j=0}^c \binom{c}{j} \binom{N-c}{l-j} s^j x^l \\ &= \sum_{j=0}^c \sum_{l=j}^{N-c+j} \binom{c}{j} \binom{N-c}{l-j} s^j x^l \\ &= \sum_{j=0}^c \binom{c}{j} (sx)^j \sum_{l=j}^{N-c+j} \binom{N-c}{l-j} x^{l-j} \\ &= \sum_{j=0}^c \binom{c}{j} (sx)^j \sum_{i=0}^{N-c} \binom{N-c}{i} x^i \\ &= (1 + sx)^c (1 + x)^{N-c}. \end{aligned}$$

Takže

$$\begin{aligned} \left(\frac{\partial}{\partial s} \right)^r Q(s, x) &= c(c-1) \cdots (c-r+1) x^r (1 + sx)^{c-r} (1 + x)^{N-c} \\ \left(\frac{\partial}{\partial s} \right)^r Q(1, x) &= c(c-1) \cdots (c-r+1) x^r (1 + x)^{N-r} \\ &= c(c-1) \cdots (c-r+1) \sum_{i=0}^{N-r} \binom{N-r}{i} x^{r+i}. \end{aligned}$$

Potom $F^{(r)}(1, n)$ je koeficient pri x^n v $(\frac{\partial}{\partial s})^r Q(1, x)$, podrobnejšie

$$F^{(r)}(1, n) = c(c-1)\cdots(c-r+1) \binom{N-r}{n-r},$$

a dostávame

$$\begin{aligned} \mathbb{E} \left[C_{N,n,k}^{\{r\}} \mid C_{N,n,k-1} = c \right] &= G^{(r)}(1, n) = F^{(r)}(1, n) / \binom{N}{n} \\ &= c(c-1)\cdots(c-(r-1)) \binom{N-r}{n-r} / \binom{N}{n} \\ &= \frac{c^{\{r\}} n^{\{r\}}}{N^{\{r\}}}. \end{aligned}$$

Konečne (vzhľadom k (1.1)) odvodíme

$$\begin{aligned} \mathbb{E} \left[C_{N,n,k}^{\{r\}} \right] &= \mathbb{E} \left[\mathbb{E} \left[C_{N,n,k}^{\{r\}} \mid C_{N,n,k-1} \right] \right] = \frac{n^{\{r\}}}{N^{\{r\}}} \mathbb{E} \left[C_{N,n,k-1}^{\{r\}} \right] \\ &= \left(\frac{n^{\{r\}}}{N^{\{r\}}} \right)^{k-1} \mathbb{E} \left[C_{N,n,1}^{\{r\}} \right] = \left(\frac{n^{\{r\}}}{N^{\{r\}}} \right)^{k-1} n^{\{r\}} \\ &= \frac{(n(n-1)\cdots(n-(r-1)))^k}{(N(N-1)\cdots(N-(r-1)))^{k-1}}. \quad \text{Q.E.D.} \end{aligned}$$

VEĽA 2.7. *Pre strednú hodnotu martingalu $\{T_k, k \in \mathbb{N}\}$ definovaného pomocou (1.2) platí*

$$\mu(k) = \mathbb{E} [T_k] = N.$$

Dôkaz. Z vety 2.5 plynie

$$\mathbb{E} [T_k] = \left(\frac{N}{n} \right)^k \mathbb{E} C_k = p^{-k} N p^k = N. \quad \text{Q.E.D.}$$

POZNÁMKA. Martingal $\{T_k, k \in \mathbb{N}\}$ je nezáporný, a preto podľa vety VI.3.3 na strane 357 v (Štěpán, 1987) existuje integrovateľná náhodná veličina T taká, že T_k konverguje k T skoro iste. Zrejme markovský čas prvého vstupu $C_{N,n,k}$ do nuly τ je markovským časom prvého vstupu do nuly martingalu $\{T_k, k \in \mathbb{N}\}$ vzhľadom k prirodzenej filtrácii. Pretože stredná hodnota náhodnej veličiny τ , daná vzťahom (2.5), je konečná, je τ skoro iste konečná náhodná veličina. Takže $T = 0$ skoro iste a $\mathbb{E} T = 0$. Z tohoto a z vety 2.7 plynie, že T_k nekonverguje k T podľa stredy.

Podmienka, že pre markovský čas τ existuje $\alpha \in (0, \infty)$ s vlastnosťou

$$\forall k \in \mathbb{N} \quad \tau > k \Rightarrow \mathbb{E} [|T_{k+1} - T_k| | T_1, T_2, \dots, T_k] \leq \alpha \quad \text{skoro iste,} \quad (2.8)$$

nemôže byť v našom prípade splnená. Ak by podmienka (2.8) splnená bola, tak by podľa vety VI.2.11 v (Štěpán, 1987) muselo byť

$$\mathbb{E}[T_\tau] = \mathbb{E}[T_1] = N.$$

To však nie je možné, keďže $T_\tau = 0$.

LEMMA 2.8. *Pre každé $k > 1$ a každé pevné $p \in (0, 1)$ platí*

$$\rho_k = \lim_{N \rightarrow \infty} \frac{1}{N} \text{var } C_{N,n,k} = p^k(1 - kp^{k-1} + (k-1)p^k). \quad (2.9)$$

Dôkaz. Z toho, že $N \rightarrow \infty$, plynie $n = pN \rightarrow \infty$. Podľa vety 2.5 je

$$\rho_k = \lim_{N \rightarrow \infty} \frac{1}{N} \text{var } C_{N,n,k} = \lim_{N \rightarrow \infty} Np_0(p_1 - p_0) + p_0(1 - p_1).$$

Dosadením za p_0 a p_1 podľa (2.6) a (2.7) je možné ρ_k prepísať do tvaru

$$\rho_k = \lim_{N \rightarrow \infty} Np^k \left(p^k \frac{\left(1 - \frac{1}{n}\right)^k}{\left(1 - \frac{1}{N}\right)^k} - p^k \right) + p^k \left(1 - p^k \frac{\left(1 - \frac{1}{n}\right)^k}{\left(1 - \frac{1}{N}\right)^k} \right)$$

Ďalej je

$$\begin{aligned} \left(\frac{1 - \frac{1}{n}}{1 - \frac{1}{N}} \right)^k &= \left(1 + \frac{\frac{1}{N} - \frac{1}{n}}{1 - \frac{1}{N}} \right)^k = \sum_{j=0}^k \binom{k}{j} \left(\frac{\frac{1}{N} - \frac{1}{n}}{1 - \frac{1}{N}} \right)^j \\ &= 1 + \frac{k}{n} \frac{p-1}{1 - \frac{1}{N}} + o\left(\frac{1}{N}\right) = 1 + \frac{k}{Np} \frac{p-1}{1 - \frac{1}{N}} + o\left(\frac{1}{N}\right). \end{aligned}$$

Potom

$$\begin{aligned} \rho_k &= \lim_{N \rightarrow \infty} Np^k \left(p^k \left(1 + \frac{k}{Np} \frac{p-1}{1 - \frac{1}{N}} + o\left(\frac{1}{N}\right) \right) - p^k \right) \\ &\quad + p^k \left(1 - p^k \left(1 + \frac{k}{Np} \frac{p-1}{1 - \frac{1}{N}} + o\left(\frac{1}{N}\right) \right) \right) \\ &= \lim_{N \rightarrow \infty} p^k \left(k p^{k-1} \frac{p-1}{1 - \frac{1}{N}} + 1 - p^k - p^{k-1} \frac{k}{N} \frac{p-1}{1 - \frac{1}{N}} \right) \\ &= p^k (1 + k p^{k-1}(p-1) - p^k) = p^k (1 + (k-1)p^k - k p^{k-1}). \quad \text{Q.E.D.} \end{aligned}$$

LEMMA 2.9. *Pre každé $k > 1$ a každé pevné $p \in (0, 1)$ platí*

$$\frac{1}{N} C_{N,n,k} \xrightarrow{P} p^k \quad \text{pre } N \rightarrow \infty. \quad (2.10)$$

Dôkaz. Podľa Čebyševovej nerovnosti pre náhodnú veličinu $\frac{1}{N} C_{N,n,k}$ je

$$\begin{aligned} P \left[\left| \frac{1}{N} C_{N,n,k} - \frac{1}{N} E C_{N,n,k} \right| \geq \varepsilon \right] &\leq \frac{1}{\varepsilon^2} \text{var} \left(\frac{1}{N} C_{N,n,k} \right) \\ &= \frac{1}{\varepsilon^2} \frac{1}{N} \left(\frac{1}{N} \text{var} C_{N,n,k} \right). \end{aligned}$$

Z vety 2.5 a lemy 2.8 dostávame

$$P \left[\left| \frac{1}{N} C_{N,n,k} - p^k \right| \geq \varepsilon \right] \rightarrow 0, \quad \text{ak } N \rightarrow \infty,$$

z čoho plynie

$$\frac{1}{N} C_{N,n,k} \xrightarrow{P} p^k. \quad \text{Q.E.D.}$$

POZNÁMKA. V článku (Hurt a kol., 1982) je možné nájsť vzorce pre výpočet všeobecných a centrálnych momentov náhodnej veličiny $C_{N,n,k}$ tretieho a štvrtého rádu a odvodenie charakteristickej funkcie.

3

ASYMPTOTICKÉ ROZDELENIE

Ako už bolo uvedené, exaktné rozdelenie náhodnej veličiny $C_{N,n,k}$ je pomerne zložité, a preto nie je ľahké s ním narábať. Často je teda vhodné použiť aproximáciu normálnym alebo Poissonovým rozdelením.

3.1 KONVERGENCIA K POISSONOVMU ROZDELENIU

V tejto chvíli nás budú zaujímať podmienky, za ktorých má $C_{N,n,k}$ pre $N \rightarrow \infty$ asymptoticky Poissonovo rozdelenie. V celej tejto podkapitole budeme predpokladať, že

$$n = n(N) \quad \text{a} \quad k = k(N),$$

teda

$$C_{N,n,k} = C_{N,n(N),k(N)} = C_N.$$

Najprv predpokladajme, že je pomer

$$\frac{n(N)}{N} = p$$

pevný. Zrejme potom je $0 < p < 1$ a $n(N) = pN$ je lineárnou funkciou N .

LEMMA 3.1. *Nech $N \rightarrow \infty$ a $k = k(N) \rightarrow \infty$ tak, že pre pevné p platí*

$$Np^k \rightarrow \lambda, \quad 0 < \lambda < \infty. \quad (3.1)$$

Potom pre každé $r \in \mathbb{N}$

$$\mathbb{E} \left[C_N^{\{r\}} \right] \rightarrow \lambda^r. \quad (3.2)$$

Dôkaz. Predpoklad (3.1) implikuje, že $k = o(N)$. Na základe vety 2.6 máme pre $r \in \mathbb{N}$

$$\begin{aligned} \mathbb{E} \left[C_N^{\{r\}} \right] &= \frac{(n(n-1)\cdots(n-(r-1)))^k}{(N(N-1)\cdots(N-(r-1)))^{k-1}} \\ &= \frac{(n^r)^k}{(N^r)^{k-1}} \frac{\left(\left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{r-1}{n}\right) \right)^k}{\left(\left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{r-1}{N}\right) \right)^{k-1}} \\ &= \left(\frac{n^k}{N^{k-1}} \right)^r \frac{\prod_{\ell=1}^{r-1} \left(1 - \frac{\ell}{N}\right) \prod_{i=1}^{r-1} \left(1 - \frac{i}{n}\right)^k}{\prod_{j=1}^{r-1} \left(1 - \frac{j}{N}\right)^k} \\ &= (Np^k)^r \prod_{\ell=1}^{r-1} \left(1 - \frac{\ell}{N}\right) \frac{\prod_{i=1}^{r-1} \left(1 - \frac{i/p}{N}\right)^k}{\prod_{j=1}^{r-1} \left(1 - \frac{j}{N}\right)^k}. \end{aligned}$$

Takže (s ohľadom na $k = o(N)$)

$$\mathbb{E} \left[C_N^{\{r\}} \right] \xrightarrow[k \rightarrow \infty]{N \rightarrow \infty} \lambda^r. \quad \text{Q.E.D.}$$

VETA 3.2 (KONVERGENCIA K POISSONOVEMU ROZDELENIU 1). *Nech $N \rightarrow \infty$ a $k = k(N) \rightarrow \infty$ tak, že pre pevné p platí*

$$Np^k \rightarrow \lambda, \quad 0 < \lambda < \infty. \quad (3.3)$$

Potom asymptotické rozdelenie náhodnej veličiny C_N je Poissonovo s parametrom λ .

Dôkaz. Vidíme, že sú splnené predpoklady lemy 3.1, takže platí (3.2). Pre faktoriálne momenty rádu r náhodnej veličiny X s Poissonovým rozdelením s parametrom λ (viď napríklad (Johnson a kol., 2005)) platí

$$\mathbb{E} X^{\{r\}} = \lambda^r.$$

Teda limita faktoriálnych momentov veličiny C_N pre $N \rightarrow \infty$ je rovná faktoriálnym momentom (rovnakého rádu) náhodnej veličiny X s Poissonovým rozdelením s parametrom λ . Očividne platí analogický vzťah aj pre všeobecné momenty. Tvrdenie dokazovanej vety dostaneme aplikovaním vety A.1 s ohľadom na to, že polomer konvergenencie mocninného radu (A-1) vo vete A.1 je pre náhodnú veličinu X s Poissonovým rozdelením nekonečný. Q.E.D.

Teraz pre zmenu budeme uvažovať fixný počet výberov k a budeme opäť skúmať podmienky konvergenzie k Poissonovmu rozdeleniu. Ukazuje sa však, že budeme musieť upustiť od predpokladu o pevnom pomere p , preto ďalej je $p = p(N)$.

VETA 3.3 (KONVERGENCIA K POISSONOVEMU ROZDELENIU 2). *Nech $N \rightarrow \infty$ a $n = n(N) \rightarrow \infty$ tak, že pre pevné k platí*

$$N [p(N)]^k \rightarrow \lambda, \quad 0 < \lambda < \infty. \quad (3.4)$$

Potom asymptotické rozdelenie C_N je Poissonovo s parametrom λ .

Dôkaz. Zopakovaním výpočtu z dôkazu lemy 3.1 pre pevné k , $N \rightarrow \infty$, $n = n(N) \rightarrow \infty$ a $N [p(N)]^k \rightarrow \lambda$ opäť dostávame (3.2). K tomu, že C_N má asymptoticky Poissonovo rozdelenie, dospejeme použitím rovnakého postupu ako v dôkaze vety 3.2. Q.E.D.

3.2 KONVERGENCIA K NORMÁLNEMU ROZDELENIU

Pri odvodení asymptotickej normality zohrá pre nás dôležitú úlohu nasledujúca veta, vid' (Rényi, 1972) veta 2 na str. 398. Uvažujeme konečný súbor jednotiek veľkosti N . Na každej z N jednotiek pozorujeme nejaký znak. Nech práve M jednotiek má daný znak a $N - M$ jednotiek daný znak nemá. Náhodne vyberieme n jednotiek bez vracania. V tomto prípade má náhodná veličina predstavujúca počet prvkov výberu, ktoré majú uvažovanú vlastnosť, hypergeometrické rozdelenie $HG(N, n, M)$. Bez obmedzenia na všeobecnosti môžeme predpokladať, že $M \leq \frac{N}{2}$ a $n \leq \frac{N}{2}$.

VETA 3.4. *Nech N , M a n sú prirodzené čísla, $1 \leq M \leq \frac{N}{2}$, $1 \leq n \leq \frac{N}{2}$. Položme $r = \frac{M}{N}$ a $p = \frac{n}{N}$. Potom*

$$\lim_{Nrp \rightarrow \infty} \sum_{j \leq nr + x \sqrt{nr(1-r)(1-p)}} \frac{\binom{M}{j} \binom{N-M}{n-j}}{\binom{N}{n}} = \Phi(x). \quad (3.5)$$

Asymptotickú normalitu $C_{N,n,k}$ vyšetříme pre pevný počet výberov k a zafixujeme aj pomer $p = \frac{n}{N}$.

VETA 3.5 (KONVERGENCIA K NORMÁLNEMU ROZDELENIU PRE $k = 2$). *Pre každé pevné $p \in (0, 1)$ platí*

$$\frac{C_{N,n,2} - Np^2}{\sqrt{N}} \xrightarrow[N \rightarrow \infty]{\mathcal{D}} N(0, \rho_2), \quad (3.6)$$

kde

$$\rho_2 = p^2(1-p)^2. \quad (3.7)$$

Dôkaz. Využijeme centrálnu limitnú vetu pre hypergeometrické rozdelenie, vid' veta 3.4. S ohľadom na (1.1) je v tomto prípade

$$n = M \quad (3.8)$$

a platí, že

$$C_{N,n,2} \sim HG(N, n, n).$$

Z (3.8) plynie (používame značenie vety 3.4)

$$r = p = \text{konštanta},$$

a preto

$$Nrp \longrightarrow \infty \quad \Leftrightarrow \quad N \longrightarrow \infty.$$

Z vety 3.4 teda vyplýva

$$\begin{aligned} \lim_{N \rightarrow \infty} \sum_{j \leq np + x\sqrt{np(1-p)^2}} \mathbb{P} [C_{N,n,2} = j] &= \Phi(x) \\ \lim_{N \rightarrow \infty} \sum_{\frac{j-np}{\sqrt{np(1-p)^2}} \leq x} \mathbb{P} \left[\frac{C_{N,n,2} - np}{\sqrt{np(1-p)^2}} = \frac{j - np}{\sqrt{np(1-p)^2}} \right] &= \Phi(x) \\ \lim_{N \rightarrow \infty} \sum_{i \leq x} \mathbb{P} \left[\frac{C_{N,n,2} - np}{\sqrt{np(1-p)^2}} = i \right] &= \Phi(x). \end{aligned}$$

Platí, že

$$np = Np^2 = \mathbb{E} C_{N,n,2},$$

ďalej jednoduchým výpočtom (s využitím (3.7)) dostaneme

$$\sqrt{np(1-p)^2} = \sqrt{\frac{n}{p} p^2(1-p)^2} = \sqrt{N\rho_2}.$$

Teda máme

$$\begin{aligned} \lim_{N \rightarrow \infty} \sum_{i \leq x} \mathbb{P} \left[\frac{C_{N,n,2} - Np^2}{\sqrt{N\rho_2}} = i \right] &= \Phi(x) \\ \lim_{N \rightarrow \infty} \mathbb{P} \left[\frac{C_{N,n,2} - Np^2}{\sqrt{N\rho_2}} \leq x \right] &= \Phi(x), \end{aligned}$$

takže

$$\frac{C_{N,n,2} - Np^2}{\sqrt{N}} \xrightarrow{\mathcal{D}} N(0, \rho_2) \quad \text{pre } N \rightarrow \infty,$$

a tým máme dokázané tvrdenie vety.

Q.E.D.

Pre $k > 2$ nie je rozšírenie tejto vety jednoduché metódami klasickej pravdepodobnosti. Napriek tomu však platí

VETA 3.6 (KONVERGENCIA K NORMÁLNEMU ROZDELENIU). *Pre každé pevné $k > 1$ a každé pevné $p \in (0, 1)$ platí*

$$\frac{C_{N,n,k} - Np^k}{\sqrt{N}} \xrightarrow[N \rightarrow \infty]{\mathcal{D}} N(0, \rho_k), \quad (3.9)$$

kde ρ_k je dané (2.9).

POZNÁMKA. Vetu 3.6 dokážeme nezávisle na vete 3.5.

Dôkaz. Vetu dokážeme indukciou podľa k .

Pre každé $N \in \mathbb{N}$ uvažujme, že výber máme reprezentovaný nasledujúcim spôsobom: Pre i -tu jednotku v populácii veľkosti N , $i = 1, \dots, N$, máme dané x_{Ni} , ktoré nadobúda hodnotu buď 0 alebo 1. Spomedzi hodnôt $x_{N1}, x_{N2}, \dots, x_{NN}$ je práve $n = n(N) = pN$ jedničiek a $N - n$ núl. Bez újmy na všeobecnosti predpokladajme, že

$$x_{Ni} = \begin{cases} 1 & \text{pre } 1 \leq i \leq n, \\ 0 & \text{pre } n + 1 \leq i \leq N. \end{cases}$$

Uvažujme náhodnú permutáciu

$$\psi_{N1}, \psi_{N2}, \dots, \psi_{NN} \quad (3.10)$$

hodnôt $x_{N1}, x_{N2}, \dots, x_{NN}$. Táto permutácia náhodne priradí n jednotkám populácie hodnotu 1 a tieto prvky budeme považovať za náhodný výber o rozsahu n z populácie veľkosti N .

V nasledujúcom bude pre nás výhodné použiť transformáciu x_{Ni} , ktorú označíme y_{Ni} , tak aby boli splnené normujúce podmienky

$$\sum_{i=1}^N y_{Ni} = 0 \quad \text{a} \quad \sum_{i=1}^N y_{Ni}^2 = 1. \quad (3.11)$$

Uvažujme transformáciu tvaru

$$y_{Ni} = a + bx_{Ni}.$$

Teda

$$y_{Ni} = \begin{cases} a_1 + b_1 x_{Ni} = a_1 + b_1 & \text{pre } 1 \leq i \leq n, \\ a_0 + b_0 x_{Ni} = a_0 & \text{pre } n + 1 \leq i \leq N. \end{cases}$$

Z (3.11) dostávame sústavu rovníc

$$n(a_1 + b_1) + (N - n)a_0 = 0 \quad (3.12)$$

$$n(a_1 + b_1)^2 + (N - n)a_0^2 = 1. \quad (3.13)$$

Vyjadrením a_0 z (3.12) a dosadením do (3.13) dostaneme ako jedno riešenie tejto sústavy

$$y_{Ni} = \begin{cases} a_1 + b_1 = \sqrt{\frac{N-n}{nN}} = \frac{1-p}{\sqrt{N(1-p)p}} & \text{pre } 1 \leq i \leq n, \\ a_0 = -\frac{n}{N-n}(a_1 + b_1) = -\frac{p}{\sqrt{N(1-p)p}} & \text{pre } n + 1 \leq i \leq N. \end{cases}$$

Takže pre $y_{N1}, y_{N2}, \dots, y_{NN}$ sú splnené podmienky (3.11) a taktiež platí

$$\max_{1 \leq i \leq N} |y_{Ni}| \rightarrow 0 \quad \text{pre } N \rightarrow \infty.$$

Uvažujme pravdepodobnostný priestor dostatočne veľký, aby bolo možné pre každé $N \in \mathbb{N}$ definovať náhodnú permutáciu

$$\xi_{N1}, \xi_{N2}, \dots, \xi_{NN} \quad (3.14)$$

hodnôt $y_{N1}, y_{N2}, \dots, y_{NN}$, pričom každá z týchto $N!$ permutácií má rovnakú pravdepodobnosť. Definujme

$$Y_t^N = \begin{cases} \sum_{i=1}^{[Nt]} \xi_{Ni} & \text{pre } \frac{1}{N} \leq t \leq 1, \\ 0 & \text{pre } 0 \leq t < \frac{1}{N}. \end{cases} \quad (3.15)$$

Potom $Y^N = \{Y_t^N, 0 \leq t \leq 1\}$ je proces s trajektóriami v priestore $D[0, 1]$. Podľa vety A.18

$$Y^N \xrightarrow[N \rightarrow \infty]{\mathcal{D}} W^0, \quad (3.16)$$

kde W^0 je Brownov most, viď definícia A.4.

Nech je najprv $k = 2$. V prvom kroku matematickej indukcie voľme špeciálne $t = p$ a dostaneme

$$Y_p^N = \sum_{i=1}^{[Np]} \xi_{Ni} = \sum_{i=1}^n \xi_{Ni}. \quad (3.17)$$

Predstavme si situáciu prieniku dvoch nezávislých výberov o rozsahu n z populácie veľkosti N . Očíslujme jednotky v populácii hodnotami $1, 2, \dots, N$ a uvažujme prvý z výberov. Bez újmy na všeobecnosti môžeme (po eventuálnom prečíslovaní prvkov populácie) predpokladať, že v prvom výbere boli vytiahnuté prvky, ktorým sme priradili čísla $1, \dots, n$. Uvažujme najprv permutáciu (3.10) n jedničiek a $N - n$ núl. Potom náhodná veličina

$$\sum_{j=1}^n \psi_{Nj} \quad (3.18)$$

predstavuje počet “zásahov” prvkami druhého výberu do prvého výberu, tj. mohutnosť ich prieniku. Avšak my uvažujeme miesto núl a jedničiek znormované hodnoty y_{Ni} , preto miesto veličiny (3.18) máme

$$D_{N,n,2} = \text{počet výskytov hodnoty } \frac{1-p}{\sqrt{N(1-p)p}} \\ \text{na prvých } n \text{ miestach permutácie (3.14)}$$

a $D_{N,n,2}$ je náhodná veličina s rovnakým rozdelením ako $C_{N,n,2}$ (ide zrejme o hypergeometrické rozdelenie $HG(N, n, n)$). Teraz (3.17) môžeme prepísať

$$\begin{aligned} Y_p^N &= \sum_{i=1}^n \xi_{Ni} = D_{N,n,2} \frac{1-p}{\sqrt{N(1-p)p}} - (n - D_{N,n,2}) \frac{p}{\sqrt{N(1-p)p}} \\ &= \frac{D_{N,n,2} - np}{\sqrt{N(1-p)p}} = \frac{D_{N,n,2} - Np^2}{\sqrt{N(1-p)p}}. \end{aligned}$$

Z (3.16) a vety A.15 špeciálne plynie

$$Y_p^N \xrightarrow[N \rightarrow \infty]{\mathcal{D}} W_p^0. \quad (3.19)$$

Z vlastností Brownovho mostu (viď veta A.5) vieme, že

$$\mathcal{L}(W_p^0) = N(0, p(1-p)).$$

Ďalej je

$$\rho_2 = (p(1-p))^2,$$

a tak (keďže $D_{N,n,2}$ má rovnaké rozdelenie ako $C_{N,n,2}$) dostávame tvrdenie vety 3.6 pre $k = 2$ a iný dôkaz vety 3.5.

Teraz pristúpme k indukčnému kroku. Predpokladajme, že dokazovaný vzťah (3.9) platí pre nejaké $k \geq 2$. Zvoľme pravdepodobnostný priestor tak,

aby bolo možné pre každé $N \in \mathbb{N}$ na ňom definovať náhodnú veličinu $D_{N,n,k}$ a náhodnú permutáciu $\xi_{N1}, \xi_{N2}, \dots, \xi_{NN}$ tak, že

$$\mathcal{L}(D_{N,n,k}) = \mathcal{L}(C_{N,n,k}) \quad (3.20)$$

a

$$D_{N,n,k} \text{ a } (\xi_{N1}, \xi_{N2}, \dots, \xi_{NN}) \text{ sú nezávislé.} \quad (3.21)$$

Opäť definujme $Y^N = \{Y_t^N, 0 \leq t \leq 1\}$ vzťahom (3.15) a dostávame (3.16). Z indukčného predpokladu plynie

$$\frac{D_{N,n,k} - Np^k}{\sqrt{N}} \xrightarrow{\mathcal{D}} N(0, \rho_k) \quad \text{pre } N \rightarrow \infty. \quad (3.22)$$

Ako dôsledok (3.21) dostaneme

$$D_{N,n,k} \text{ a } Y^N \text{ sú nezávislé } \forall N \in \mathbb{N}.$$

Preto podľa vety A.14

$$\frac{D_{N,n,k} - Np^k}{\sqrt{N}} + Y^N \xrightarrow[N \rightarrow \infty]{\mathcal{D}} U + W^0, \quad \text{kde } U \sim N(0, \rho_k)$$

a U, W^0 sú nezávislé (na vhodnom pravdepodobnostnom priestore). Ďalej ako dôsledok jednoduchej transformácie dostaneme

$$\frac{p(D_{N,n,k} - Np^k)}{\sqrt{Np(1-p)}} + Y^N \xrightarrow[N \rightarrow \infty]{\mathcal{D}} V + W^0,$$

kde

$$V = U \sqrt{\frac{p}{1-p}} \sim N\left(0, \frac{p\rho_k}{1-p}\right).$$

Využijeme náhodnú zmenu času, vo vete A.16 položíme $T_N = \frac{D_{N,n,k}}{N}$. Lemma 2.9 hovorí o konvergencii v pravdepodobnosti náhodnej veličiny $\frac{C_{N,n,k}}{N}$. Keďže konvergencia v pravdepodobnosti implikuje konvergenciu v distribúcii a predpokladáme (3.20), dostávame, že $\frac{D_{N,n,k}}{N} \xrightarrow{\mathcal{D}} p^k$, kde p^k je konštanta. Z vety III.4.22 v (Štěpán, 1987) preto plynie

$$T_N = \frac{D_{N,n,k}}{N} \xrightarrow{\mathbb{P}} p^k = a.$$

Odtiaľ podľa vety A.16 dostávame

$$\frac{p(D_{N,n,k} - Np^k)}{\sqrt{Np(1-p)}} + Y_{\frac{D_{N,n,k}}{N}}^N \xrightarrow{\mathcal{D}} V + W_{p^k}^0.$$

Z vlastností Brownovho mostu (viď veta A.5) plynie, že

$$W_{p^k}^0 \sim N(0, p^k(1-p^k)).$$

Teda $V + W_{p^k}^0$ je konvolúcia nezávislých normálne rozdelených náhodných veličín a platí

$$V + W_{p^k}^0 \sim N\left(0, \frac{p\rho_k}{1-p} + p^k(1-p^k)\right).$$

Pre $N \rightarrow \infty$ dostávame, že

$$\frac{p(D_{N,n,k} - Np^k)}{\sqrt{Np(1-p)}} + Y_{\frac{D_{N,n,k}}{N}}^N \xrightarrow{\mathcal{D}} N\left(0, \frac{p\rho_k}{1-p} + p^k(1-p^k)\right).$$

Dá sa postupne odvodiť, že

$$\begin{aligned} \frac{p(D_{N,n,k} - Np^k)}{\sqrt{Np(1-p)}} + Y_{\frac{D_{N,n,k}}{N}}^N &= \frac{p(D_{N,n,k} - Np^k)}{\sqrt{Np(1-p)}} + \sum_{i=1}^{D_{N,n,k}} \xi_{Ni} \\ &= \frac{p(D_{N,n,k} - Np^k)}{\sqrt{Np(1-p)}} + \frac{D_{N,n,k+1}(1-p)}{\sqrt{N(1-p)p}} \\ &\quad - \frac{(D_{N,n,k} - D_{N,n,k+1})p}{\sqrt{N(1-p)p}} \\ &= \frac{D_{N,n,k+1} - Np^{k+1}}{\sqrt{N(1-p)p}}, \end{aligned} \quad (3.23)$$

pričom pri úprave (3.23) možno použiť rovnaký argument ako pri (3.17). Opäť uvažujme, že sú jednotky v populácii očíslované hodnotami $1, \dots, N$. Predstavme si situáciu postupnosti k nezávislých výberov o rozsahu n z tejto populácie. Označme počet prvkov vybraných v každom z k výberov $D_{N,n,k}$. Bez újmy na všeobecnosti predpokladajme, že sme týmto prvkom priradili čísla $1, \dots, D_{N,n,k}$. Teraz nás zaujíma počet “zásahov” prvkami $(k+1)$. výberu do prieniku prvých k výberov, tj. mohutnosť prieniku týchto $k+1$ výberov. Označme teda

$$D_{N,n,k+1} = \text{počet výskytov hodnoty } \frac{1-p}{\sqrt{N(1-p)p}} \\ \text{na prvých } D_{N,n,k} \text{ miestach permutácie (3.14),}$$

a pretože predpokladáme (3.21), platí

$$\mathcal{L}(D_{N,n,k+1}) = HG(N, n, D_{N,n,k}) = \mathcal{L}(C_{N,n,k+1}).$$

Ďalej je

$$\begin{aligned} \frac{p\rho_k}{1-p} + p^k(1-p^k) &= p^{k+1} \frac{p(1-kp^{k-1} + (k-1)p^k) + (1-p^k)(1-p)}{p(1-p)} \\ &= \frac{p^{k+1}(1-(k+1)p^k + kp^{k+1})}{p(1-p)} \\ &= \frac{\rho_{k+1}}{p(1-p)}. \end{aligned}$$

Teda dostávame

$$\begin{aligned} \frac{D_{N,n,k+1} - Np^{k+1}}{\sqrt{N(1-p)p}} &\xrightarrow{\mathcal{D}} N\left(0, \frac{\rho_{k+1}}{p(1-p)}\right) \\ \frac{D_{N,n,k+1} - Np^{k+1}}{\sqrt{N}} &\xrightarrow{\mathcal{D}} N(0, \rho_{k+1}) \end{aligned}$$

a odvodili sme platnosť dokazovaného vzťahu pre $D_{N,n,k+1}$. Keďže náhodná veličina $C_{N,n,k+1}$ má rovnaké rozdelenie ako $D_{N,n,k+1}$, je tým dôkaz dokončený.

Q.E.D.

3.3 SIMULÁCIA

V predchádzajúcej časti práce sme dokázali konvergenciu náhodnej veličiny $C_{N,n,k}$ k Poissonovmu a normálnemu rozdeleniu. Teraz sa budeme zaoberať simuláciou, ktorej cieľom je vyšetriť vhodnosť aproximácie Poissonovým a normálnym rozdelením. Závery o rýchlosti konvergenzie k Poissonovmu a normálnemu rozdeleniu je možné urobiť na základe výstupov simulácie v R. Zdrojové kódy funkcií použitých v našej simulácii je možné nájsť v sekcii A.2 v Appendixe.

Uvažujme najprv situáciu z vety 3.3, teda prípad konvergenzie k Poissonovmu rozdeleniu pre pevné k . Budeme skúmať konvergenciu k Poissonovmu rozdeleniu s pevne zvoleným parametrom λ . Budeme postupne uvažovať rastúce hodnoty veľkosti populácie N . Hodnoty parametrov N , n a k však nebudeme voliť celkom ľubovoľne. Na to, aby bola približne splnená podmienka (3.4) z vety 3.3, budeme vychádzať zo vzťahu

$$n \doteq \sqrt[k]{\lambda N^{1-\frac{1}{k}}}, \quad (3.24)$$

ktorý plynie elementárnymi úpravami z (3.4). Alebo ekvivalentne k (3.24) je

$$\begin{aligned} p &= \sqrt[k]{\frac{\lambda}{N}} \\ n &\doteq pN. \end{aligned}$$

Výpočet tejto aproximácie je úlohou funkcie `poisson.2(p.N,p.lambda, p.k)` pre zadanú veľkosť populácie `p.N`, hodnotu parametra Poissonovho rozdelenia `p.lambda` a fixný počet výberov `p.k`.

Analogicky pristupujeme k vete 3.2, ktorá hovorí o konvergencii k Poissonovmu rozdeleniu pre pevné p . Opäť použijeme približný vzťah

$$k \doteq \frac{\ln\left(\frac{\lambda}{N}\right)}{\ln p}, \quad (3.25)$$

teda máme približne splnenú podmienku (3.3). Funkcia `poisson.1(p.N, p.lambda,p.p)` uskutočňuje tento výpočet pre zadanú veľkosť populácie `p.N`, hodnotu parametra Poissonovho rozdelenia `p.lambda` a fixný pomer `p.p`.

Funkcia `poisson.sample(p.N,p.lambda,p.size,p.fix,p.pk)` pre danú veľkosť populácie `p.N` a hodnotu parametra Poissonovho rozdelenia `p.lambda` vracia `p.size` realizácií náhodnej veličiny $C_{p,N}$, pričom pomocou parametra `p.fix` určujeme, či držíme pevný parameter p (`p.fix= 1`) alebo k (`p.fix= 2`). Potom `p.pk` predstavuje hodnotu p , respektíve k .

V prípade vety 3.6 overujeme rýchlosť konvergencie normovanej náhodnej veličiny

$$\frac{C_{N,n,k} - Np^k}{\sqrt{N}} \quad (3.26)$$

k normálnemu rozdeleniu $N(0, \rho_k)$, kde ρ_k je dané vzťahom (2.9). Funkcia `normal.sample(p.N,p.size,p.k,p.p)` pre zadanú veľkosť populácie `p.N`, počet výberov `p.k` a pomer `p.p` vracia `p.size` realizácií náhodnej veličiny znormovaných v zmysle (3.26).

Pomocou práve spomínaných funkcií sme pre rastúce hodnoty N a pre rôzne pevné hodnoty ostatných parametrov zostrojili 5000 realizácií náhodnej veličiny C_N , respektíve $C_{N,n,k}$, a graficky ich znázornili pomocou histogramov. Pre lepšiu prehľad sú v obrázkoch zobrazené aj pravdepodobnosti Poissonovho rozdelenia alebo hustota normálneho rozdelenia s príslušnými

parametrami. Obrázok 3.1 ilustruje konvergenciu k Poissonovmu rozdeleniu s parametrom $\lambda = 5$ pre pevné $k = 2$. Podobne obrázok 3.2 zobrazuje konvergenciu k Poissonovmu rozdeleniu $Po(5)$ pre pevné $p = \frac{2}{5}$. Pretože veľkosť výberu n a počet výberov k by mali byť celé čísla, používame pri ich výpočte vo vzťahoch (3.24) a (3.25) zaokrúhľovanie. Na obrázku 3.2 si môžeme všimnúť, že ak (pred zaokrúhľením) bola hodnota k celé číslo alebo veľmi blízka celému číslu, s rastúcim N sa príslušný histogram postupne tesnejšie primyká k hustote $Po(5)$. Z obrázkov je však zrejmé, že zaokrúhľovanie u k ovplyvňuje hodnoty, ktoré nadobúda $C_{N,n,k}$, oveľa výraznejšie ako rovnaké zaokrúhľenie u n . V súlade s tým, že pri rastúcom počte výberov je mohutnosť ich prieniku nerastúca, je histogram v obrázku 3.2 pri zaokrúhľení k nadol posunutý doprava oproti hustote $Po(5)$ a naopak pri zaokrúhľení nahor je histogram posunutý smerom doľava. Preto sme skúsili voliť N tak, aby k nadobúdalo celočíselné hodnoty alebo hodnoty blízke celočíselným hodnotám. Presnejšie, pre zvolené $\lambda = 5$, $p = \frac{2}{5}$ a pre k nadobúdajúce postupne hodnoty $1, 2, \dots$ sme určili hodnotu N ako

$$N \doteq \frac{\lambda}{p^k}. \quad (3.27)$$

Takto sme dostali hodnoty N , ktoré sme zaokrúhlili na celé číslo, a zopakovali sme postup ako u obrázku 3.2, avšak v tomto prípade sa už hodnoty k zaokrúhľením takmer nezmenili. Výsledok je znázornený na obrázku 3.3. Vidíme, že rýchlosť konverencie je porovnateľná s tou u pevného k (obrázok 3.1). A napokon, obrázok 3.4 predstavuje konvergenciu k normálnemu rozdeleniu $N(0, \rho_2)$ pre pevné $k = 2$ a $p = \frac{2}{5}$.

Na priloženom CD je možné nájsť analogicky vytvorené obrázky pre rozlične volené hodnoty parametrov (opäť generujeme 5000 realizácií náhodnej veličiny). Na základe týchto obrázkov môžeme urobiť záver, že nami skúmaná konvergencia je dosť pomalá. Taktiež sa zdá, že pre relatívne malé hodnoty parametra λ limitného Poissonovho rozdelenia sa histogram rýchlejšie približuje hustote $Po(\lambda)$.

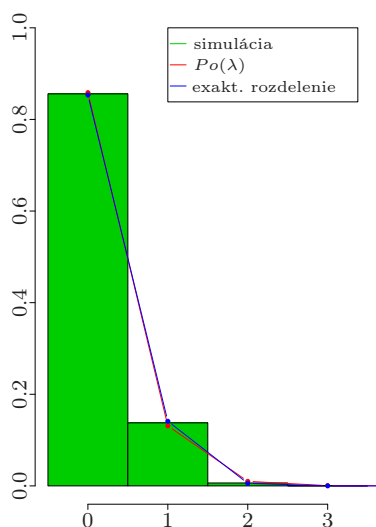
POZNÁMKA. Pre hodnoty parametra λ väčšie než 15 je podľa (Johnson a kol., 2005) pre Poissonovo rozdelenie vhodnou aproximáciou normálne rozdelenie. Ak by sme chceli aproximovať rozdelenie $C_{N,n,k}$ Poissonovým rozdelením, ktorého hodnota parametra by bola väčšia než 15, bolo by vhodné miesto aproximácie Poissonovým rozdelením zvažovať normálnu aproximáciu.

PRÍKLAD 3.7 (POKRAČOVANIE PRÍKLADU 2.3). Uvažujme opäť náhodnú veličinu $C_{16,5,4}$ z príkladu 2.3. Zaujímalo by nás, či by nebolo možné jej exaktné rozdelenie aproximovať Poissonovým, prípadne normálnym rozdelením. Parameter Poissonovho rozdelenia položíme (na základe (3.3) alebo (3.4)) rovný

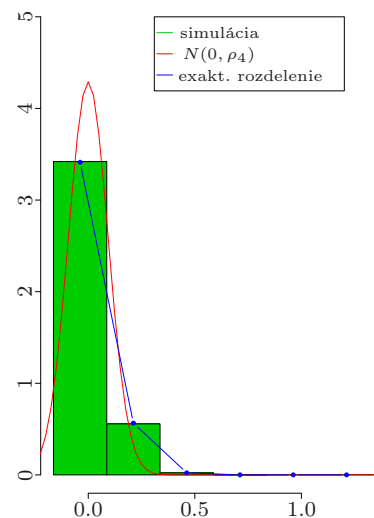
$$\lambda = Np^k = \frac{n^k}{N^{k-1}} = \frac{625}{4096} \doteq 0.1525879. \quad (3.28)$$

Inými slovami, pokúsime sa rozdelenie $C_{16,5,4}$ aproximovať Poissonovým rozdelením s rovnakou strednou hodnotou ako má $C_{16,5,4}$. A ďalej podľa (2.9) dostaneme

$$\rho_4 \doteq 0.008645438. \quad (3.29)$$

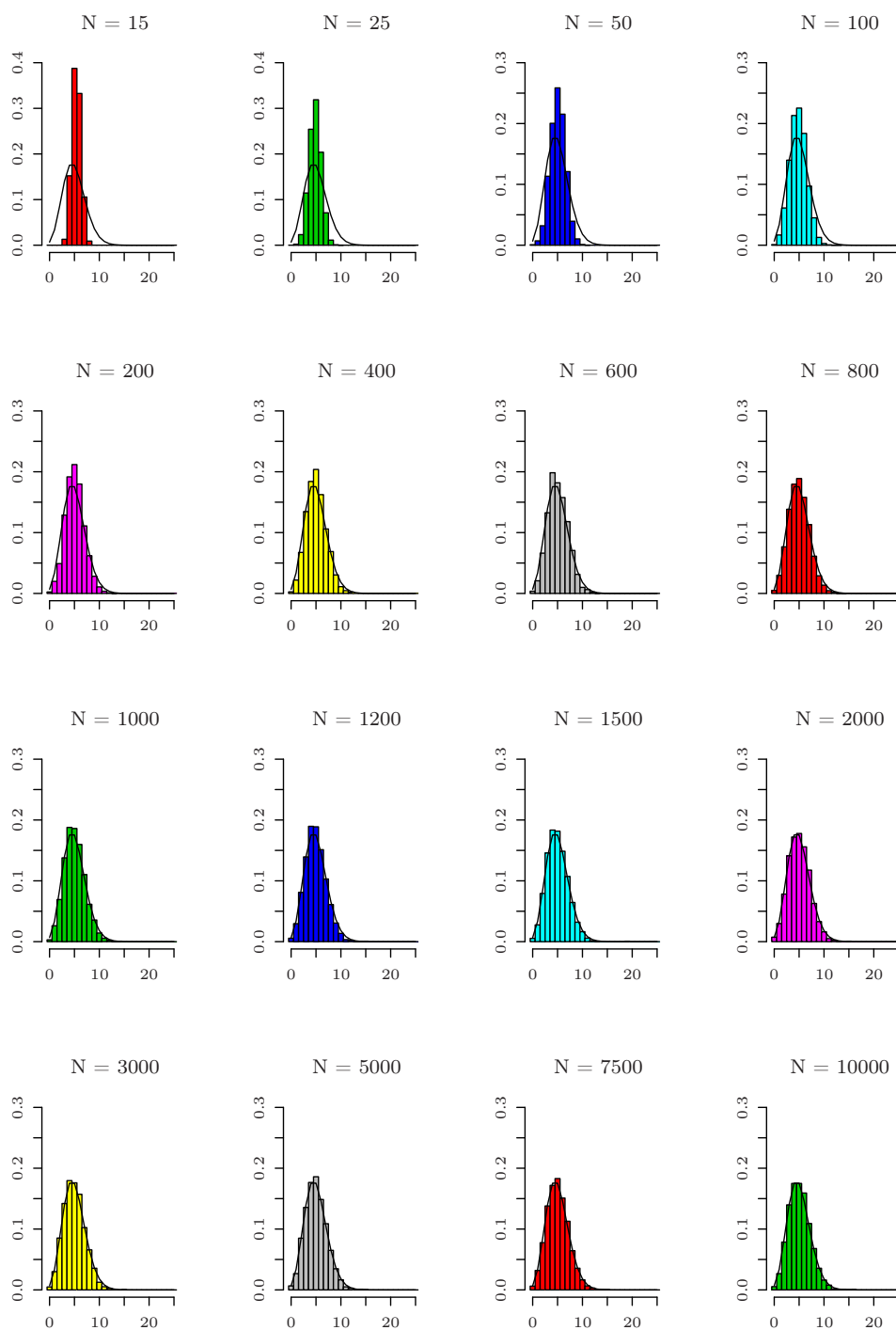


Obrázok 3.5: Konvergencia k Poissonovmu rozdeleniu.

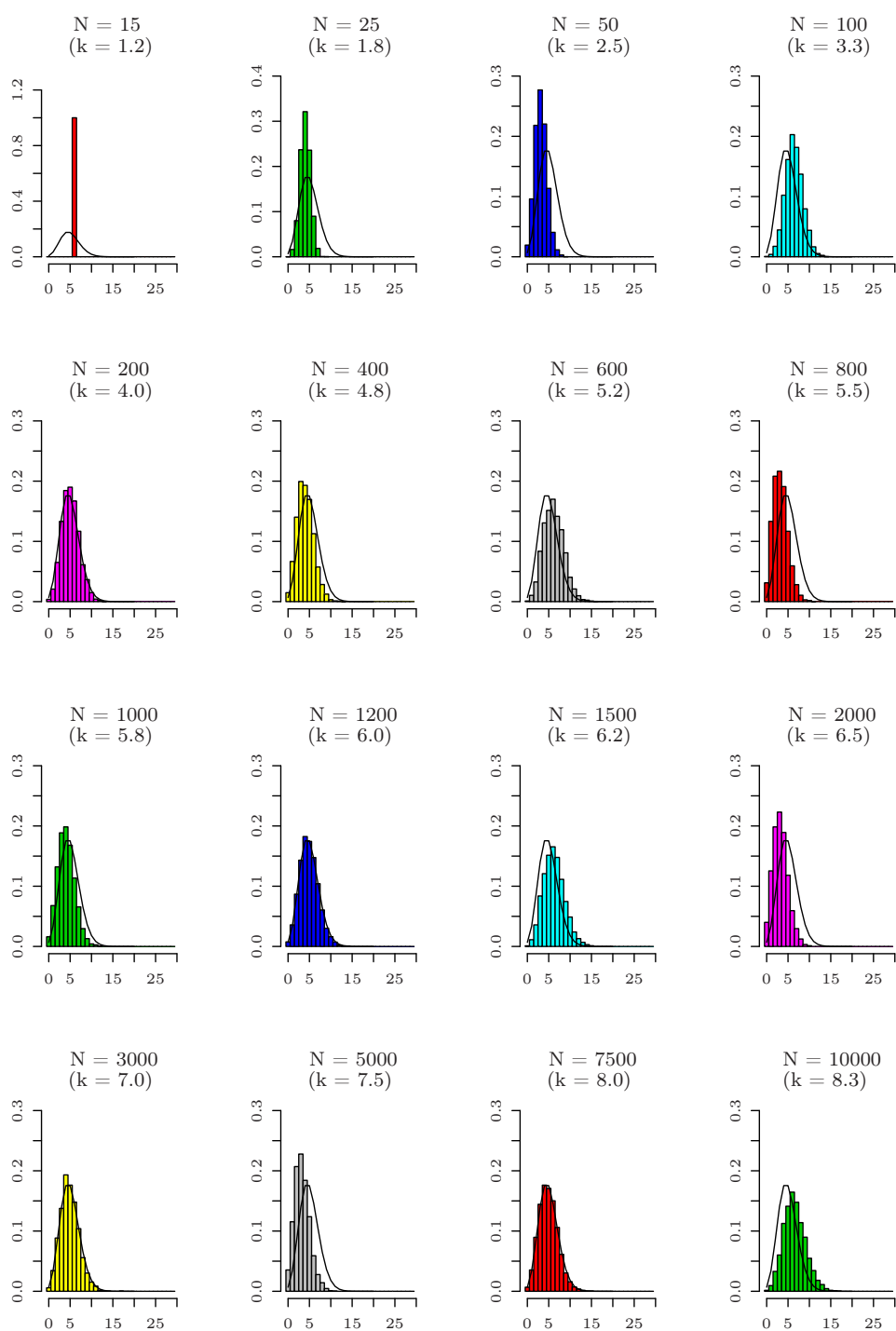


Obrázok 3.6: Konvergencia k normálnemu rozdeleniu.

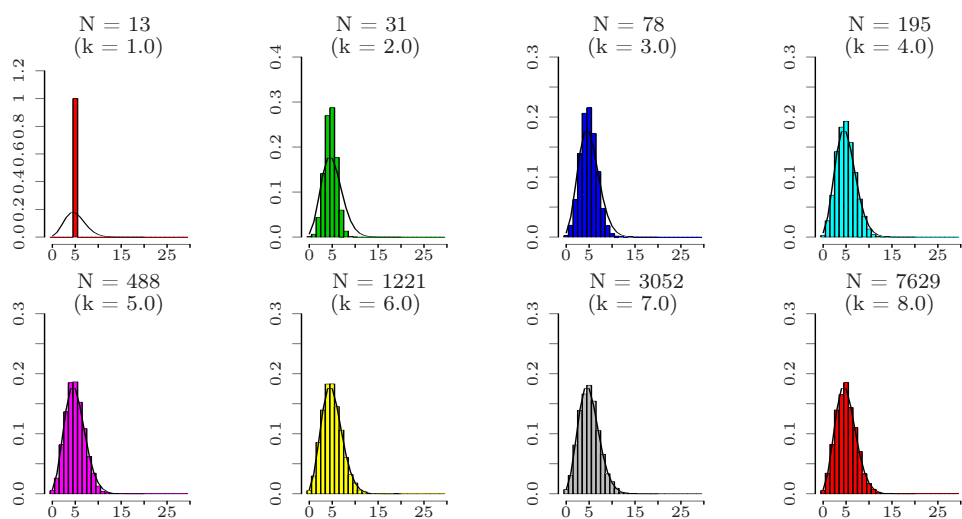
Jednoduchou modifikáciou predchádzajúcej simulácie sme vygenerovali 5000 realizácií náhodnej veličiny $C_{16,5,4}$. Takto vygenerované hodnoty sme znázornili pomocou histogramu do obrázku 3.5 spolu s pravdepodobnosťami exaktného a Poissonovho rozdelenia s parametrom λ daným (3.28). Podobne sme na obrázku 3.6 nechali vykresliť histogram vygenerovaných hodnôt znormovaných v zmysle (3.26), pravdepodobnosti náhodnej veličiny $C_{16,5,4}$ znormovanej ako v (3.26) a hustotu normálneho rozdelenia $N(0, \rho_4)$, kde hodnota ρ_4 je daná (3.29). Na základe týchto obrázkov môžeme usúdiť, že aproximáciu Poissonovým rozdelením by sme použiť mohli, avšak normálna aproximácia sa nezdá byť vhodná.



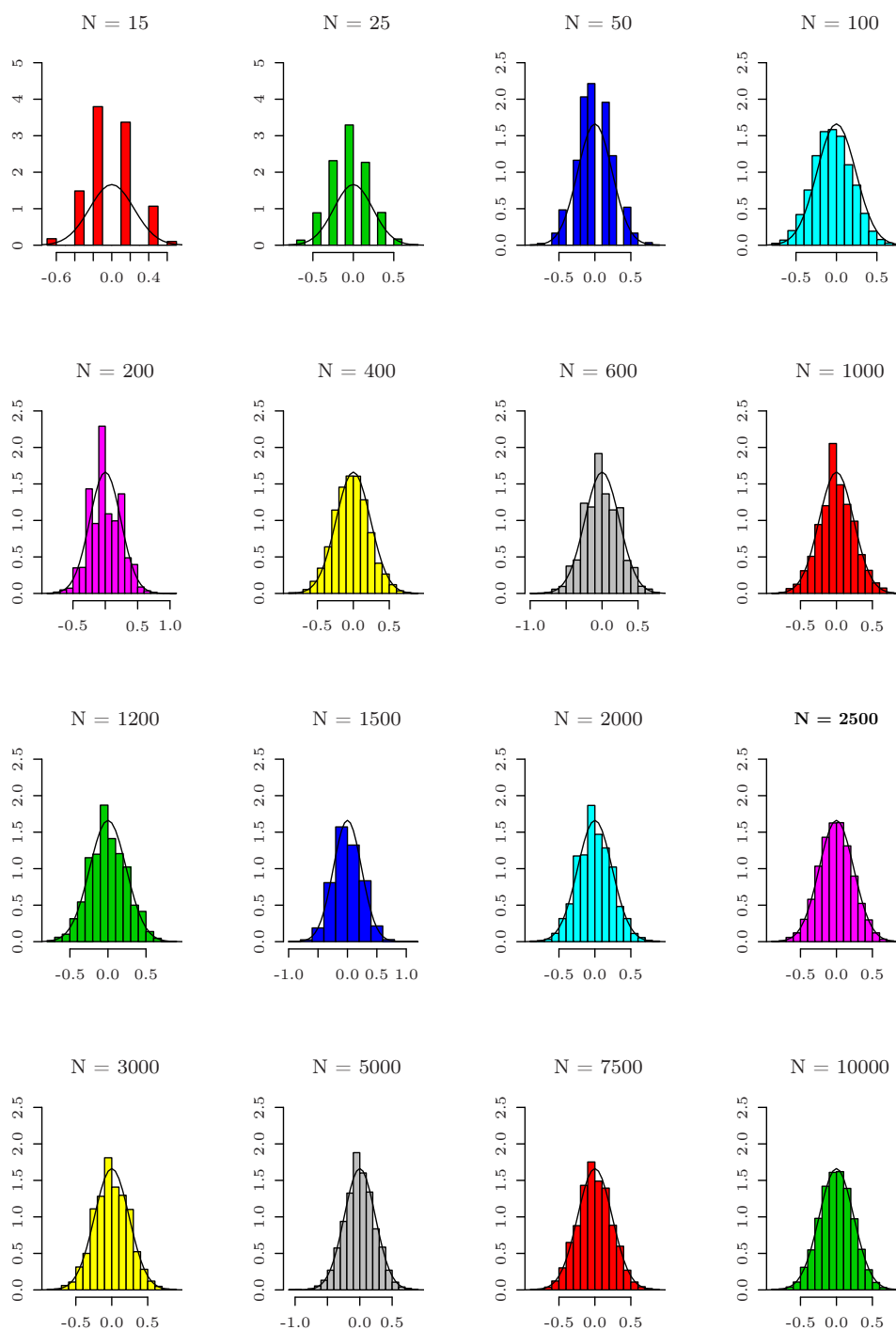
Obrázok 3.1: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(5)$ pre pevné $k = 2$ a pre rastúce hodnoty N .



Obrázok 3.2: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(5)$ pre pevné $p = \frac{2}{5}$ a pre rastúce hodnoty N .



Obrázok 3.3: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(5)$ pre pevné $p = \frac{2}{5}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



Obrázok 3.4: Histogramy ilustrujúce rýchlosť konvergencie k normálnemu rozdeleniu $N(0, \rho_2)$ pre pevné $k = 2$, pre pevné $p = \frac{2}{5}$ a pre rastúce hodnoty N .

A

APPENDIX

A.1 POUŽITÉ TVRDENIA

VETA A.1. *Nech X_n , $n \in \mathbb{N}$, X sú náhodné veličiny a $m_j^{(n)} = E (X_n)^j$, $m_j = E X^j$ ich momenty. Ak má mocninný rad*

$$m(z) = \sum_{j=0}^{\infty} \frac{m_j}{j!} z^j \quad \text{pre } z \in \mathbb{C}, \quad (\text{A-1})$$

kladný polomer konvergencie a

$$\lim_{n \rightarrow \infty} m_j^{(n)} = m_j \in \mathbb{R} \quad \text{pre } j \in \mathbb{N},$$

potom

$$X_n \xrightarrow[n \rightarrow \infty]{\mathcal{D}} X.$$

Dôkaz. Vid' (Štěpán, 1987), str. 229, dôkaz vety III.4.30.

Q.E.D.

Hovoríme, že náhodný proces $\{X_t, 0 \leq t \leq 1\}$ je *spojitý*, ak všetky jeho trajektórie sú spojité funkcie na $[0, 1]$. V tejto práci je pre nás zo spojitých náhodných procesov dôležitý Wienerov proces a predovšetkým Brownov most. Oba tieto procesy patria do rodiny spojitých centrovaných Gaussovských procesov. Náhodný proces $\{X_t, 0 \leq t \leq 1\}$ sa nazýva *Gaussovský*, ak pre každé $k \in \mathbb{N}$ a ľubovlnú voľbu $0 \leq t_1 < \dots < t_k \leq 1$ má vektor $(X_{t_1}, \dots, X_{t_k})$ k -rozmerné normálne rozdelenie. Proces $\{X_t, 0 \leq t \leq 1\}$ je *centrovaný*, ak $E|X_t| < \infty$ a $E X_t = 0$ pre každé t , $0 \leq t \leq 1$. Pre centrovaný proces definujeme kovariančnú funkciu ako $K(s, t) = E X_s X_t$ pre $0 \leq s, t \leq 1$.

DEFINÍCIA A.2. *Spojitély proces $\{W_t, 0 \leq t \leq 1\}$, kde $W_0 = 0$, sa nazýva Wienerov proces, ak je centrovany Gaussovský s kovariančnou funkciou*

$$K(s, t) = \min\{s, t\} = s \wedge t.$$

Ekvivalentnú definíciu poskytuje napríklad veta III.5.9 v (Štěpán, 1987) na strane 234:

VETA A.3. *Spojitély proces $\{W_t, 0 \leq t \leq 1\}$ je Wienerov práve vtedy, keď*

$$(i) \mathcal{L}(W_t - W_s) = N(0, |t - s|) \quad \text{pre } 0 \leq s, t \leq 1,$$

(ii) $W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}}$ sú nezávislé náhodné veličiny pre každú konečnú postupnosť $0 \leq t_1 < \dots < t_k \leq 1$, ide teda o proces s nezávislými prírastkami,

(iii) $W_0 = 0$.

DEFINÍCIA A.4. *Nech $\{W_t, 0 \leq t \leq 1\}$ je Wienerov proces. Brownov most (z bodu 0 do bodu 0) na intervale $[0, 1]$ je náhodný proces $\{W_t^0, 0 \leq t \leq 1\}$ definovaný vzťahom*

$$W_t^0 = W_t - tW_1, \quad 0 \leq t \leq 1.$$

Elementárnym výpočtom sa dá odvodiť

VETA A.5. *Brownov most je centrovany Gaussovský proces s kovariančnou funkciou*

$$K(s, t) = s(1 - t) \quad \text{pre } 0 \leq s \leq t \leq 1.$$

Pripomeňme známy pojem konvergenzie v distribúcii.

DEFINÍCIA A.6. *Nech X^n a X sú náhodné veličiny s hodnotami v metrickom priestore (S, d) . Hovoríme, že náhodné veličiny X^n konvergujú k veličine X v distribúcii, ak pre každú spojitú ohraničenú funkciu $H : S \mapsto \mathbb{R}$ platí*

$$\mathbb{E} H(X^n) \xrightarrow{n \rightarrow \infty} \mathbb{E} H(X).$$

Píšeme $X^n \xrightarrow[n \rightarrow \infty]{\mathcal{D}} X$ (v S).

Poznamenajme, že na spojitý náhodný proces $\{X_t, 0 \leq t \leq 1\}$ sa môžeme pozerat ako na náhodnú veličinu s hodnotami v priestore spojitých funkcií na intervale $[0, 1]$, označovanom $C[0, 1]$. Dôležitým priestorom pre konvergenziu v distribúcii je Skorochodov priestor, ktorý značíme $D[0, 1]$. Ide o priestor všetkých funkcií

$$f : [0, 1] \mapsto \mathbb{R},$$

ktoré sú zprava spojité na intervale $[0, 1]$ a majú konečné limity zľava v každom bode $t \in [0, 1]$. Funkcie f s takýmito vlastnosťami sú označované ako tzv. cadlag alebo corlol funkcie.

DEFINÍCIA A.7. Hovoríme, že funkcia

$$\lambda : [0, 1] \mapsto [0, 1]$$

je transformácia času, ak

$$\lambda(0) = 0, \quad \lambda(1) = 1 \quad \text{a} \quad \lambda \quad \text{je rastúca a spojitá.}$$

DEFINÍCIA A.8. V $D[0, 1]$ uvažujeme konvergenciu $f_n \xrightarrow{s} f$, ktorá je definovaná nasledovne: existujú transformácie času λ_n také, že

$$\max_{0 \leq t \leq 1} |\lambda_n(t) - t| \rightarrow 0 \quad \text{a} \quad \max_{0 \leq t \leq 1} |f_n(\lambda_n(t)) - f(t)| \rightarrow 0 \quad \text{pre} \quad n \rightarrow \infty \quad (\text{A-2})$$

alebo ekvivalentne: existujú transformácie času λ_n také, že

$$\max_{0 \leq t \leq 1} |\lambda_n(t) - t| \rightarrow 0 \quad \text{a} \quad \max_{0 \leq t \leq 1} |f_n(t) - f(\lambda_n(t))| \rightarrow 0 \quad \text{pre} \quad n \rightarrow \infty. \quad (\text{A-3})$$

PRÍKLAD A.9. Nech $f_n, f \in D[0, 1]$. Ak $f_n \rightarrow f$ rovnomerne na intervale $[0, 1]$, potom $f_n \xrightarrow{s} f$. (Volíme $\lambda_n(t) = t$.)

PRÍKLAD A.10. Definujme

$$f_n(t) = I_{[0, \frac{1}{2} + \frac{1}{n})}(t) \quad \text{a} \quad f(t) = I_{[0, \frac{1}{2})}(t) \quad \text{pre} \quad t \in [0, 1].$$

Volíme transformáciu času $\lambda_n(t)$ ako po dvoch častiach lineárnu funkciu takú, že $\lambda_n(\frac{1}{2}) = \frac{1}{2} + \frac{1}{n}$, teda

$$\lambda_n(t) = \begin{cases} (1 + \frac{2}{n})t & \text{pre} \quad 0 \leq t < \frac{1}{2}, \\ (1 - \frac{2}{n})t + \frac{2}{n} & \text{pre} \quad \frac{1}{2} \leq t \leq 1. \end{cases}$$

Potom zrejme je $f_n(\lambda_n(t)) = f(t)$ a $f_n \xrightarrow{s} f$. Avšak napriek tomu, že $f_n \xrightarrow{s} f$, máme

$$\lim_{n \rightarrow \infty} f_n\left(\frac{1}{2}\right) = 1 \neq 0 = f\left(\frac{1}{2}\right).$$

Dôležitá je nasledujúca vlastnosť.

VETA A.11. Nech $f_n \xrightarrow{s} f$ a f buď spojitá funkcia. Potom

$$\max_{0 \leq t \leq 1} |f_n(t) - f(t)| \rightarrow 0.$$

Dôkaz. Zvoľme $\lambda_n(t)$ ako v (A-3). Potom pre $n \rightarrow \infty$

$$\max_{0 \leq t \leq 1} |f_n(t) - f(t)| \leq \max_{0 \leq t \leq 1} |f_n(t) - f(\lambda_n(t))| + \max_{0 \leq t \leq 1} |f(\lambda_n(t)) - f(t)| \rightarrow 0,$$

pretože $\max_{0 \leq t \leq 1} |\lambda_n(t) - t| \rightarrow 0$ a f je rovnomerne spojitá na $[0, 1]$. Q.E.D.

POZNÁMKA. Poznamenajme, že príklad A.9 spolu s vetou A.11 hovoria, že pre $f_n, f \in C[0, 1]$ platí, že

$$\max_{0 \leq t \leq 1} |f_n(t) - f(t)| \rightarrow 0 \quad \Leftrightarrow \quad f_n \xrightarrow{s} f.$$

VETA A.12. *Nech $f_n \xrightarrow{s} f$ a $t_n \rightarrow t$, $0 \leq t_n, t \leq 1$. Ak je f spojitá funkcia, potom*

$$\lim_{n \rightarrow \infty} f_n(t_n) = f(t).$$

Dôkaz. Z vety A.11 a z predpokladu spojitosti funkcie f plynie, že

$$\begin{aligned} |f_n(t_n) - f(t)| &\leq |f_n(t_n) - f(t_n)| + |f(t_n) - f(t)| \\ &\leq \max_{0 \leq t \leq 1} |f_n(t) - f(t)| + |f(t_n) - f(t)| \rightarrow 0 \end{aligned}$$

pre $n \rightarrow \infty$, a tým je dôkaz hotový.

Q.E.D.

Podstatné je však nasledujúce netriviálne zistenie (viď veta A.2.2 na str. 458 v (Kallenberg, 1997)):

VETA A.13. *Na priestore $D = D[0, 1]$ možno definovať metriku $d_s(f, g)$ tak, že (D, d_s) je separabilný úplný metrický priestor a*

$$f_n \xrightarrow{s} f \quad \Leftrightarrow \quad d_s(f_n, f) \rightarrow 0.$$

Každý stochastický proces s trajektóriami v $D[0, 1]$ je náhodná veličina s hodnotami v metrickom priestore (D, d_s) . Môžeme teda vyšetriť konvergenciu $Y^n \xrightarrow{\mathcal{D}} Y$, kde $Y^n = (Y_t^n, 0 \leq t \leq 1)$ a $Y = (Y_t, 0 \leq t \leq 1)$ sú procesy s trajektóriami v $D[0, 1]$.

Platia nasledujúce vety, ktoré sú pre nás v tejto práci užitočné:

VETA A.14. *Nech $Y^n = \{Y_t^n, 0 \leq t \leq 1\}$, $Y = \{Y_t, 0 \leq t \leq 1\}$ sú náhodné veličiny s hodnotami v $D[0, 1]$ a nech N^n a N sú reálne náhodné veličiny také, že*

$$Y^n \xrightarrow{\mathcal{D}} Y \quad \text{v} \quad D[0, 1]$$

a

$$N^n \xrightarrow{\mathcal{D}} N \quad \text{v} \quad \mathbb{R}.$$

Nech pre každé $n \in \mathbb{N}$ platí, že

$$Y^n \quad \text{a} \quad N^n \quad \text{sú} \quad \text{nezávislé} \quad \text{veličiny}.$$

Potom

$$Y^n + N^n \xrightarrow{\mathcal{D}} Y + N \quad \text{v} \quad D[0, 1],$$

kde Y a N sú nezávislé veličiny.

Dôkaz. Plynie z III.4.14 na strane 222 v (Štěpán, 1987). Q.E.D.

VĚTA A.15. *Nech $Y^n = \{Y_t^n, 0 \leq t \leq 1\}$, $Y = \{Y_t, 0 \leq t \leq 1\}$ sú náhodné veličiny s hodnotami v $D[0, 1]$. Nech*

$$Y^n \xrightarrow{\mathcal{D}} Y \quad \text{v} \quad D[0, 1] \quad \text{a} \quad Y \quad \text{je} \quad \text{spojitý} \quad \text{proces.}$$

Potom

$$(Y_{t_1}^n, Y_{t_2}^n, \dots, Y_{t_k}^n) \xrightarrow{\mathcal{D}} (Y_{t_1}, Y_{t_2}, \dots, Y_{t_k}) \quad \text{v} \quad \mathbb{R}^k$$

pre každú konečnú postupnosť $0 \leq t_1 < t_2 < \dots < t_k \leq 1$.

Dôkaz. Zvolíme $0 \leq t_1 < t_2 < \dots < t_k \leq 1$ pevne a definujeme zobrazenie

$$F : D[0, 1] \mapsto \mathbb{R}^k$$

predpisom

$$F(f) = (f(t_1), \dots, f(t_k)).$$

Z vety A.11 plynie, že $C[0, 1] \subset D[0, 1]$ je množina bodov spojitosti zobrazenia F . Pretože $P[Y \in C[0, 1]] = 1$, plynie naše tvrdenie z vety III.4.12 v (Štěpán, 1987). Q.E.D.

VĚTA A.16. *Nech $Y^n = \{Y_t^n, 0 \leq t \leq 1\}$, $Y = \{Y_t, 0 \leq t \leq 1\}$ sú náhodné veličiny s hodnotami v $D[0, 1]$. Nech*

$$Y^n \xrightarrow{\mathcal{D}} Y \quad \text{v} \quad D[0, 1] \quad \text{a} \quad Y \quad \text{je} \quad \text{spojitý} \quad \text{proces.}$$

Uvažujme náhodnú veličinu T^n takú, že

$$0 \leq T^n \leq 1 \quad \text{a} \quad T^n \xrightarrow[n \rightarrow \infty]{\mathcal{P}} a \in [0, 1].$$

Potom

$$Y_{T^n}^n \xrightarrow{\mathcal{D}} Y_a \quad \text{v} \quad \mathbb{R}.$$

Dôkaz. Z vety III.4.28 v (Štěpán, 1987) plynie

$$(Y^n, T^n) \xrightarrow{\mathcal{D}} (Y, a).$$

Definujeme zobrazenie

$$F : D[0, 1] \times \mathbb{R} \mapsto \mathbb{R}$$

predpisom

$$F(f, t) = f(t).$$

Z vety A.12 plynie, že bod (f, t) je bodom spojitosti zobrazenia F , ak f je spojitá funkcia na $[0, 1]$. Veta III.4.12 v (Štěpán, 1987) potom implikuje (keďže $P[Y \in C[0, 1]] = 1$), že

$$Y_{T^n}^n = Y^n(T^n) = F(Y^n, T^n) \xrightarrow{\mathcal{D}} F(Y, a) = Y(a) = Y_a \quad \text{v } \mathbb{R},$$

a tým je veta dokázaná.

Q.E.D.

Pre procesy s trajektóriami v $D[0, 1]$ platí nasledujúca dvojica hlbokých rozšírení centrálnej limitnej vety.

VETA A.17 (PRINCÍP INVARIANCIE PRE I.I.D. NÁHODNÉ VELIČINY). *Nech*

$$\xi_1, \xi_2, \dots, \xi_n$$

sú i.i.d. náhodné veličiny s

$$E \xi_j = 0 \quad \text{a} \quad E \xi_j^2 = \sigma^2 \in (0, \infty) \quad \text{pre } 1 \leq j \leq n.$$

Definujme náhodný proces $X^n = \{X_t^n, 0 \leq t \leq 1\}$ s trajektóriami v priestore $D[0, 1]$ predpisom

$$X_t^n = \begin{cases} \frac{1}{\sigma\sqrt{n}} \sum_{j=1}^{[nt]} \xi_j & \text{pre } \frac{1}{n} \leq t \leq 1, \\ 0 & \text{pre } 0 \leq t < \frac{1}{n}, \end{cases}$$

potom

$$X^n \xrightarrow[n \rightarrow \infty]{\mathcal{D}} W,$$

kde W je Wienerov proces (na intervale $[0, 1]$).

Dôkaz. Vid' napríklad (Billingsley, 1968), str. 137, dôkaz vety 16.1. Q.E.D.

VETA A.18 (PRINCÍP INVARIANCIE PRE PROSTÝ NÁHODNÝ VÝBER). *Nech pre $n \in \mathbb{N}$ je*

$$x_{n1}, x_{n2}, \dots, x_{nk_n}$$

postupnosť reálnych čísel takých, že

$$\sum_{i=1}^{k_n} x_{ni} = 0, \quad \sum_{i=1}^{k_n} x_{ni}^2 = 1$$

a

$$\max_{1 \leq i \leq k_n} |x_{ni}| \rightarrow 0 \quad \text{pre } n \rightarrow \infty.$$

Ďalej buď

$$\xi_{n1}(\omega), \xi_{n2}(\omega), \dots, \xi_{nk_n}(\omega)$$

náhodná permutácia $x_{n1}, x_{n2}, \dots, x_{nk_n}$, pričom každé z takýchto $k_n!$ náhodných poradí má rovnakú pravdepodobnosť (rovnú $\frac{1}{k_n!}$). Definujme náhodný proces $X^n = \{X_t^n, 0 \leq t \leq 1\}$ s trajektóriami v priestore $D[0, 1]$ vzťahom

$$X_t^n = \begin{cases} \sum_{i=1}^{[k_n t]} \xi_{ni}(\omega) & \text{pre } \frac{1}{k_n} \leq t \leq 1, \\ 0 & \text{pre } 0 \leq t < \frac{1}{k_n}. \end{cases}$$

Potom

$$X^n \xrightarrow[n \rightarrow \infty]{\mathcal{D}} W^0,$$

kde W^0 je Brownov most (z 0 do 0) na intervale $[0, 1]$.

Dôkaz. Vid' (Billingsley, 1968), str. 209, dôkaz vety 24.1.

Q.E.D.

A.2 ZDROJOVÝ KÓD

Nižšie uvádzame zdrojové kódy funkcií použitých v našej simulácii v R.

```
cnk <- function(p.N, p.n, p.k) {
  .vyber <- matrix(0,nrow=p.k,ncol=p.n)
  if (p.n==1) {
    for (i in 1:p.k) {
      .vyber[i] <- sample(1:p.N,size=p.n,replace = FALSE);
      if (i>1) .prienik <- intersect(.prienik,.vyber[i])
      else .prienik <- .vyber[1]
    }
  }
  else {
    for (i in 1:p.k) {
      .vyber[i,] <- sample(1:p.N,size=p.n,replace = FALSE);
      if (i>1) .prienik <- intersect(.prienik,.vyber[i,])
      else .prienik <- .vyber[1,]
    }
  }
  return(length(.prienik))
}
```

```
poisson.1 <- function(p.N, p.lambda, p.p) {
  .n <- round(p.p*p.N)
  .k <- round(log(p.lambda/p.N)/log(p.p))
  .k <- max(.k,1)
  return(cnk(p.N,.n,.k))
}
```

```
poisson.2 <- function(p.N, p.lambda, p.k) {
  .p <- (p.lambda/p.N)^(1/p.k)
  .n <- round(.p*p.N)
  return(cnk(p.N,.n,p.k))
}
```

```
poisson.sample <- function(p.N, p.lambda, p.size, p.fix, p.pk) {
  # p.fix 1 - fix p
  #       2 - fix k
  .sample <- numeric(p.size)
  for (i in 1:p.size) {
```

```
    if (p.fix == 1) {
      .sample[i] <- poisson.1(p.N,p.lambda,p.pk)
    }
    else {
      .sample[i] <- poisson.2(p.N,p.lambda,p.pk)
    }
  }
  return(.sample)
}
```

```
normal.sample <- function(p.N, p.size, p.k, p.p) {
  .sample <- numeric(p.size)
  .n <- round(p.p*p.N)
  for (i in 1:p.size) {
    .sample[i] <- cnk(p.N,.n,p.k)
  }
  .sample <- (.sample - p.N*(p.p^p.k))/sqrt(p.N)
  return(.sample)
}
```

LITERATÚRA

BILLINGSLEY, P. (1968): *Convergence of Probability Measures*. John Wiley & Sons, New York.

HURT, J., MACHEK, J., ŠTĚPÁN, J. A VORLÍČKOVÁ, D. (1982): The Intersections of Random Finite Sets. *Mathematica Slovaca*, 32(3), 229–237.

JOHNSON, B. R. (1980): An Elementary Proof of Inclusion–Exclusion Formulas. *The American Mathematical Monthly*, 87(9), 750–751.

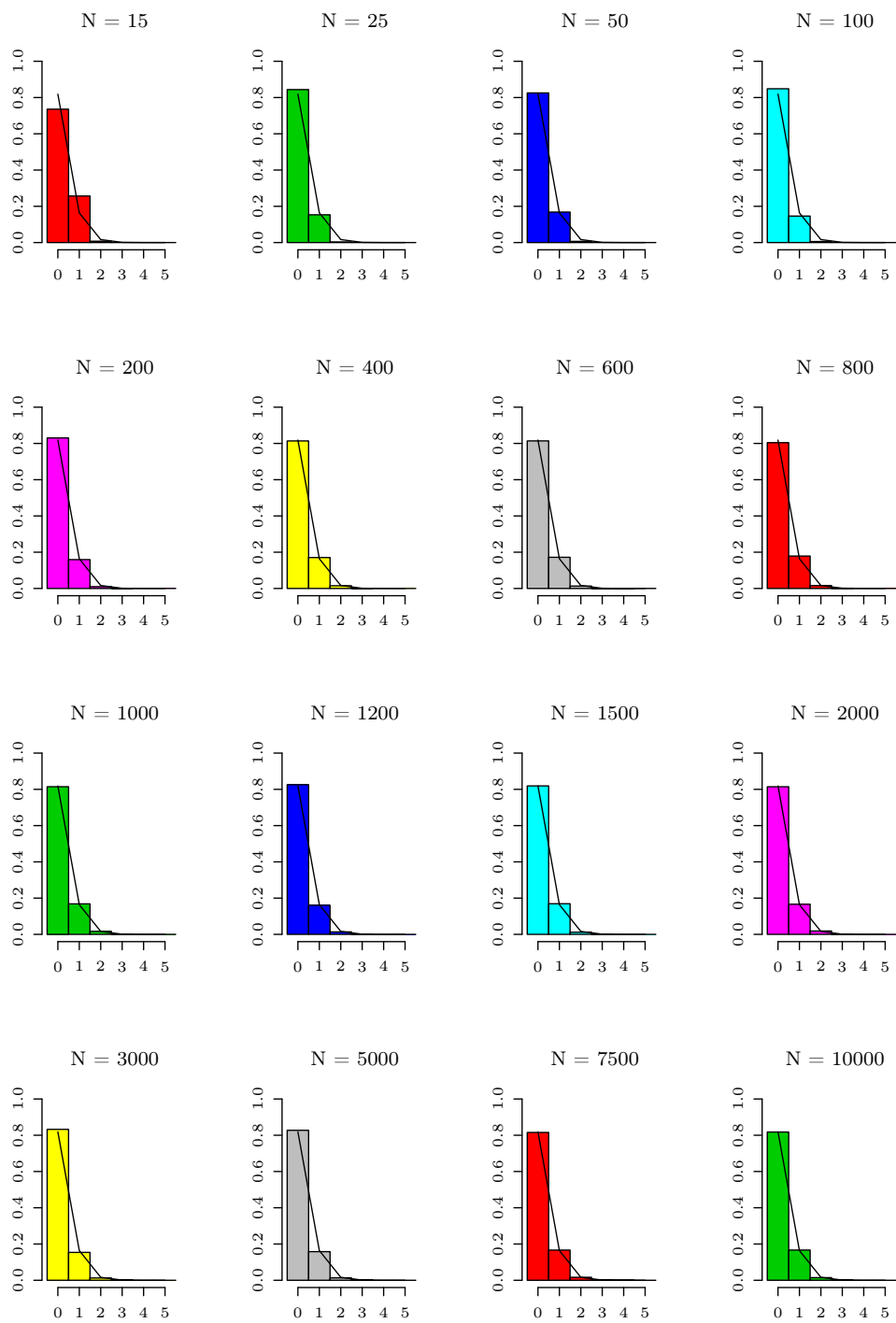
JOHNSON, N. L., KEMP, A. W. A KOTZ, S. (2005): *Univariate Discrete Distributions*. John Wiley & Sons, Hoboken (NJ), 3. edícia.

KALLENBERG, O. (1997): *Foundations of Modern Probability*. Springer, New York.

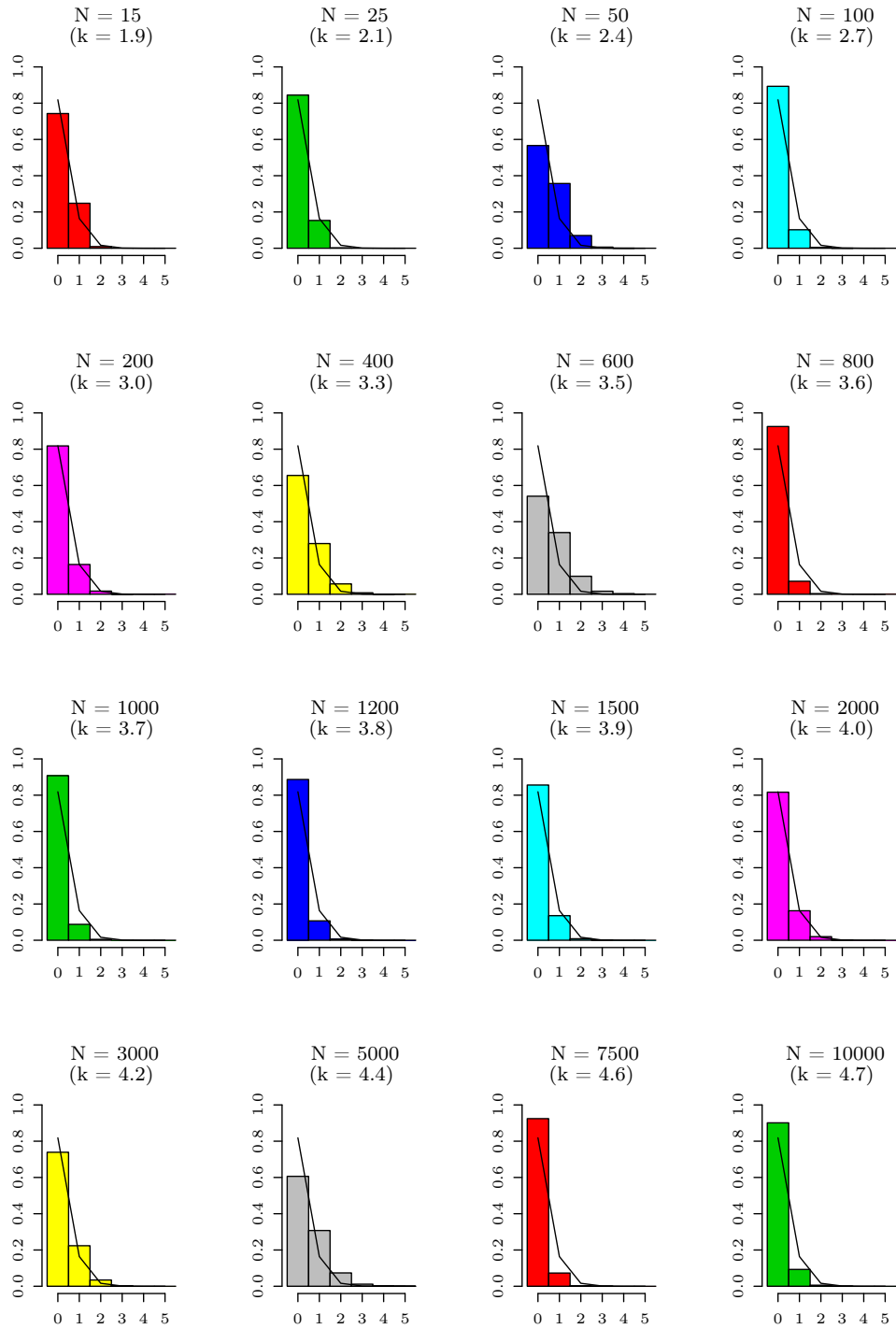
RÉNYI, A. (1972): *Teorie pravděpodobnosti*. Academia, Praha.

ŠTĚPÁN, J. (1987): *Teorie pravděpodobnosti*. Academia, Praha.

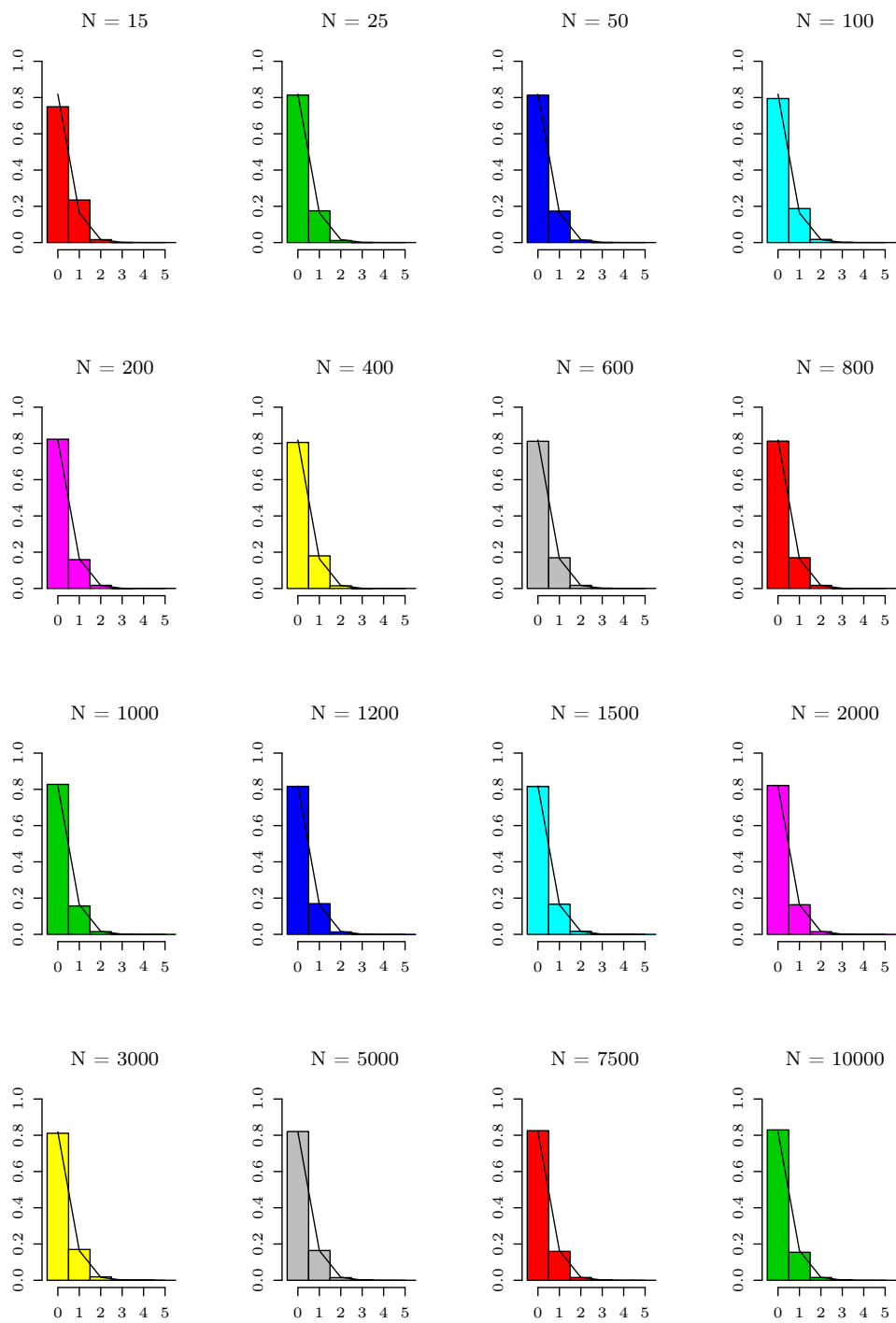
VALÁŠKOVÁ, Z. (2007): *Nespravedlivá losování*. Bakalárska práca, Univerzita Karlova v Praze.



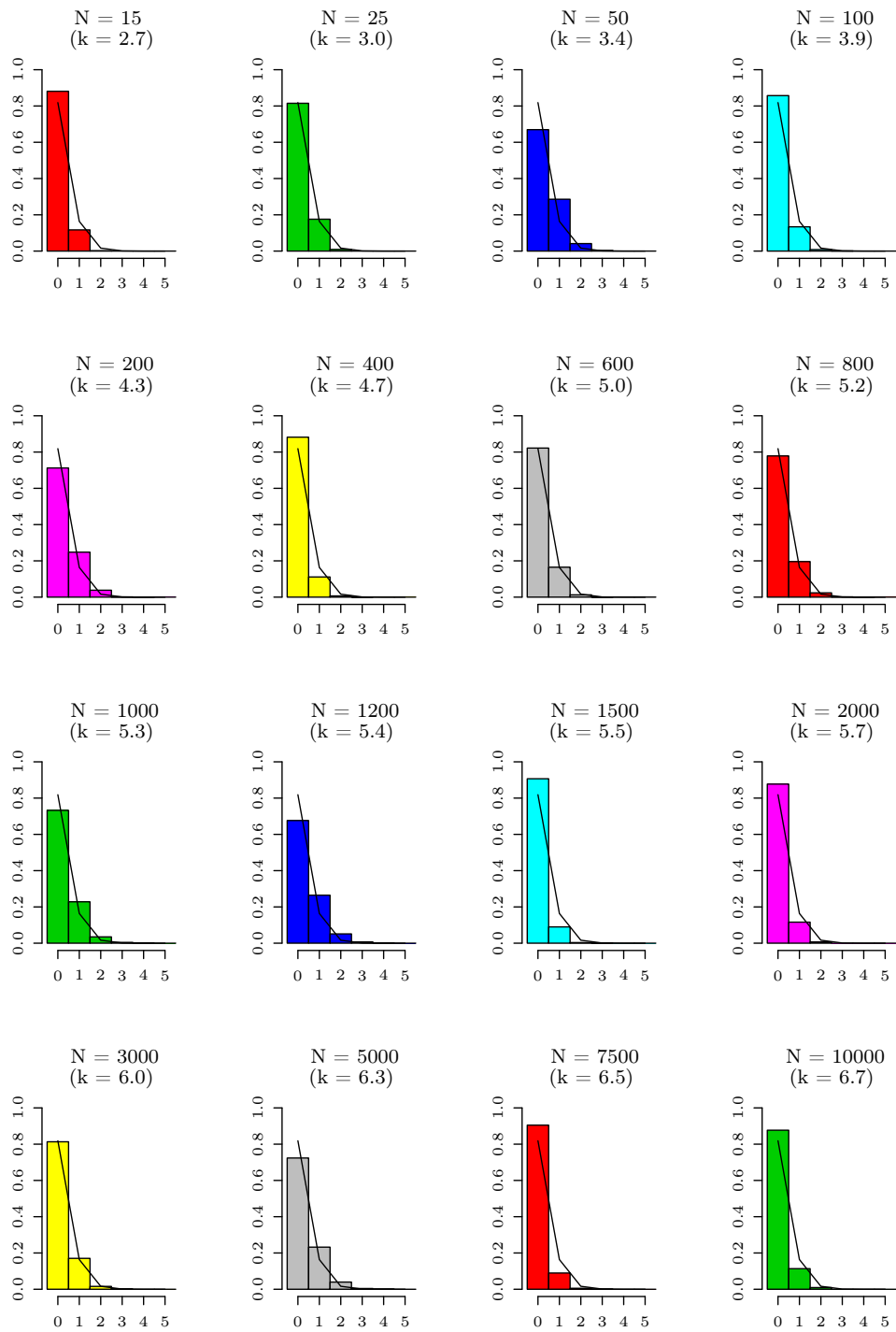
Obrázok 1: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(0.2)$ pre pevné $k = 2$ a pre rastúce hodnoty N .



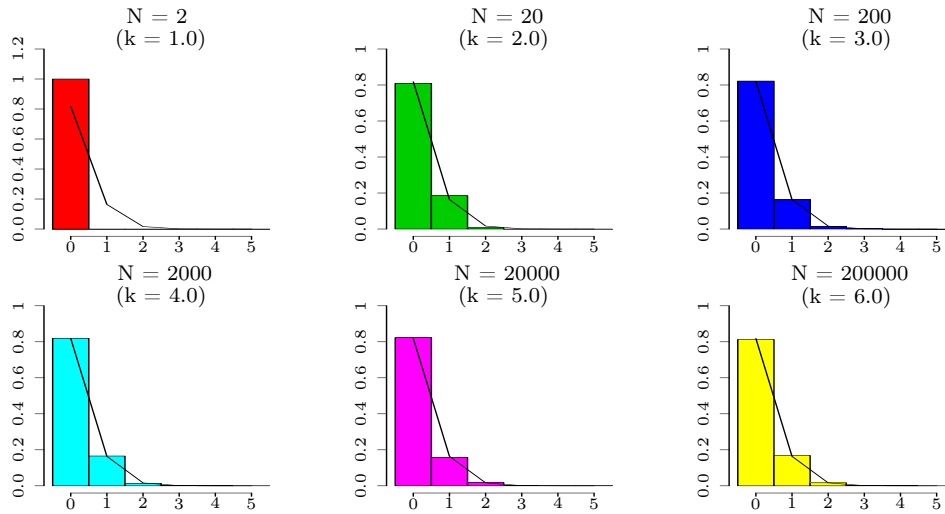
Obrázok 2: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(0.2)$ pre pevné $p = \frac{1}{10}$ a pre rastúce hodnoty N .



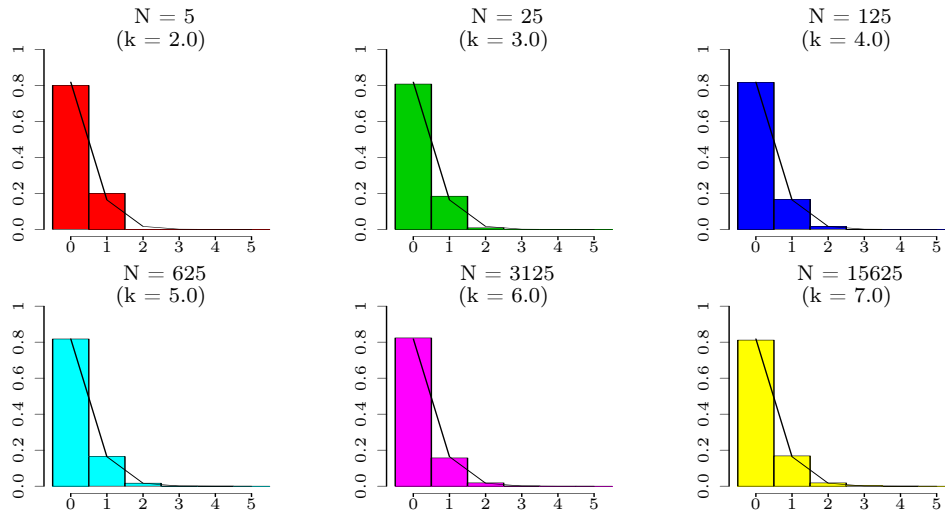
Obrázok 3: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(0.2)$ pre pevné $k = 3$ a pre rastúce hodnoty N .



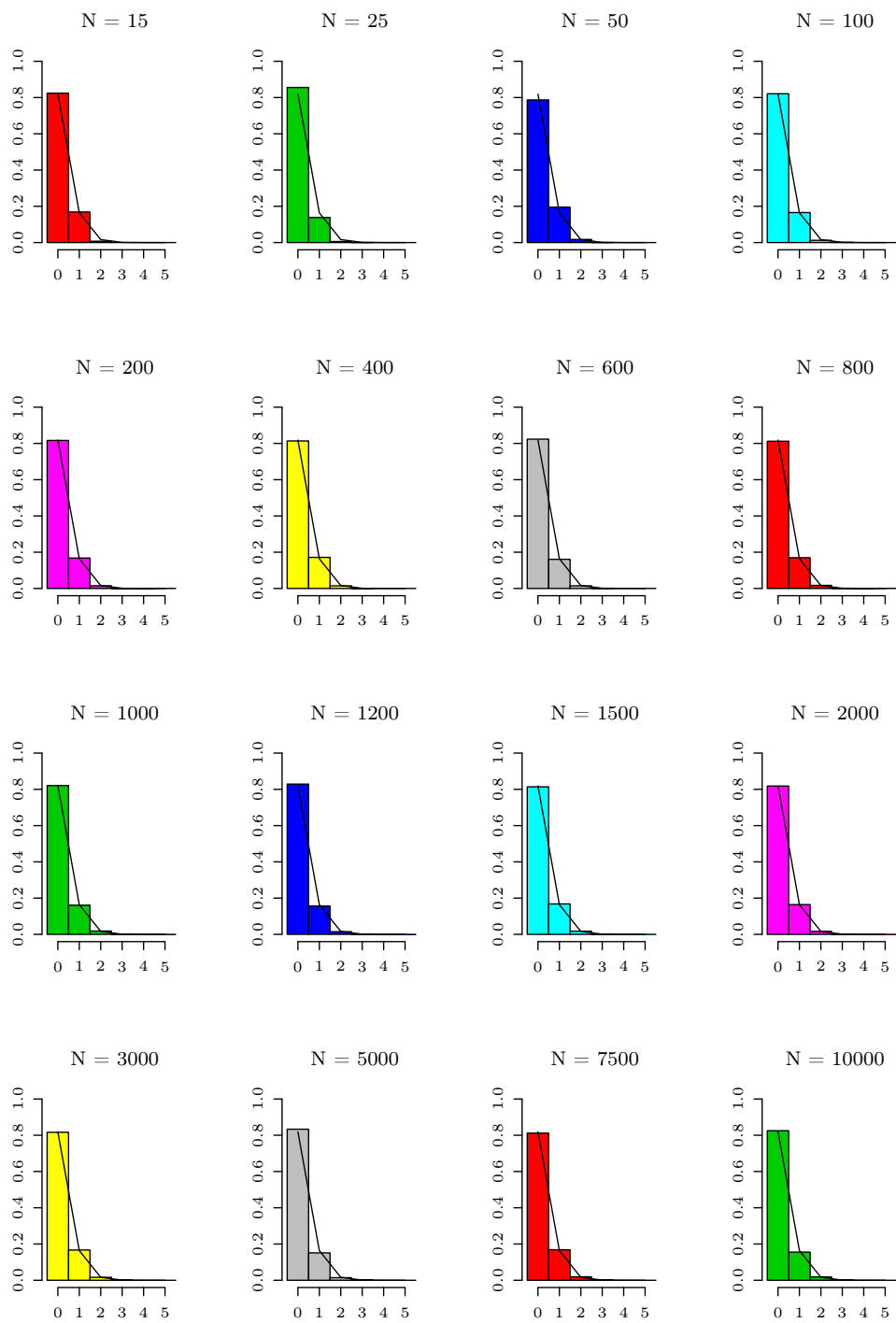
Obrázok 4: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(0.2)$ pre pevné $p = \frac{1}{5}$ a pre rastúce hodnoty N .



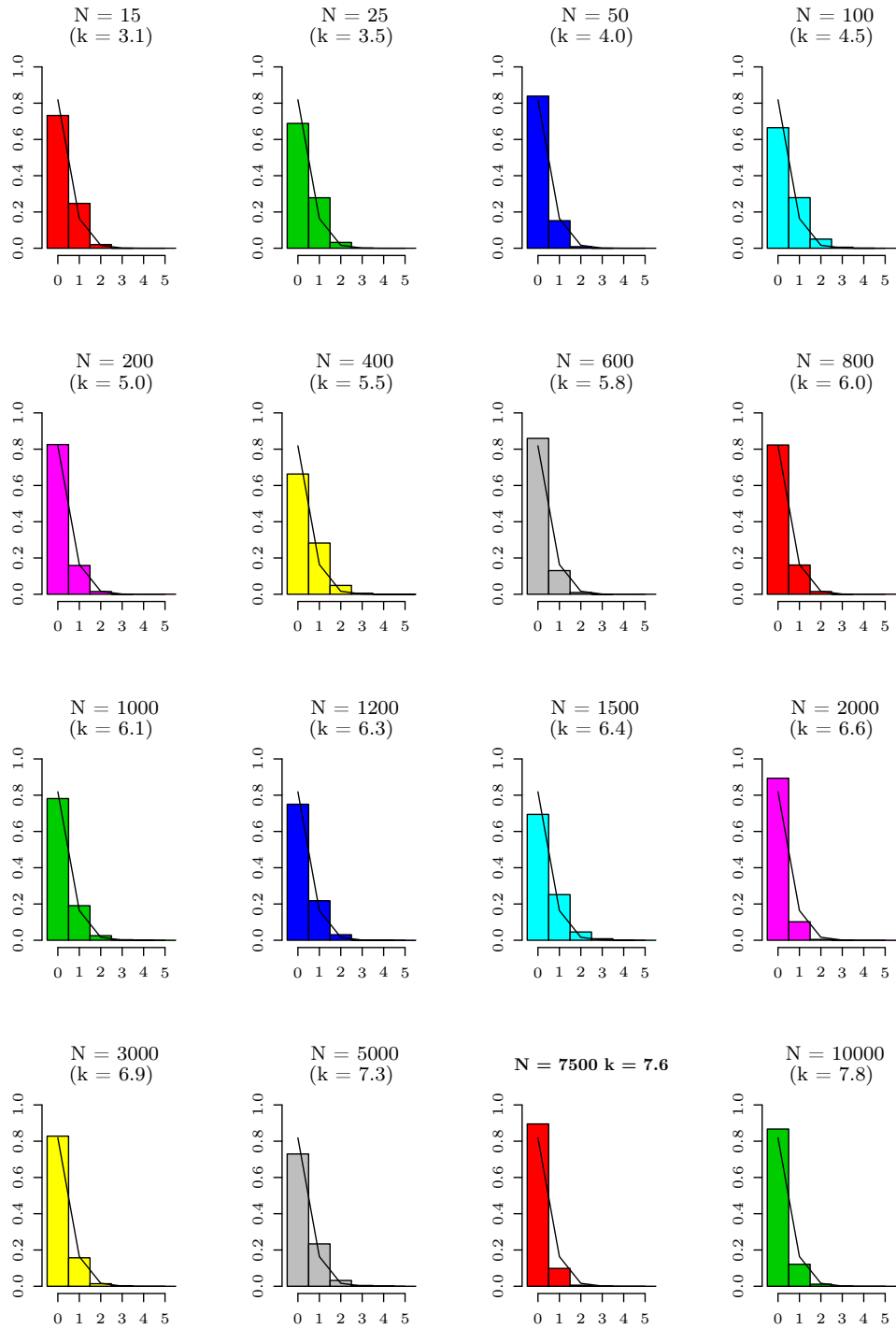
Obrázok 5: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(0.2)$ pre pevné $p = \frac{1}{10}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



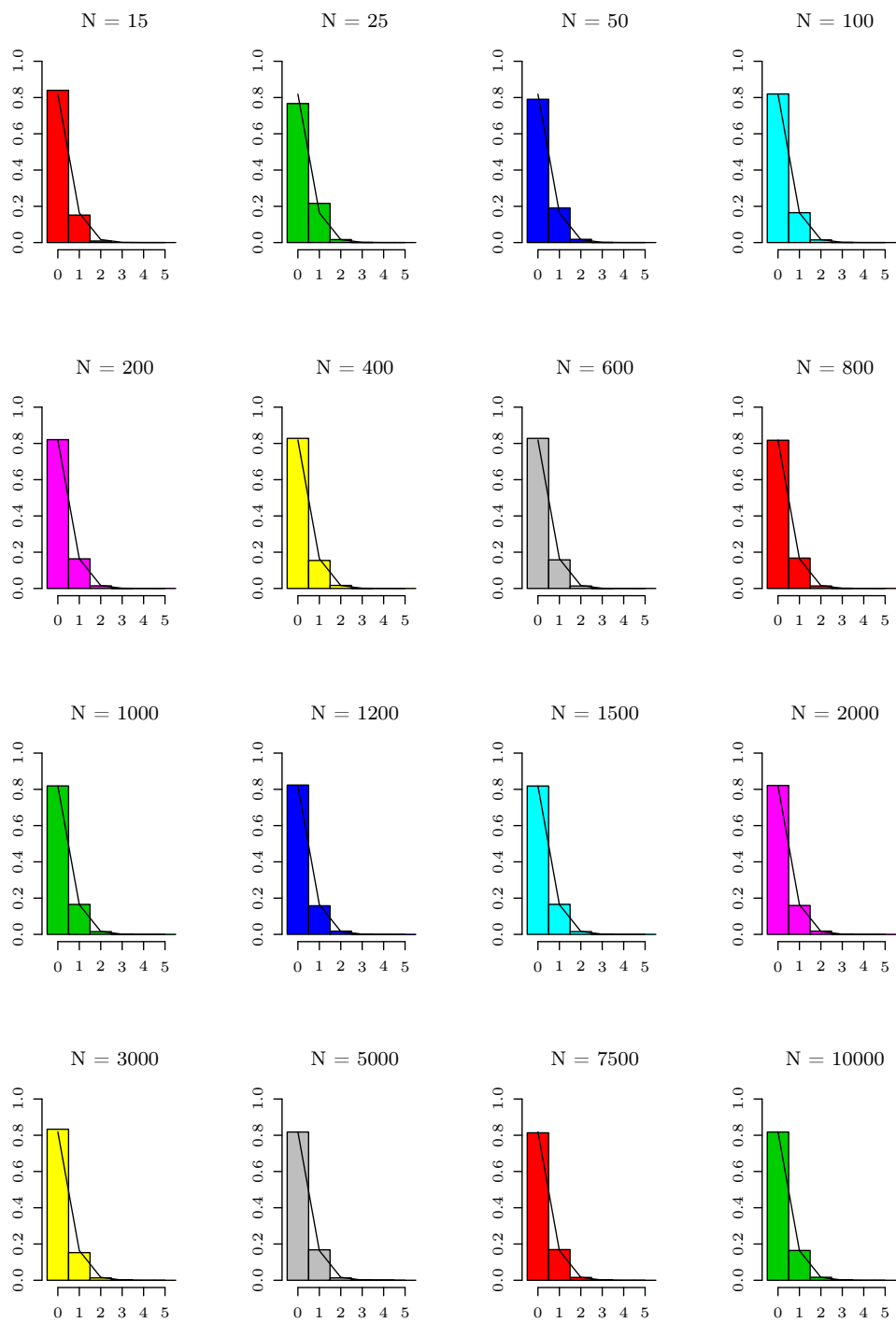
Obrázok 6: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(0.2)$ pre pevné $p = \frac{1}{5}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



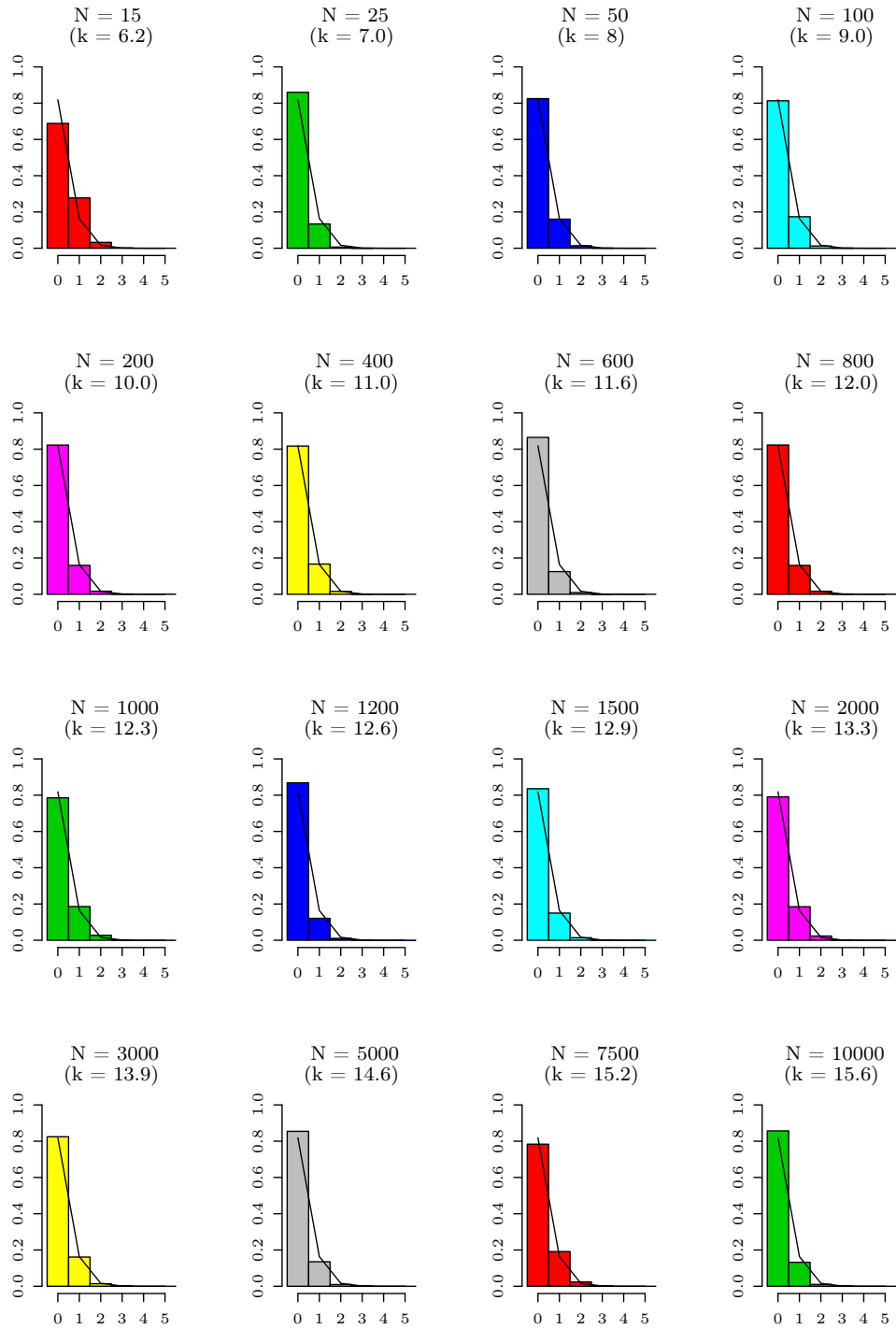
Obrázok 7: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(0.2)$ pre pevné $k = 4$ a pre rastúce hodnoty N .



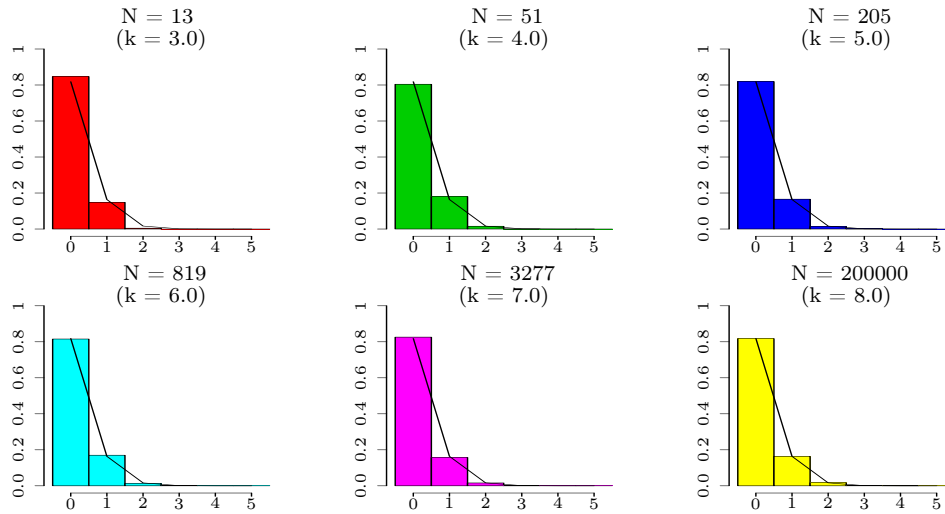
Obrázok 8: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(0.2)$ pre pevné $p = \frac{1}{4}$ a pre rastúce hodnoty N .



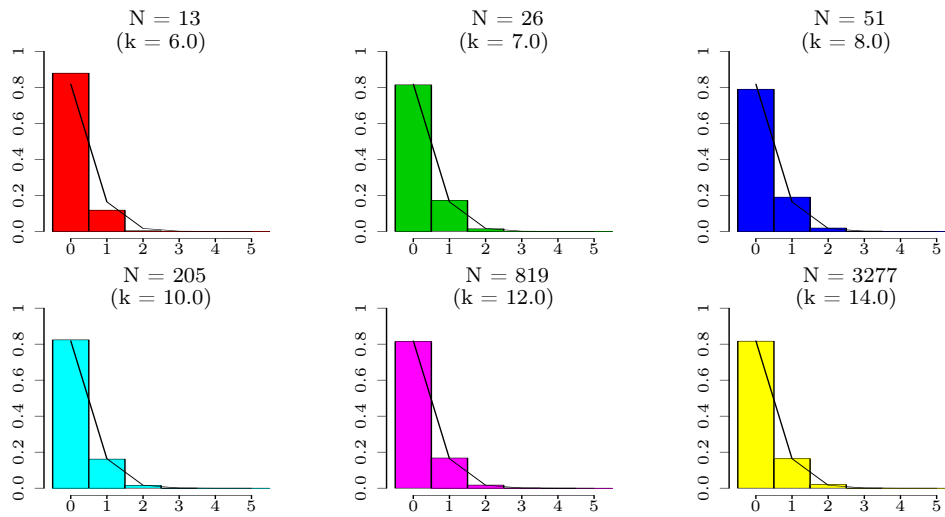
Obrázok 9: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(0.2)$ pre pevné $k = 5$ a pre rastúce hodnoty N .



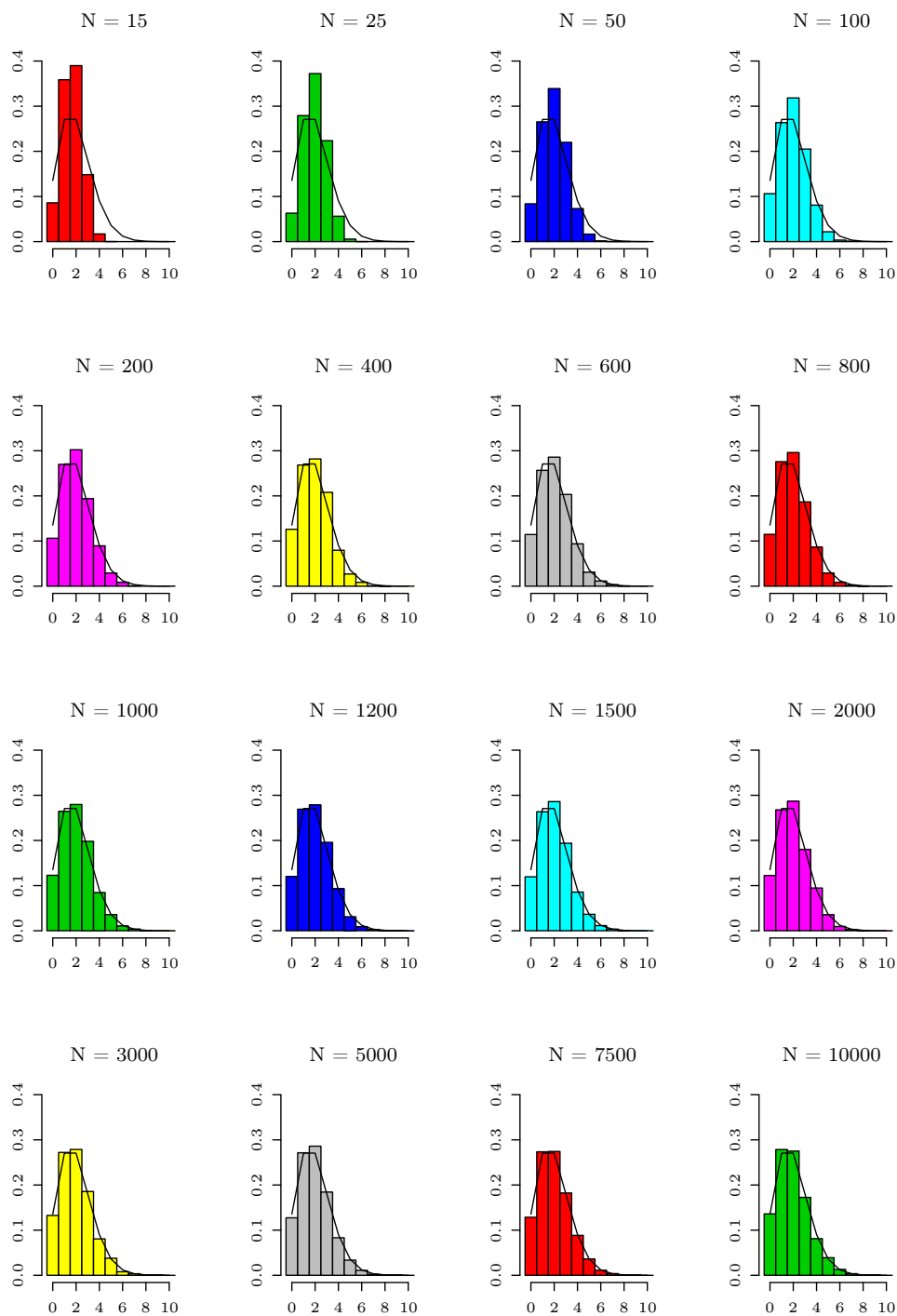
Obrázok 10: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(0.2)$ pre pevné $p = \frac{1}{2}$ a pre rastúce hodnoty N .



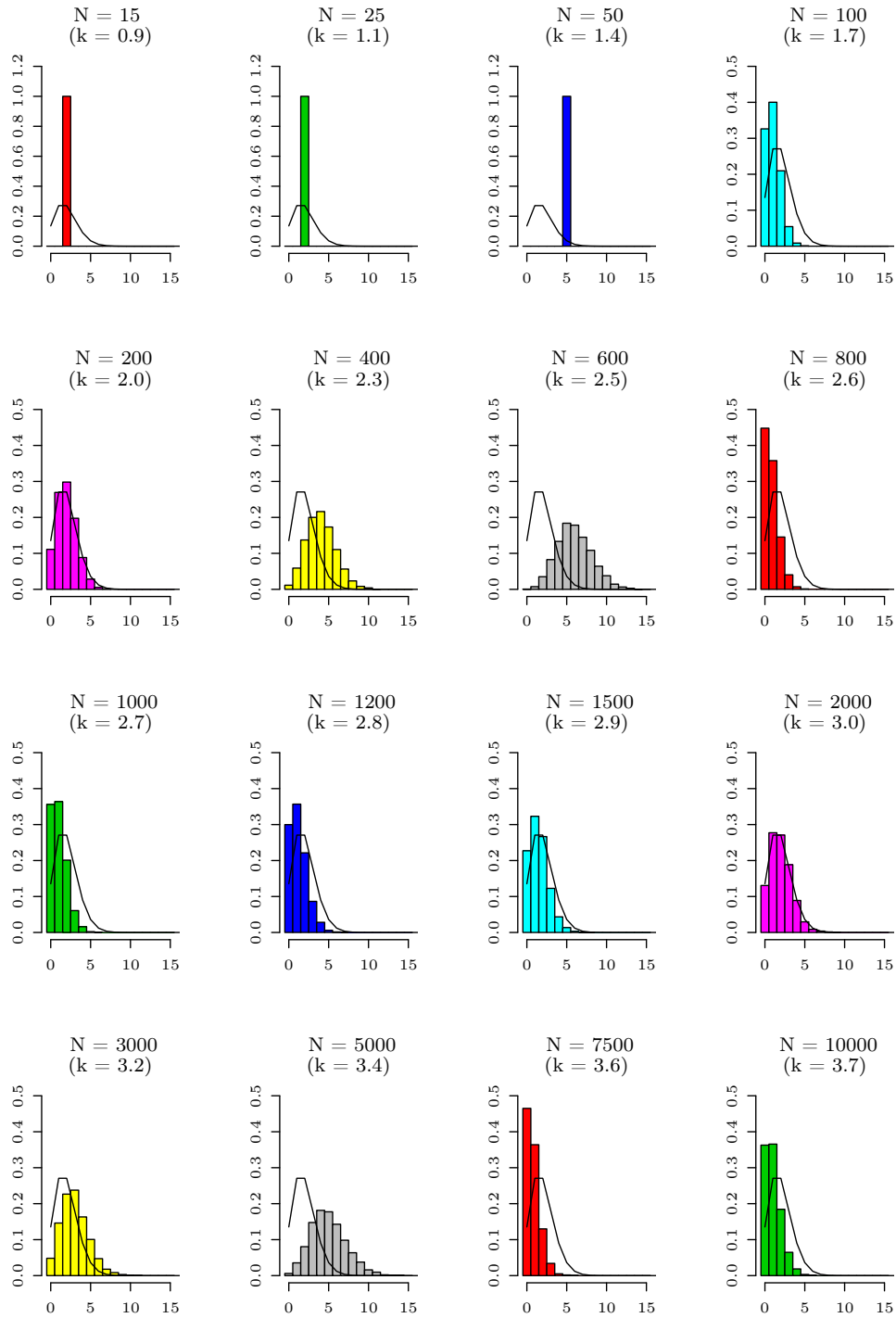
Obrázok 11: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(0.2)$ pre pevné $p = \frac{1}{4}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



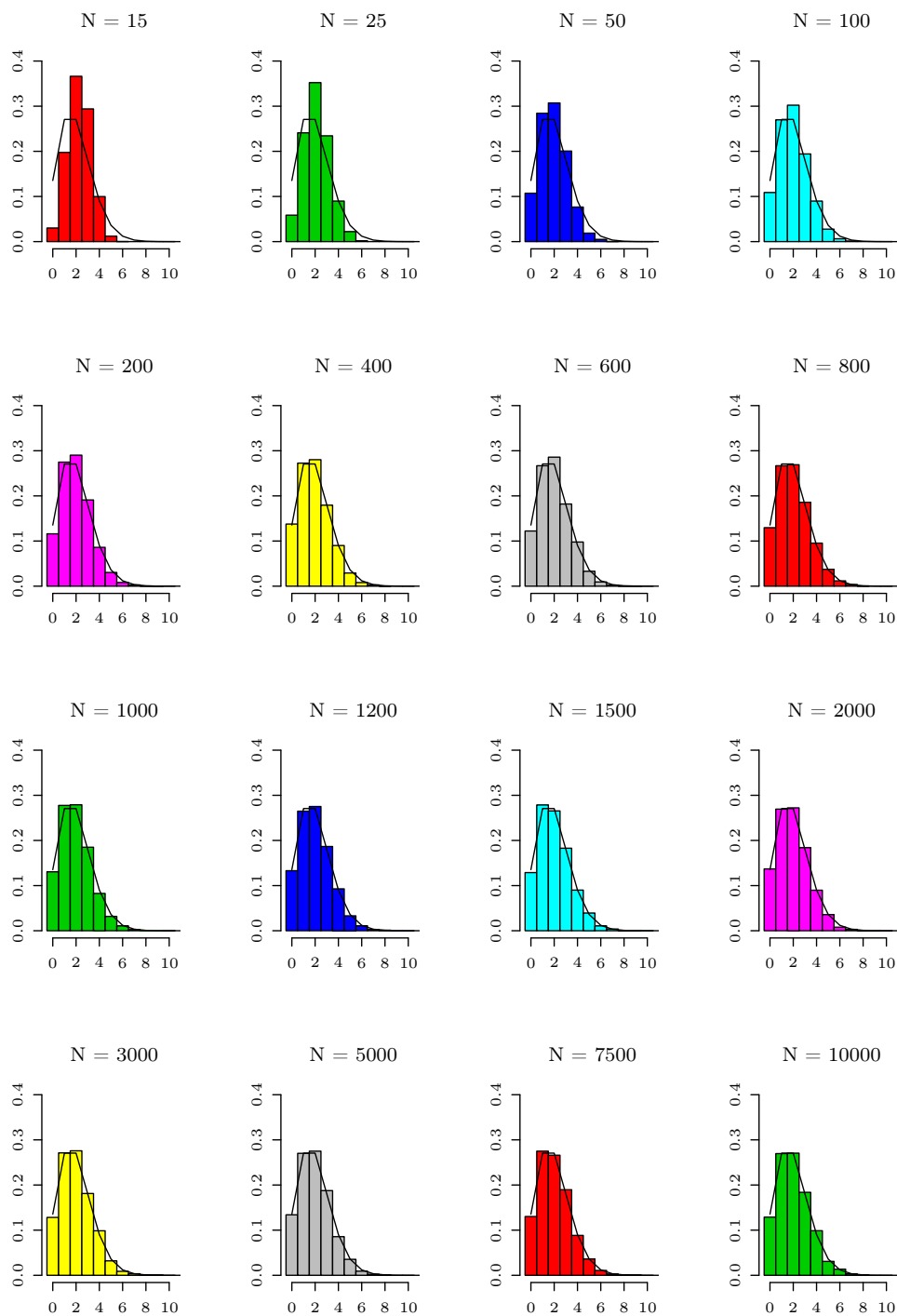
Obrázok 12: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(0.2)$ pre pevné $p = \frac{1}{2}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



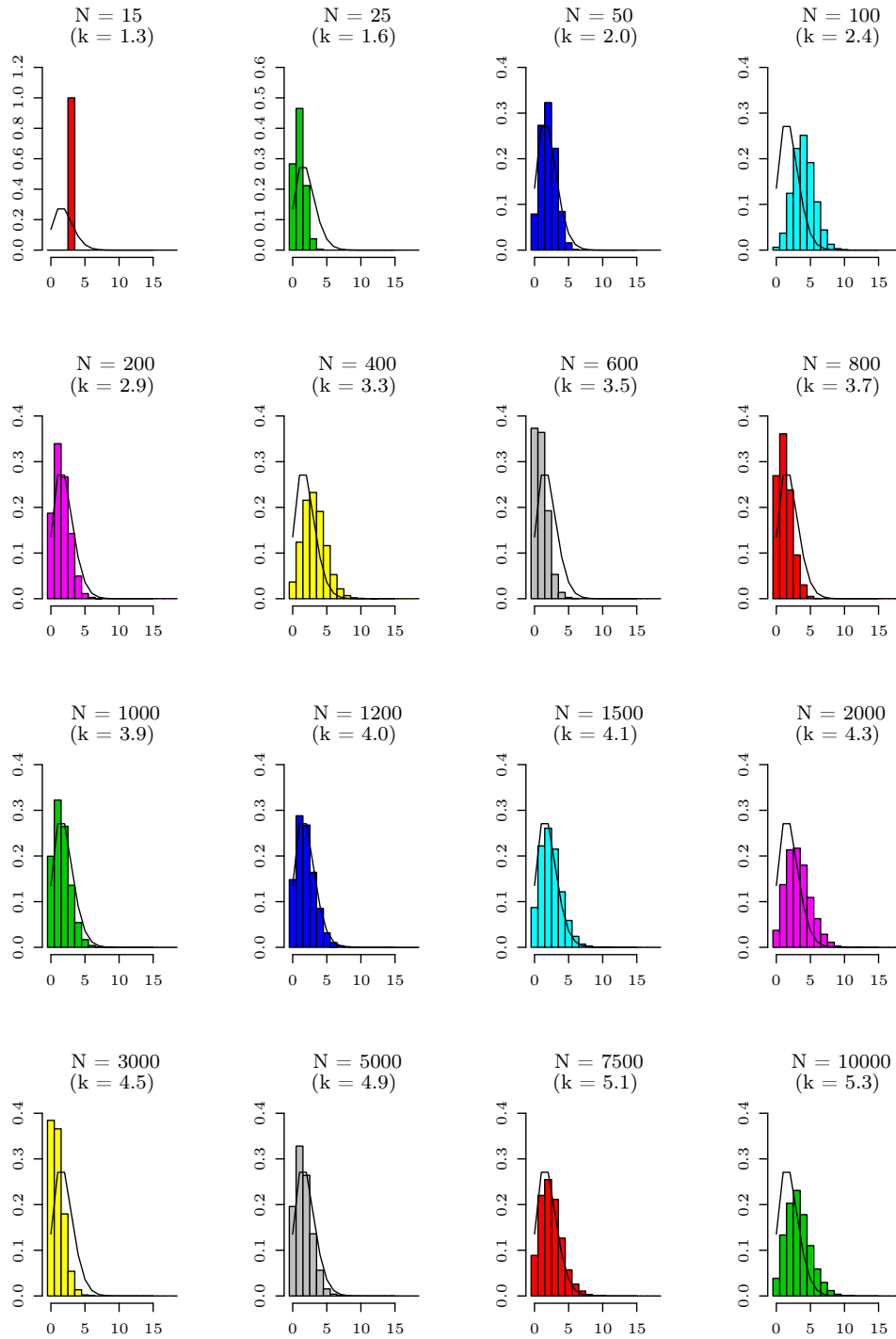
Obrázok 13: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(2)$ pre pevné $k = 2$ a pre rastúce hodnoty N .



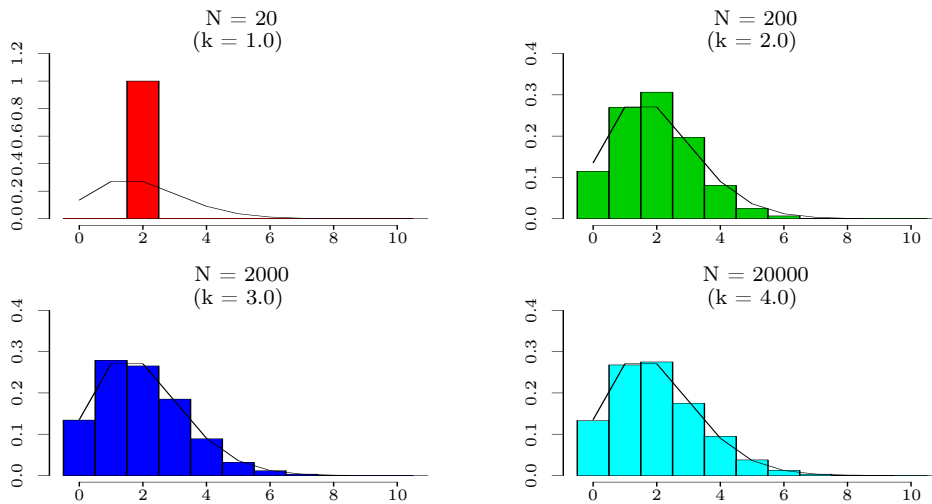
Obrázok 14: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(2)$ pre pevné $p = \frac{1}{10}$ a pre rastúce hodnoty N .



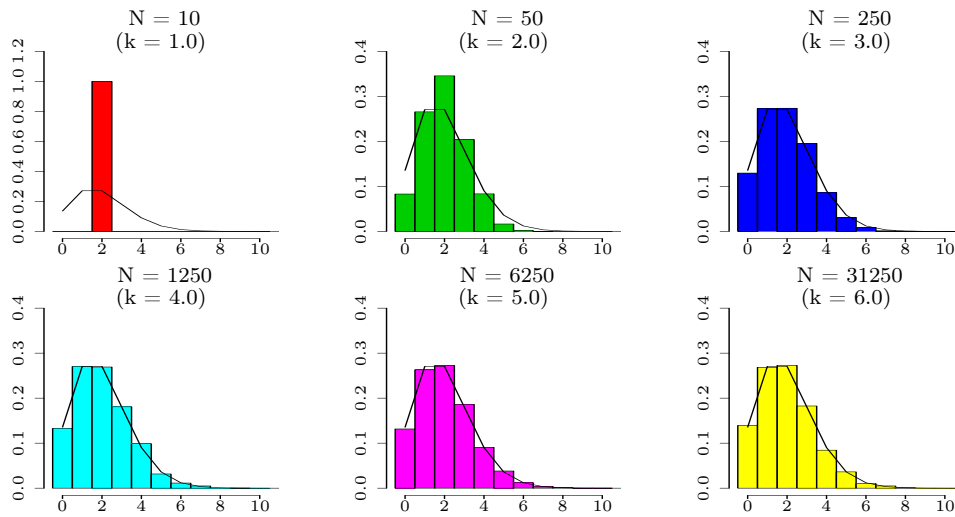
Obrázok 15: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(2)$ pre pevné $k = 3$ a pre rastúce hodnoty N .



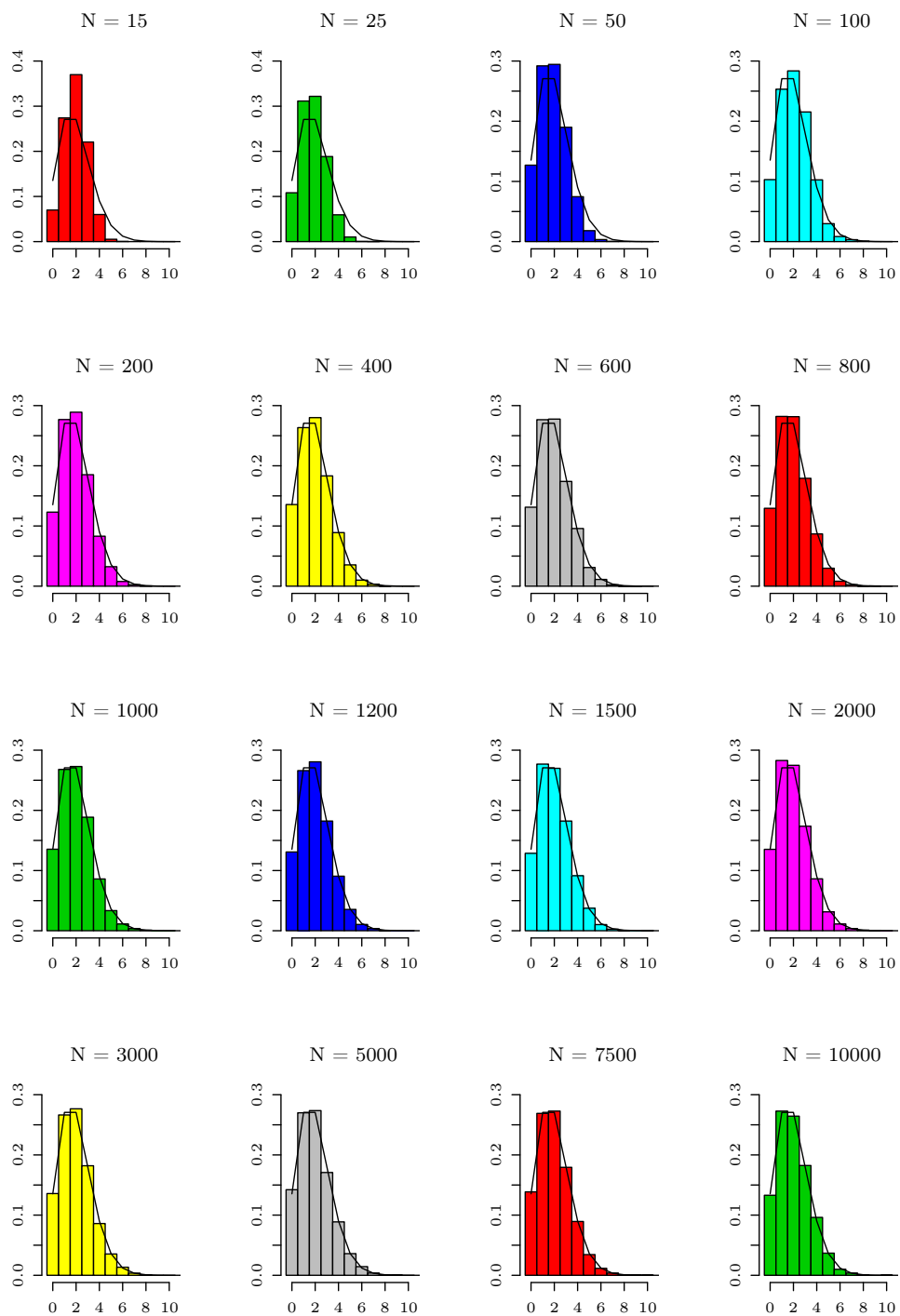
Obrázok 16: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(2)$ pre pevné $p = \frac{1}{5}$ a pre rastúce hodnoty N .



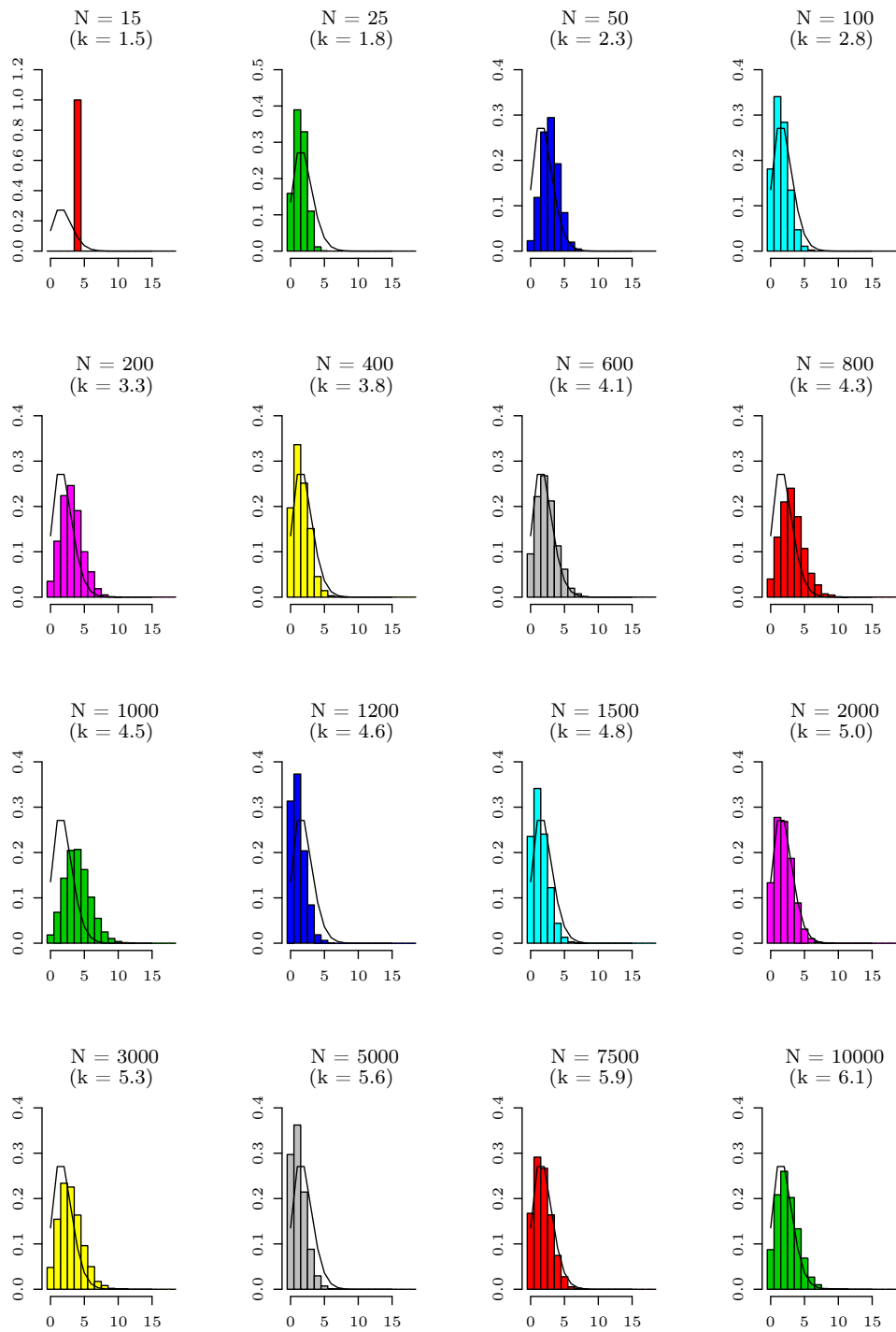
Obrázok 17: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(2)$ pre pevné $p = \frac{1}{10}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



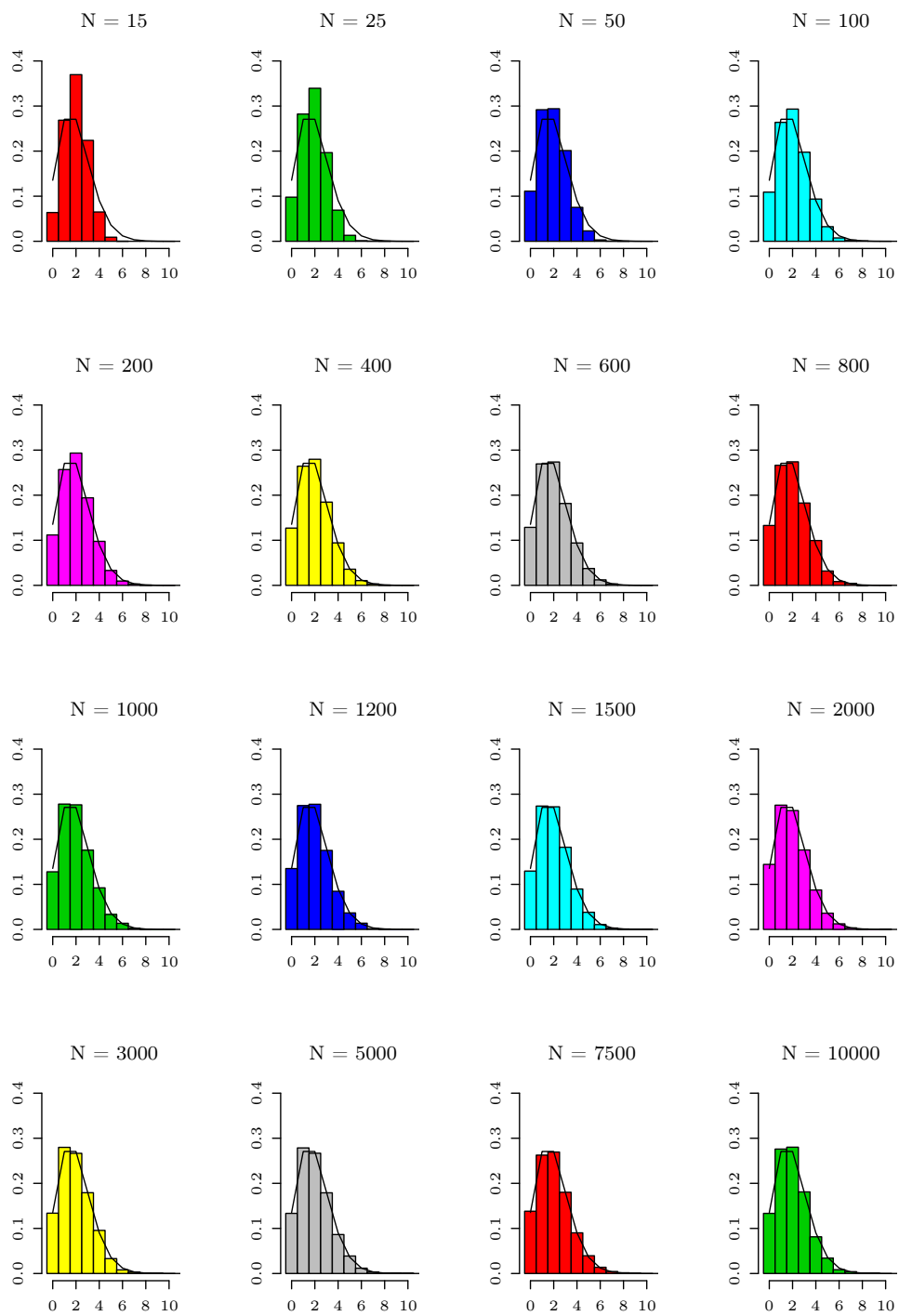
Obrázok 18: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(2)$ pre pevné $p = \frac{1}{5}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



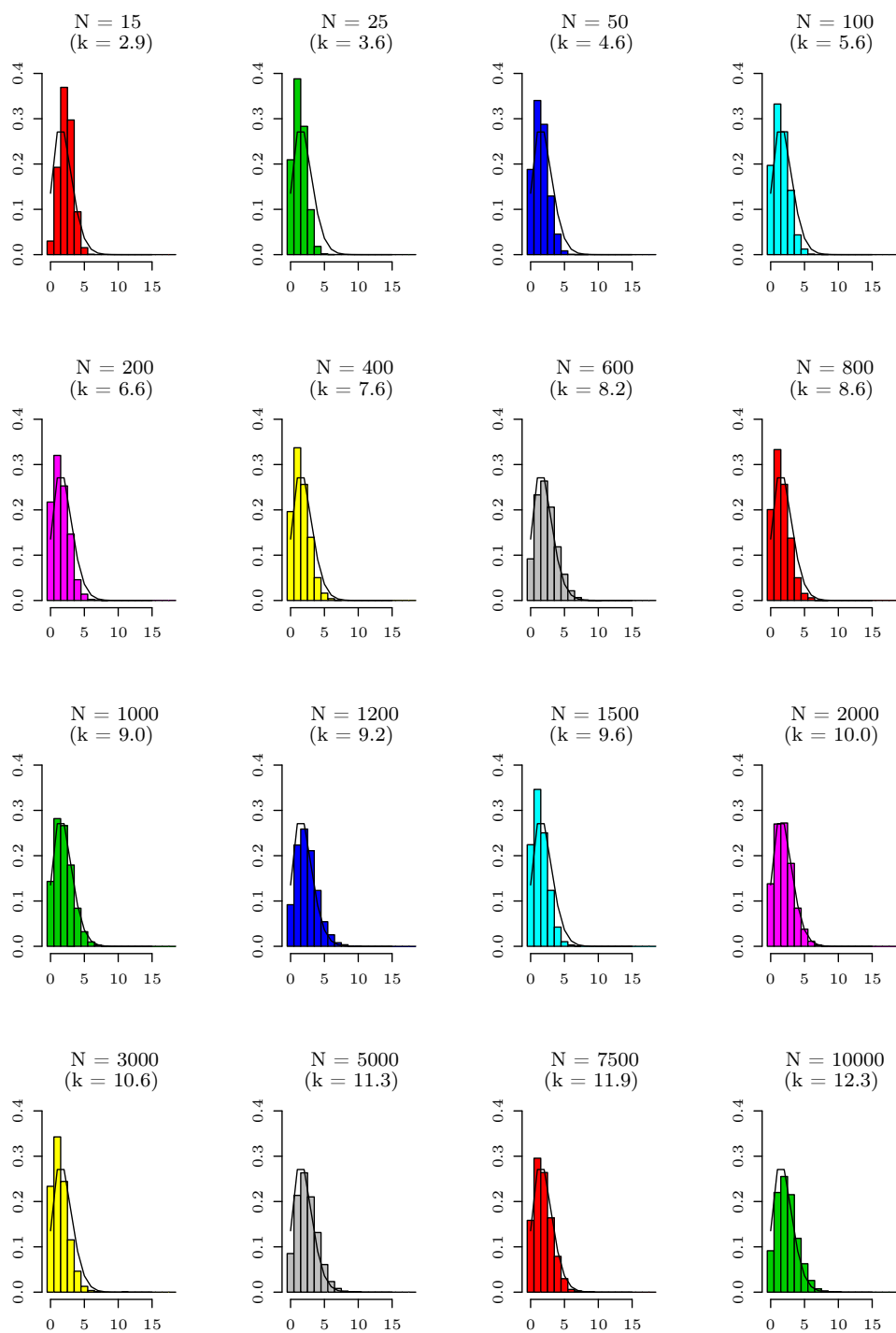
Obrázok 19: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(2)$ pre pevné $k = 4$ a pre rastúce hodnoty N .



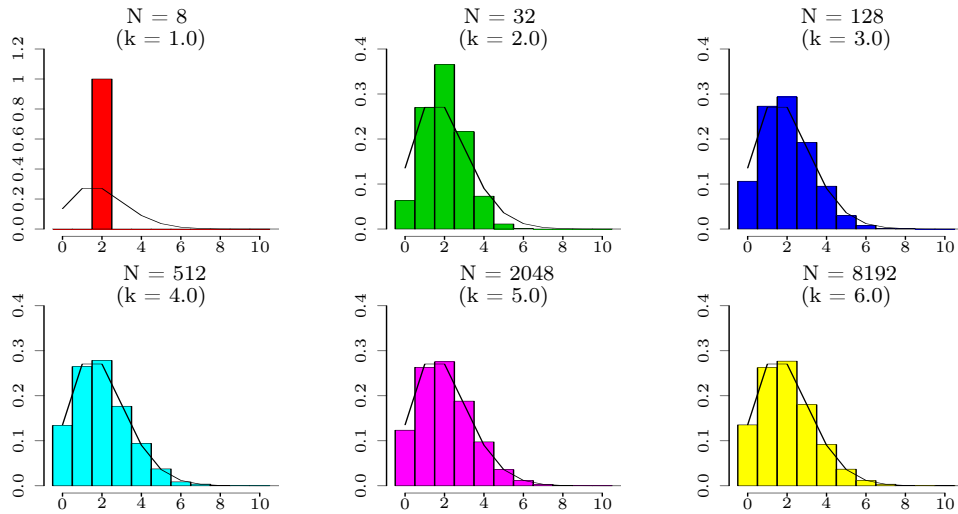
Obrázok 20: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(2)$ pre pevné $p = \frac{1}{4}$ a pre rastúce hodnoty N .



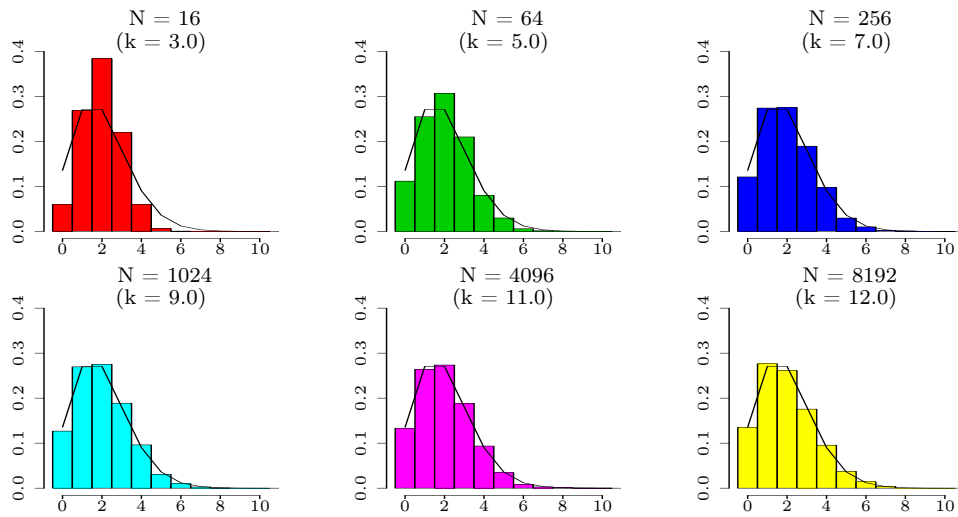
Obrázok 21: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(2)$ pre pevné $k = 5$ a pre rastúce hodnoty N .



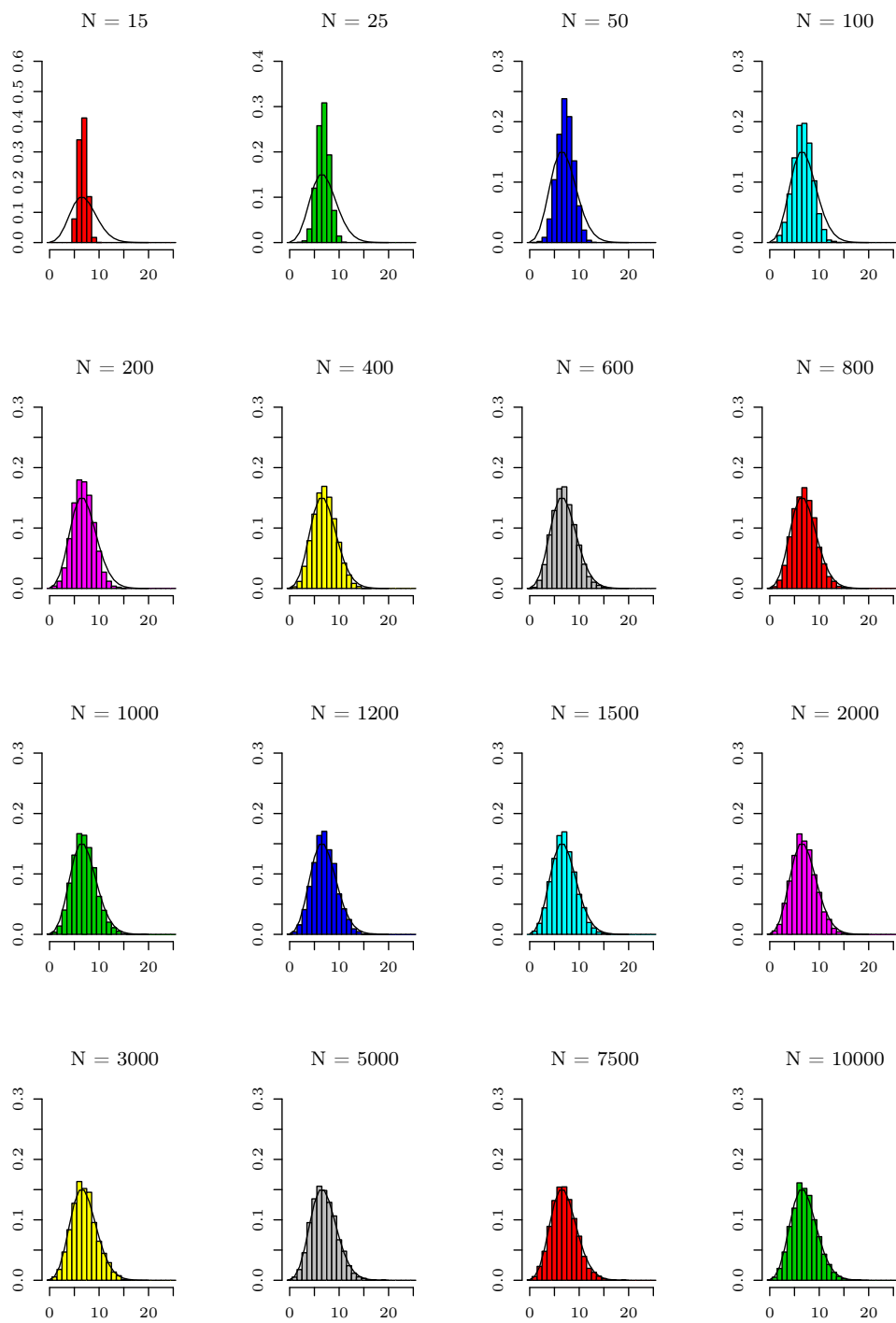
Obrázok 22: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(2)$ pre pevné $p = \frac{1}{2}$ a pre rastúce hodnoty N .



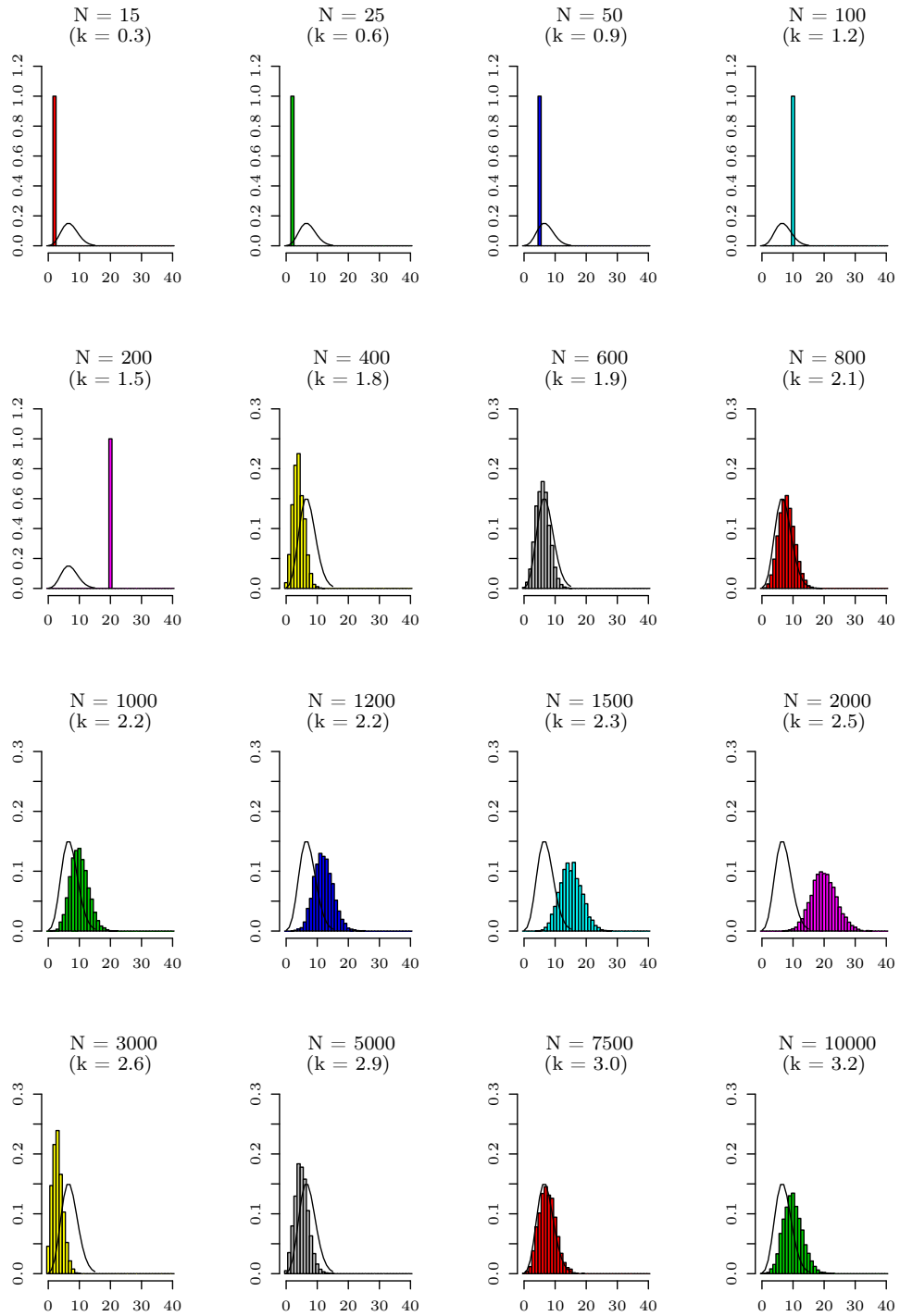
Obrázok 23: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(2)$ pre pevné $p = \frac{1}{4}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



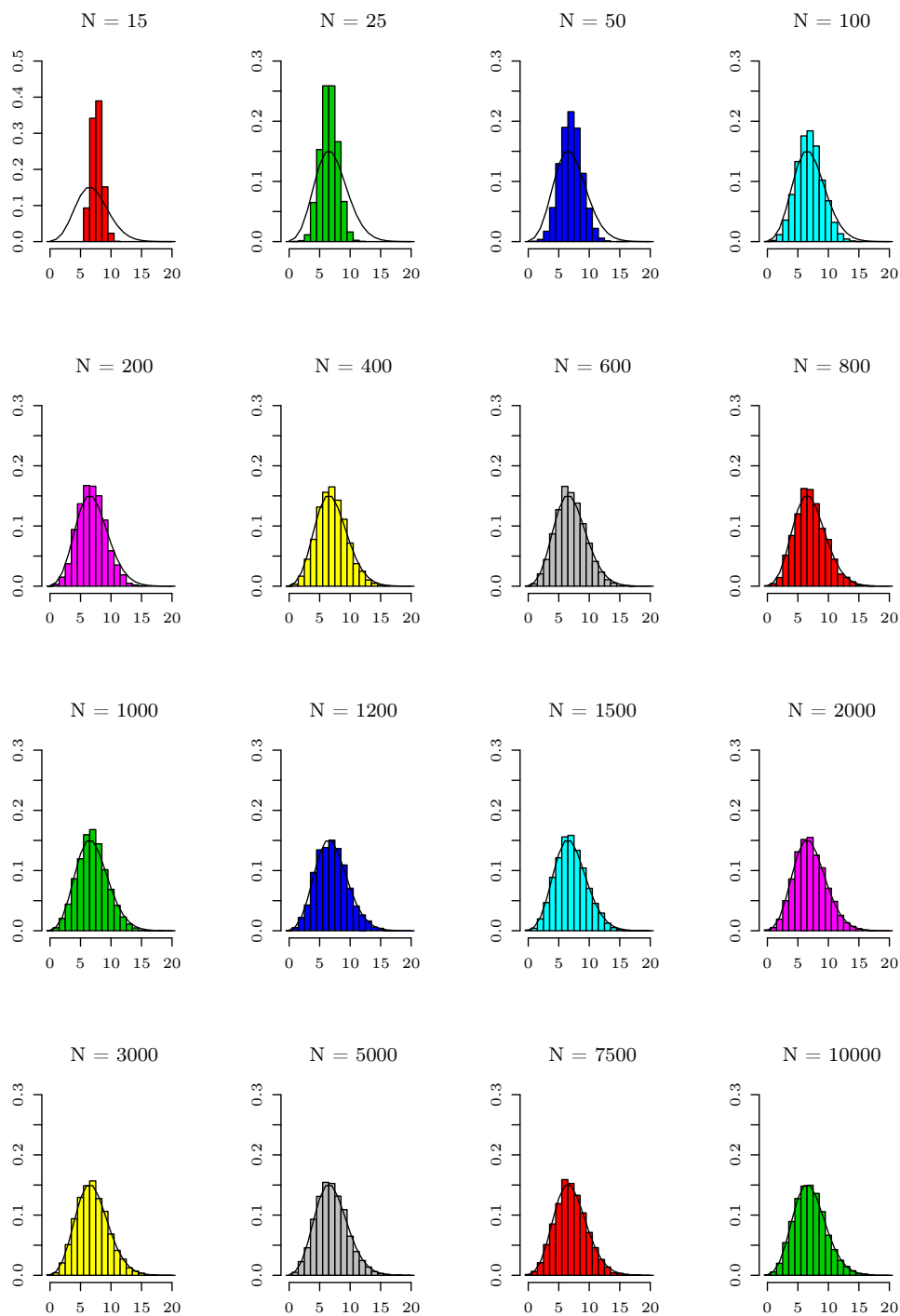
Obrázok 24: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(2)$ pre pevné $p = \frac{1}{2}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



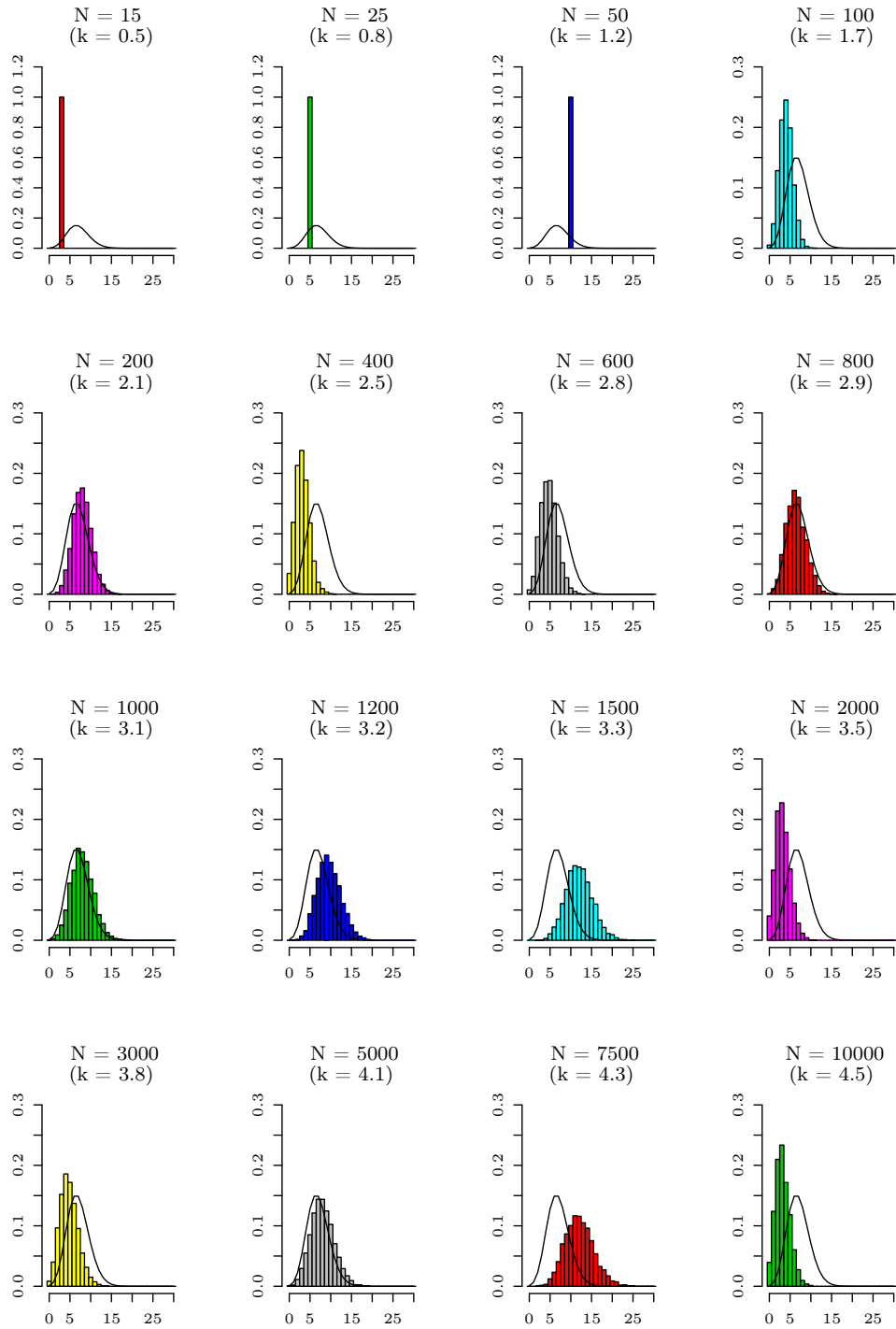
Obrázok 25: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(7)$ pre pevné $k = 2$ a pre rastúce hodnoty N .



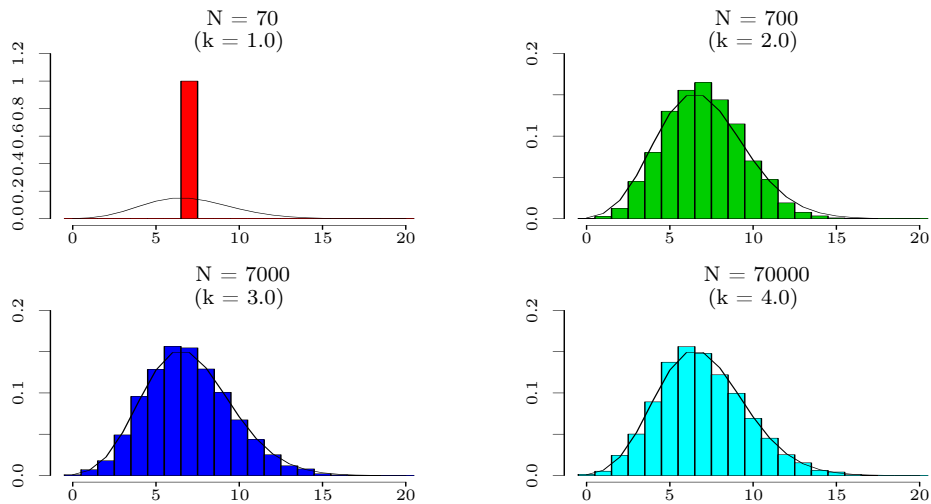
Obrázok 26: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(7)$ pre pevné $p = \frac{1}{10}$ a pre rastúce hodnoty N .



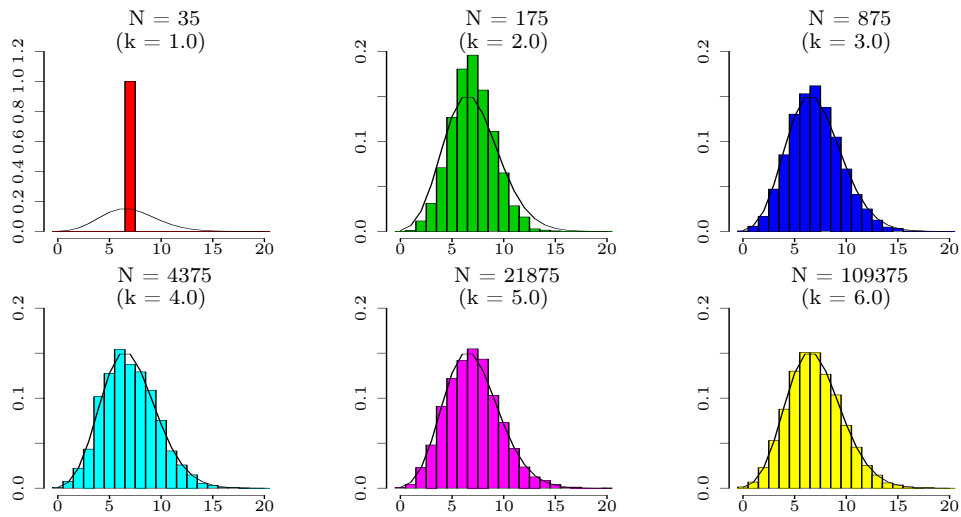
Obrázok 27: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(7)$ pre pevné $k = 3$ a pre rastúce hodnoty N .



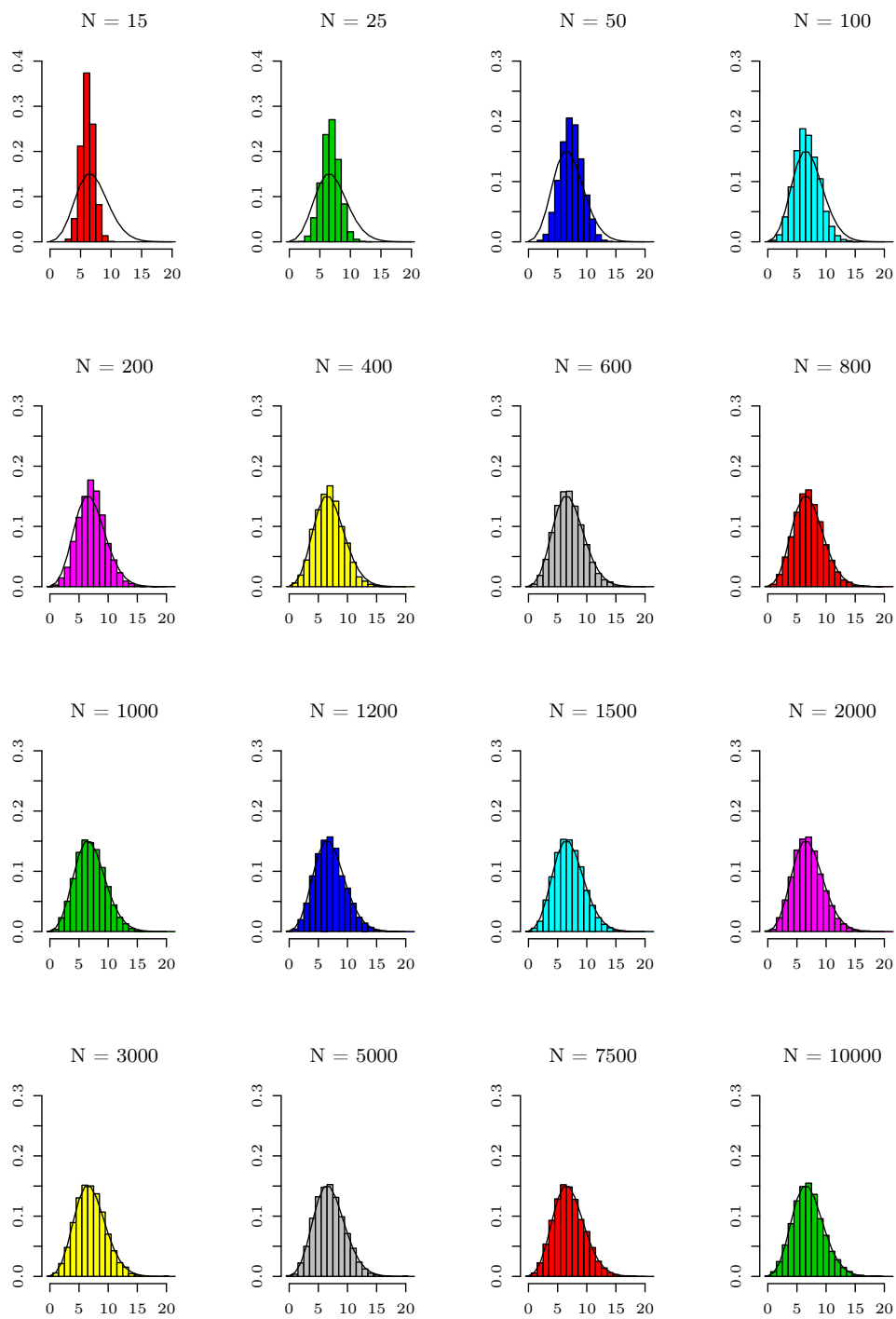
Obrázok 28: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(7)$ pre pevné $p = \frac{1}{5}$ a pre rastúce hodnoty N .



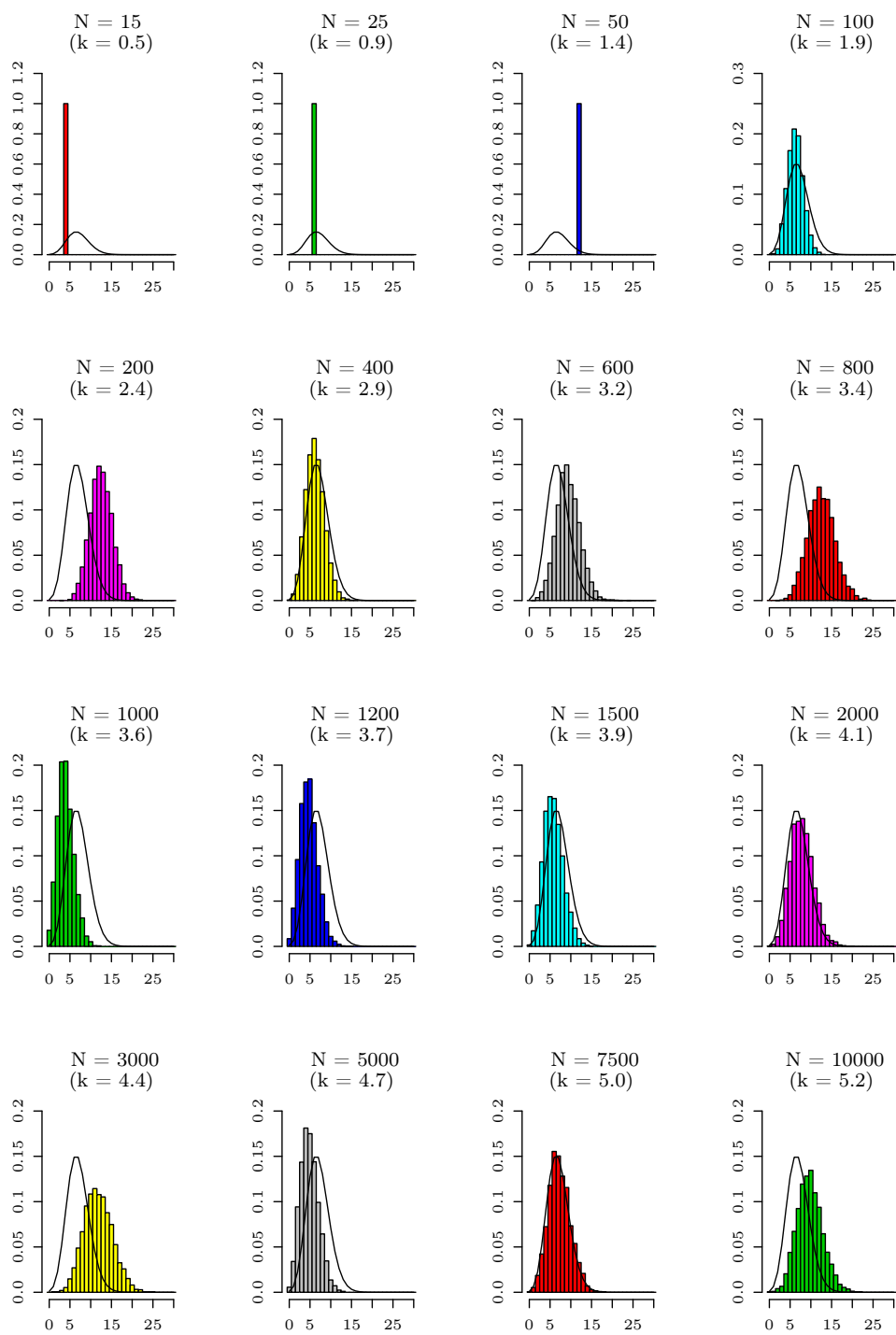
Obrázok 29: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(7)$ pre pevné $p = \frac{1}{10}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



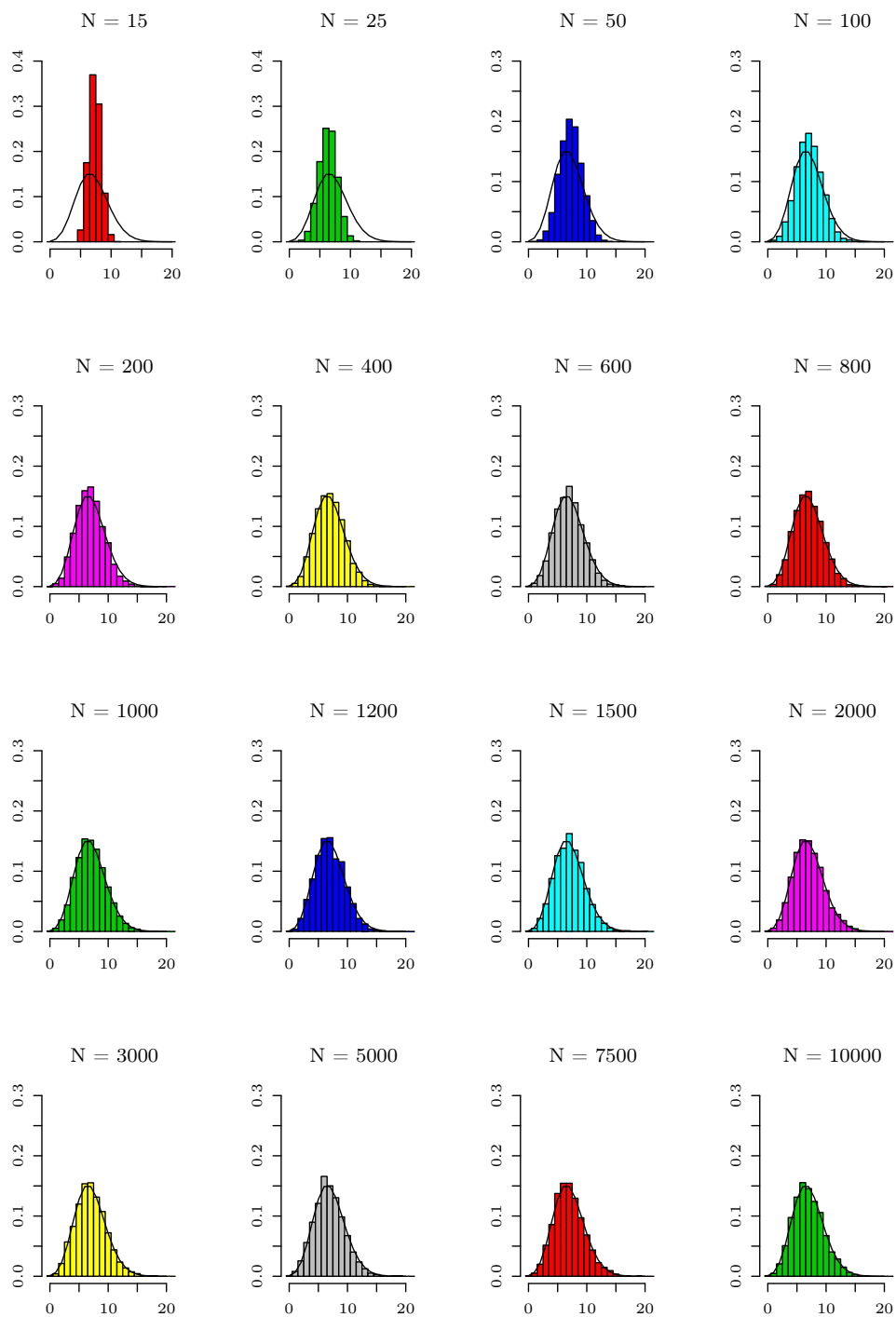
Obrázok 30: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(7)$ pre pevné $p = \frac{1}{5}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



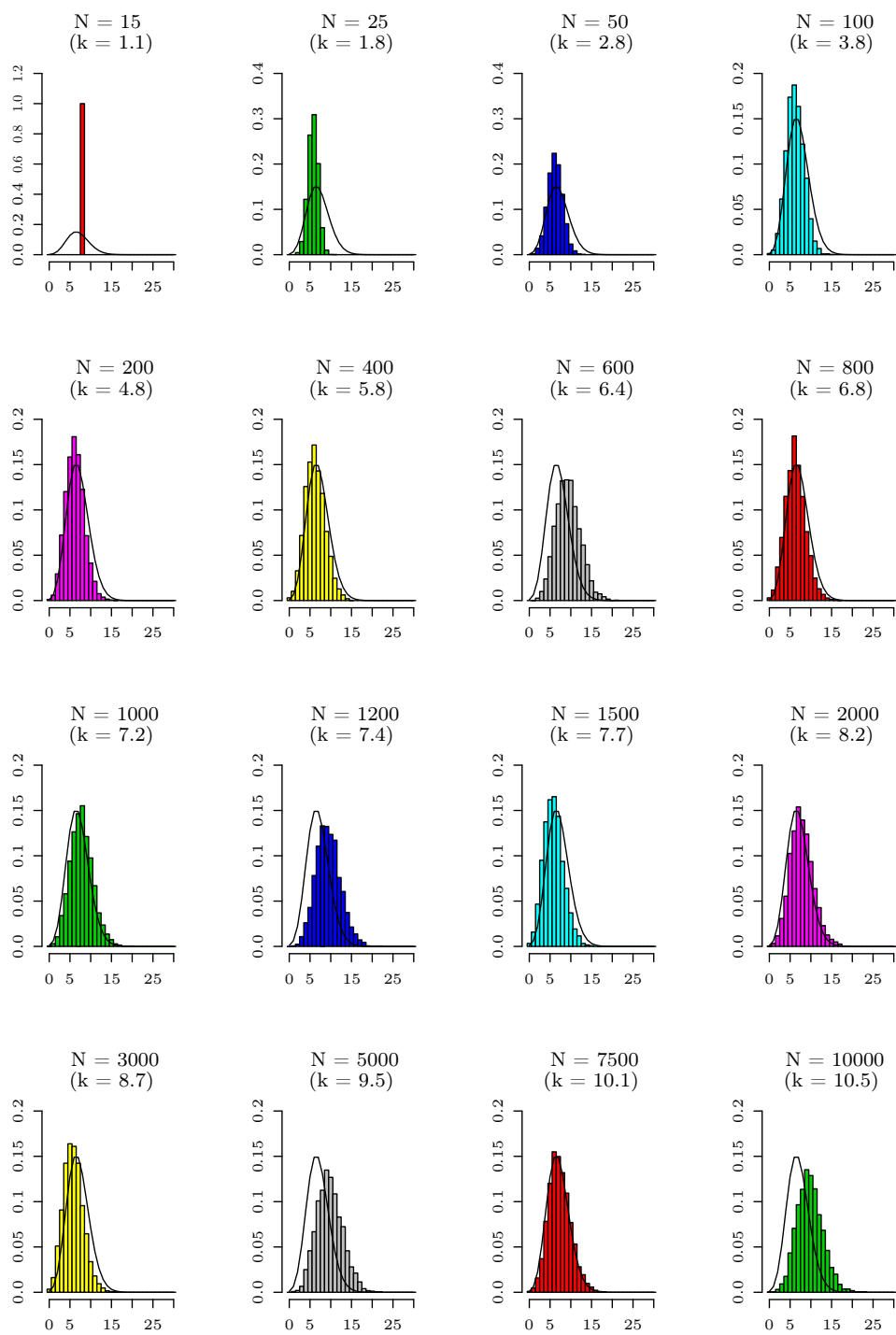
Obrázok 31: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(7)$ pre pevné $k = 4$ a pre rastúce hodnoty N .



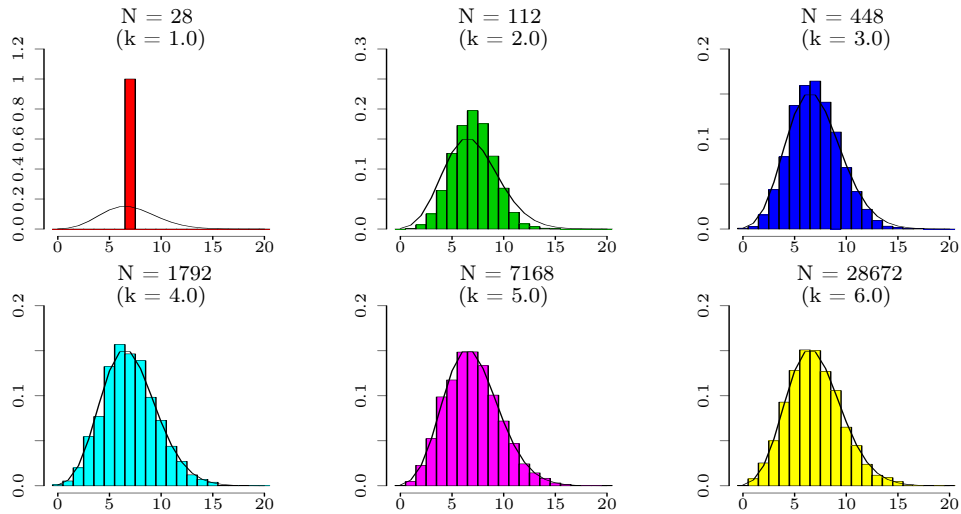
Obrázok 32: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(7)$ pre pevné $p = \frac{1}{4}$ a pre rastúce hodnoty N .



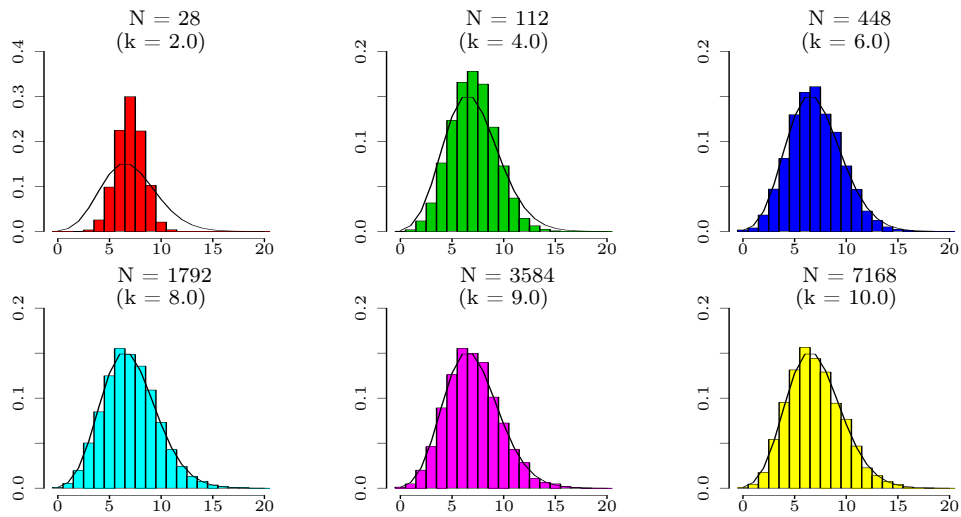
Obrázok 33: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(7)$ pre pevné $k = 5$ a pre rastúce hodnoty N .



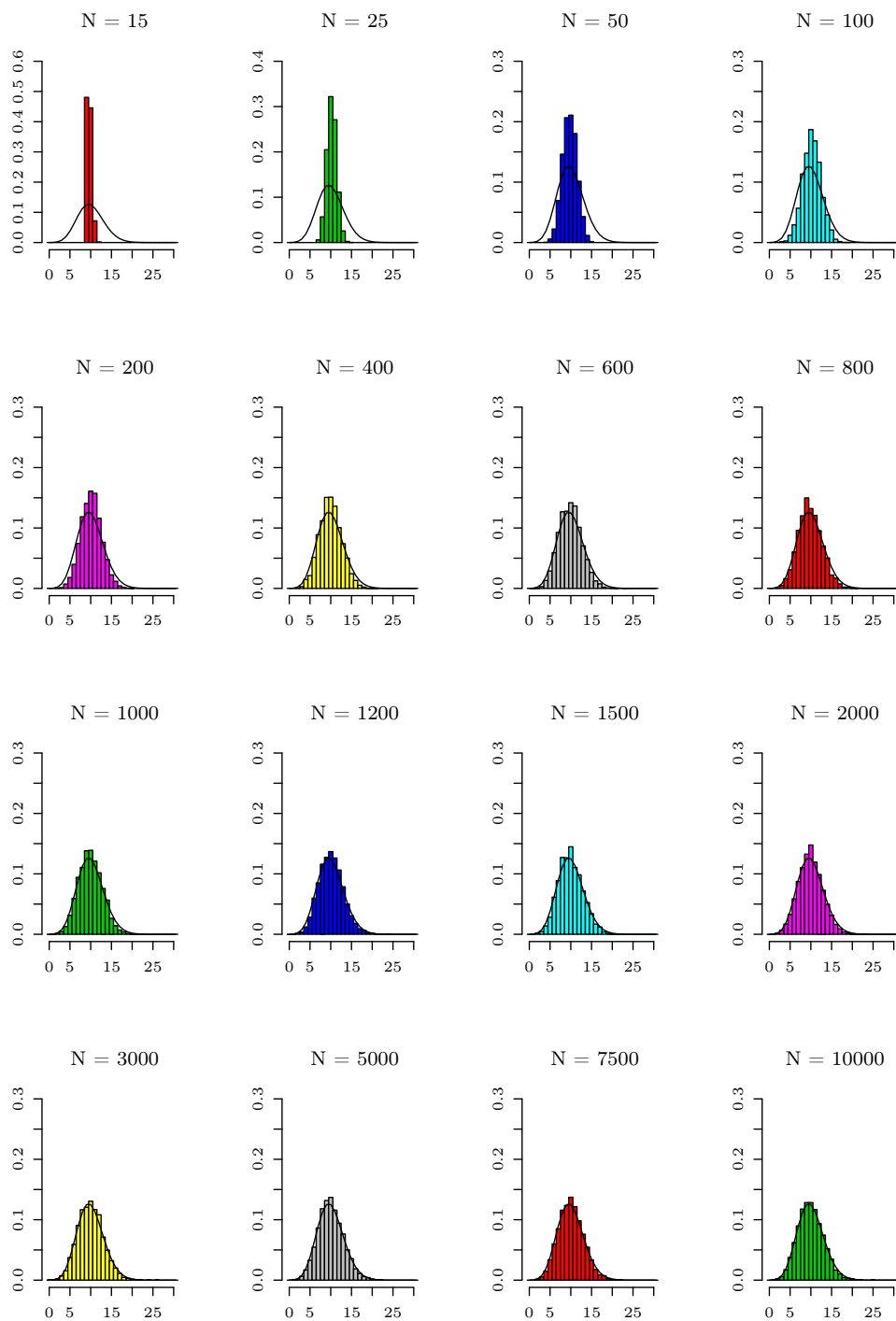
Obrázok 34: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(7)$ pre pevné $p = \frac{1}{2}$ a pre rastúce hodnoty N .



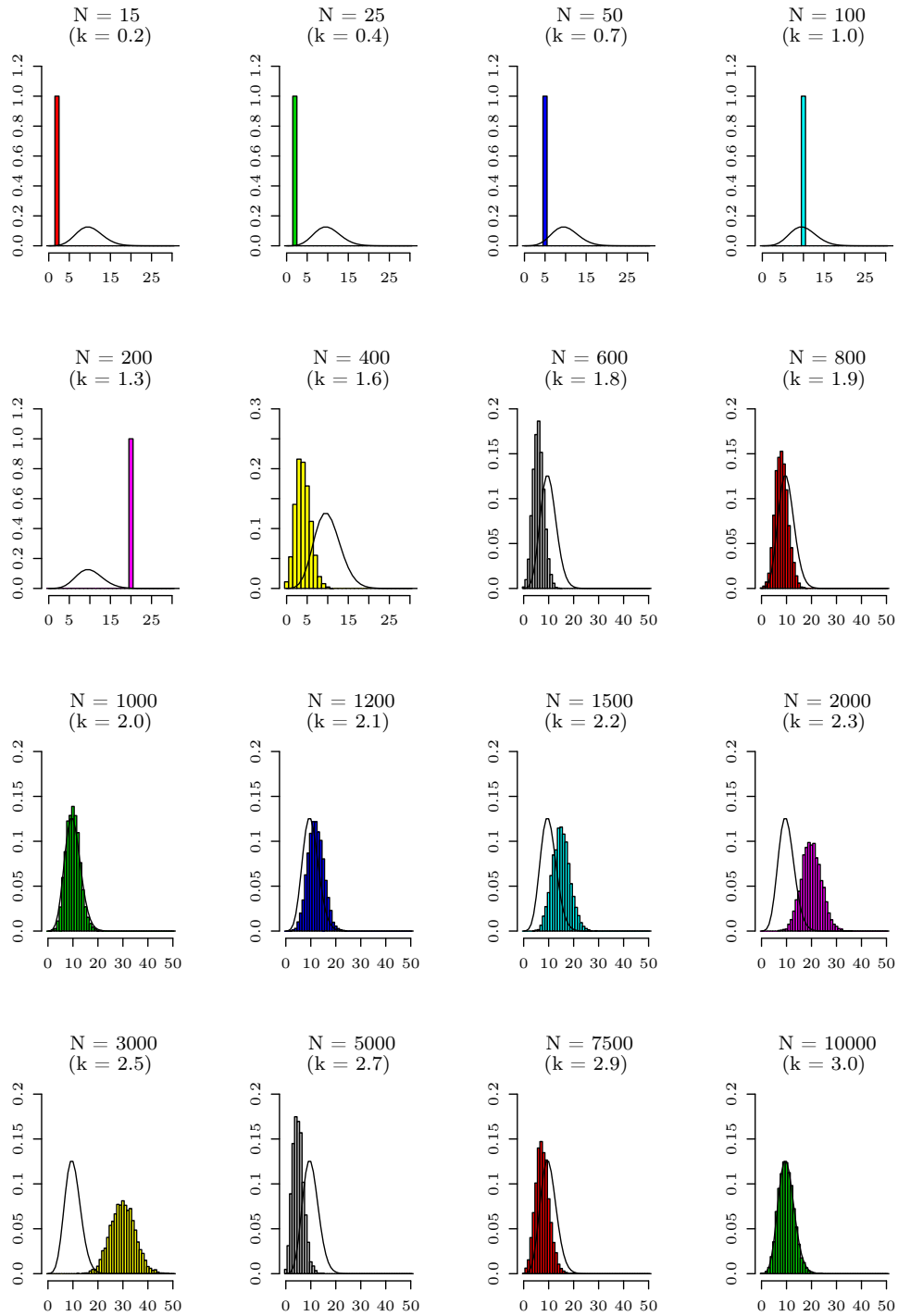
Obrázok 35: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(7)$ pre pevné $p = \frac{1}{4}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



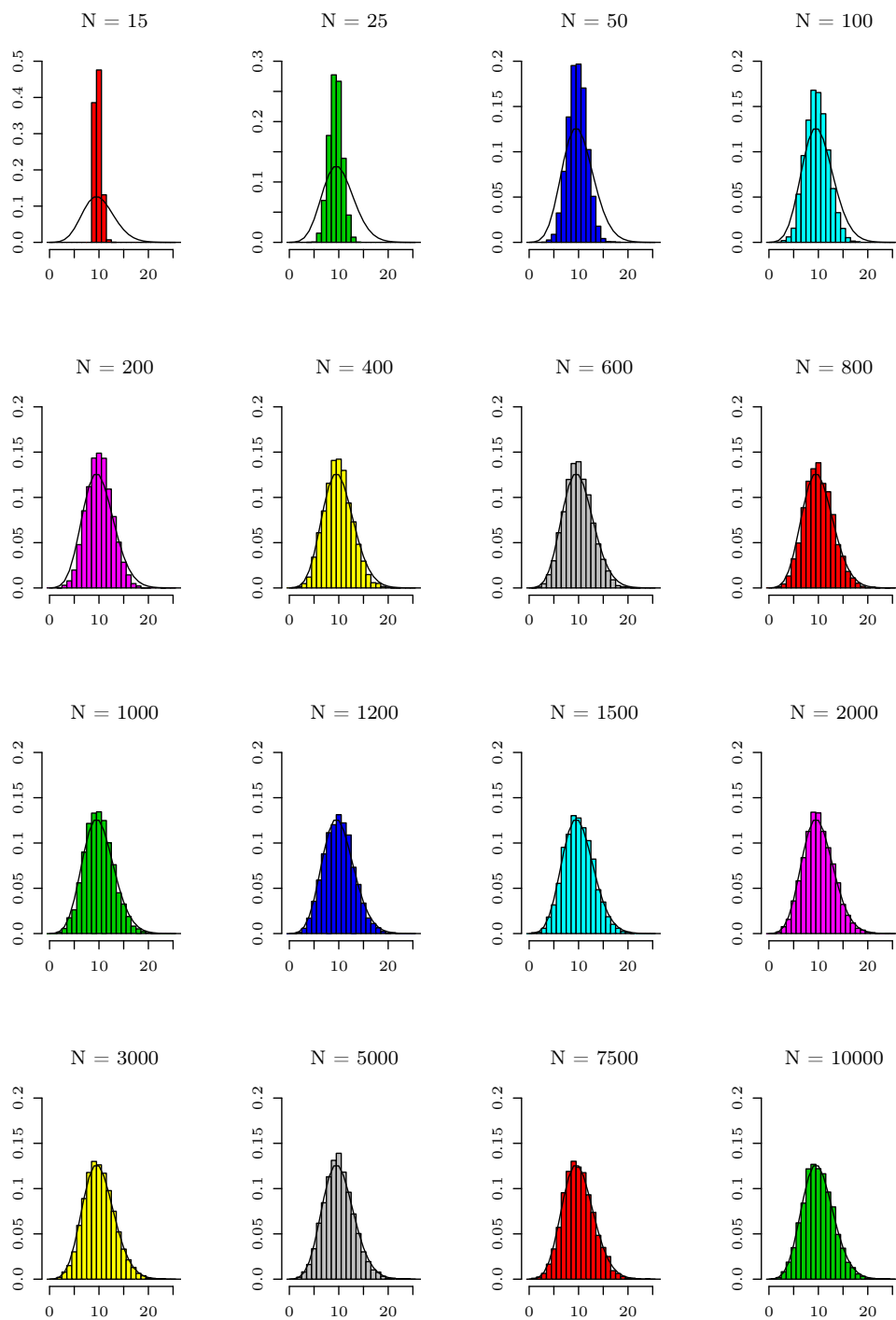
Obrázok 36: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(7)$ pre pevné $p = \frac{1}{2}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



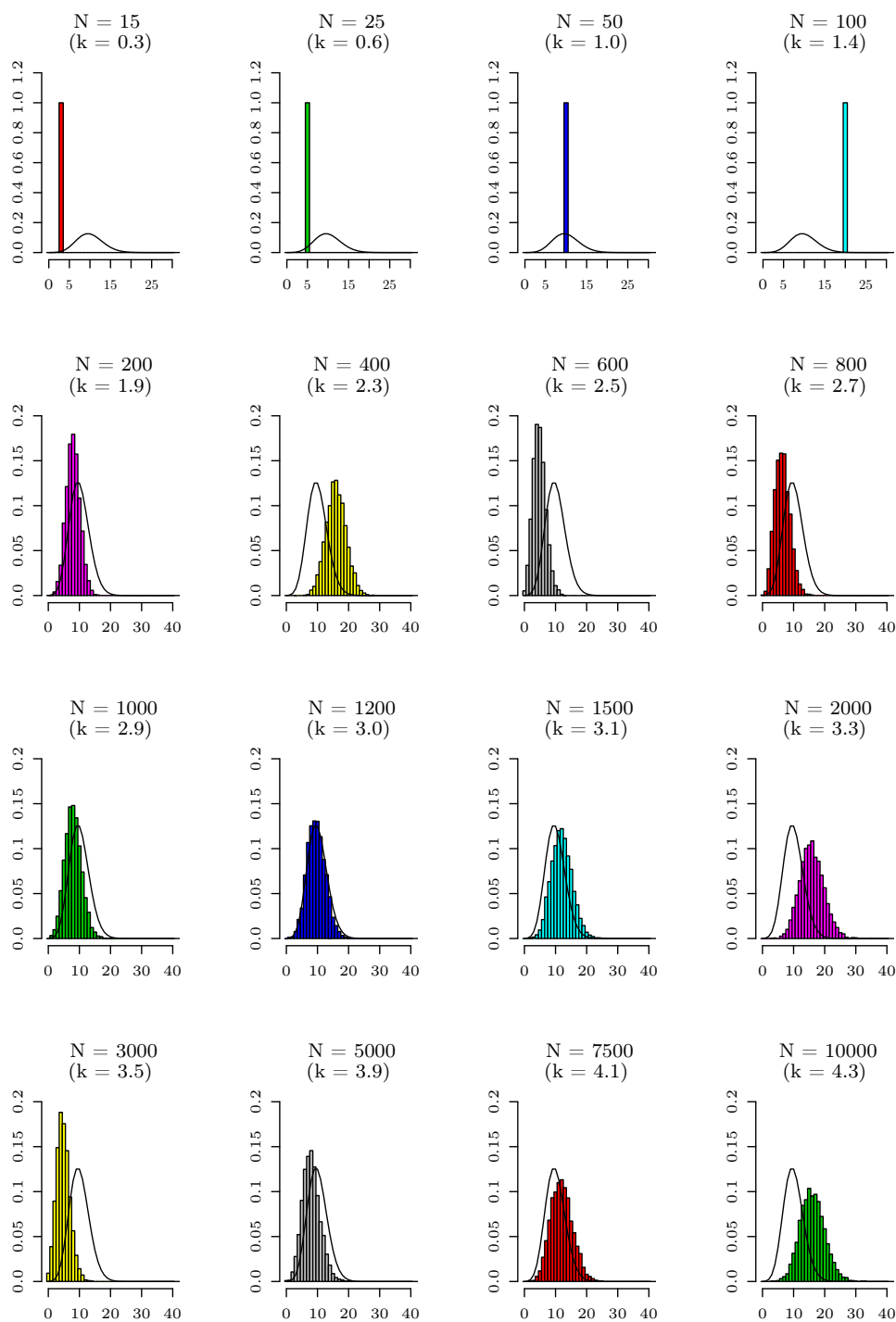
Obrázok 37: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(10)$ pre pevné $k = 2$ a pre rastúce hodnoty N .



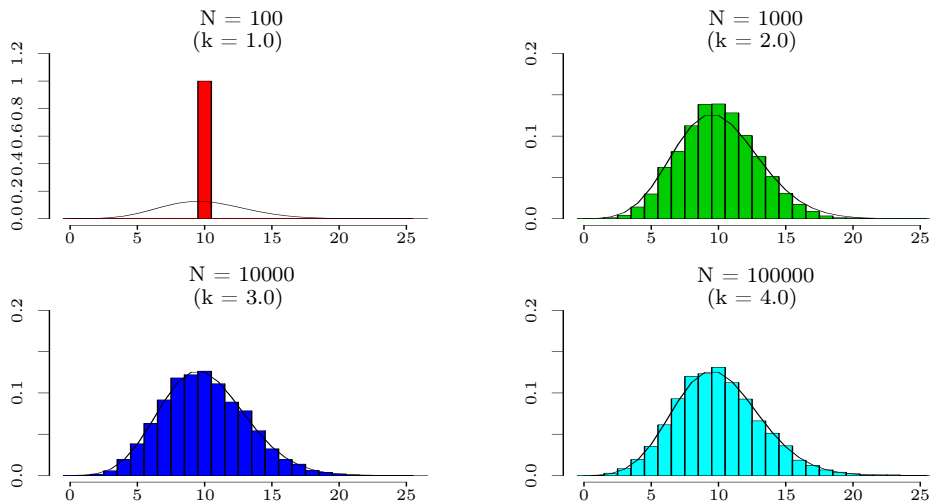
Obrázok 38: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(10)$ pre pevné $p = \frac{1}{10}$ a pre rastúce hodnoty N .



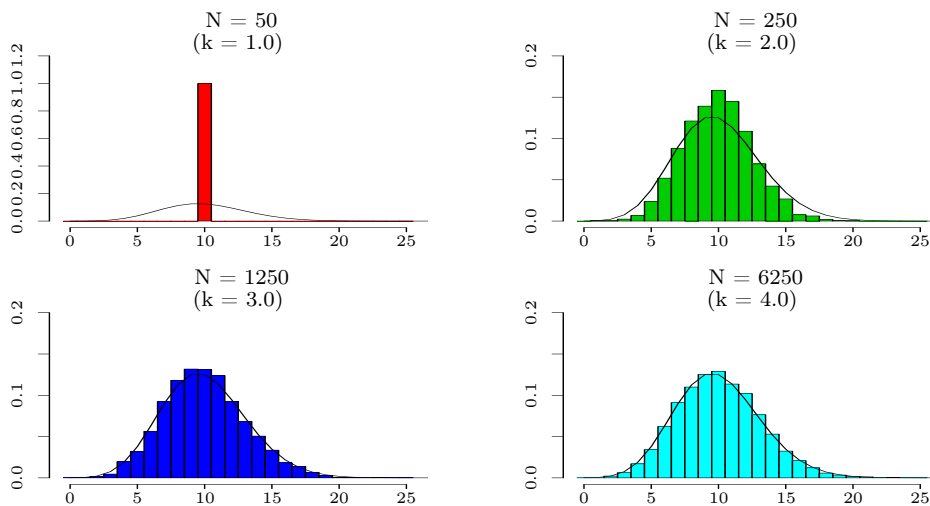
Obrázok 39: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(10)$ pre pevné $k = 3$ a pre rastúce hodnoty N .



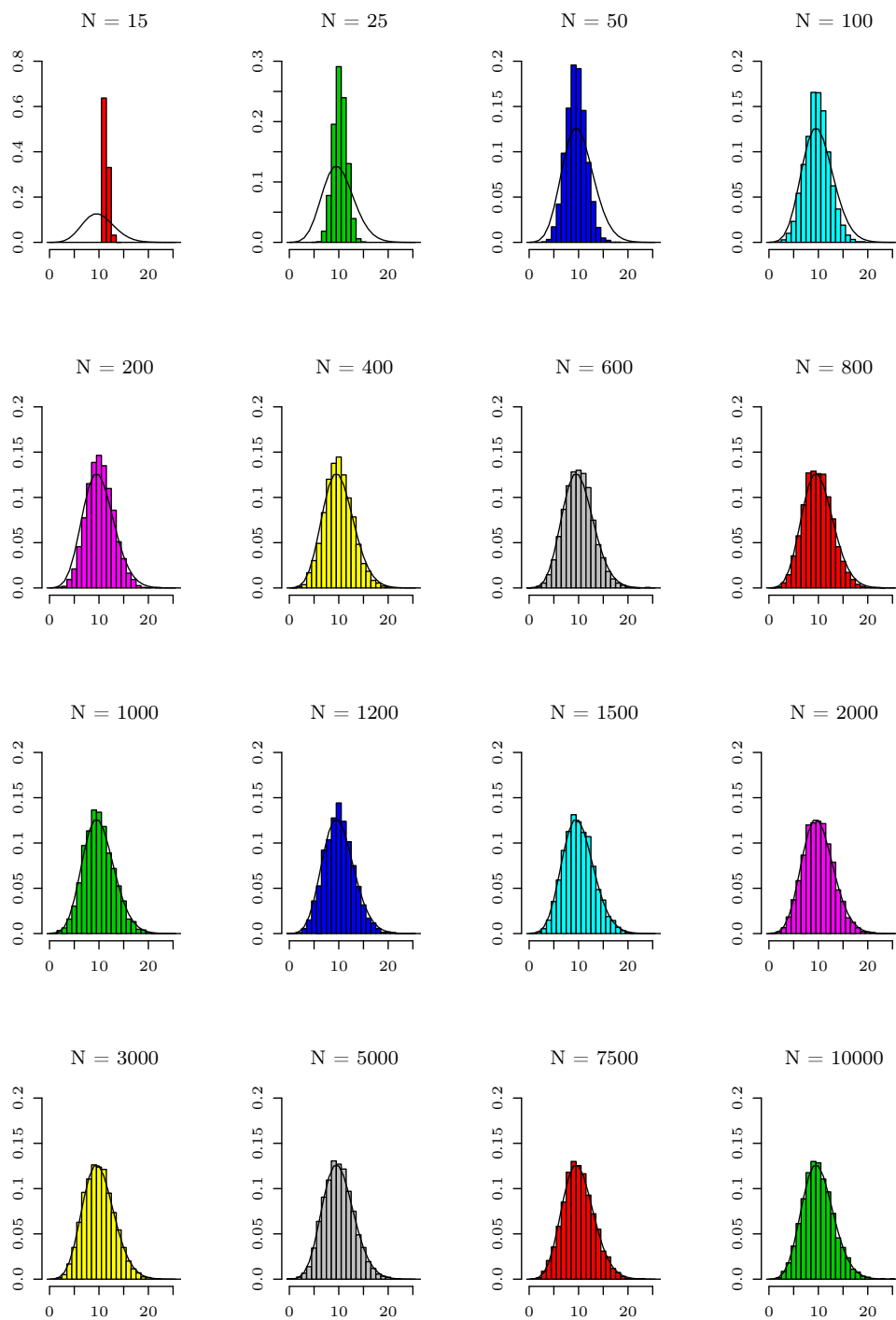
Obrázok 40: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(10)$ pre pevné $p = \frac{1}{5}$ a pre rastúce hodnoty N .



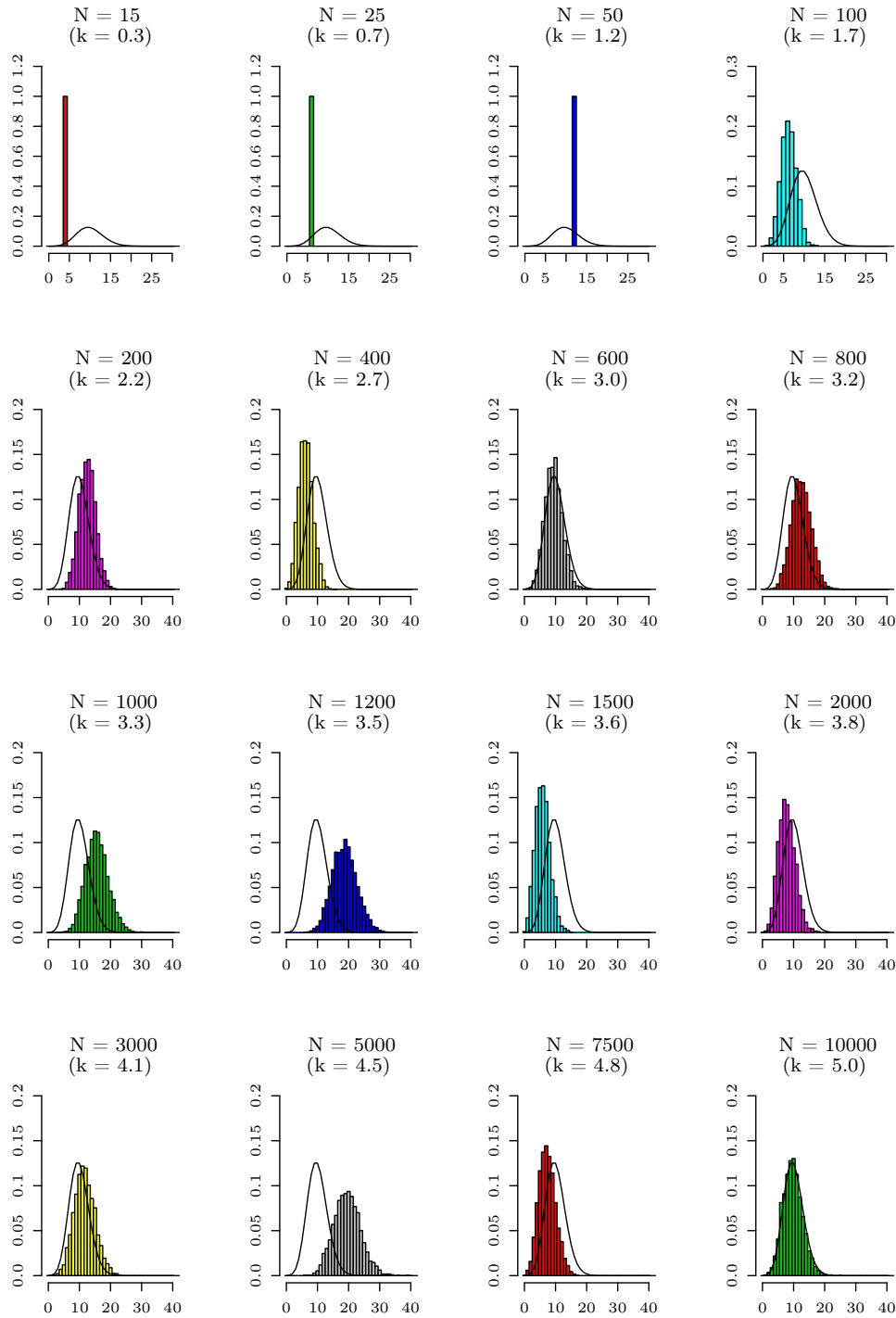
Obrázok 41: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(10)$ pre pevné $p = \frac{1}{10}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



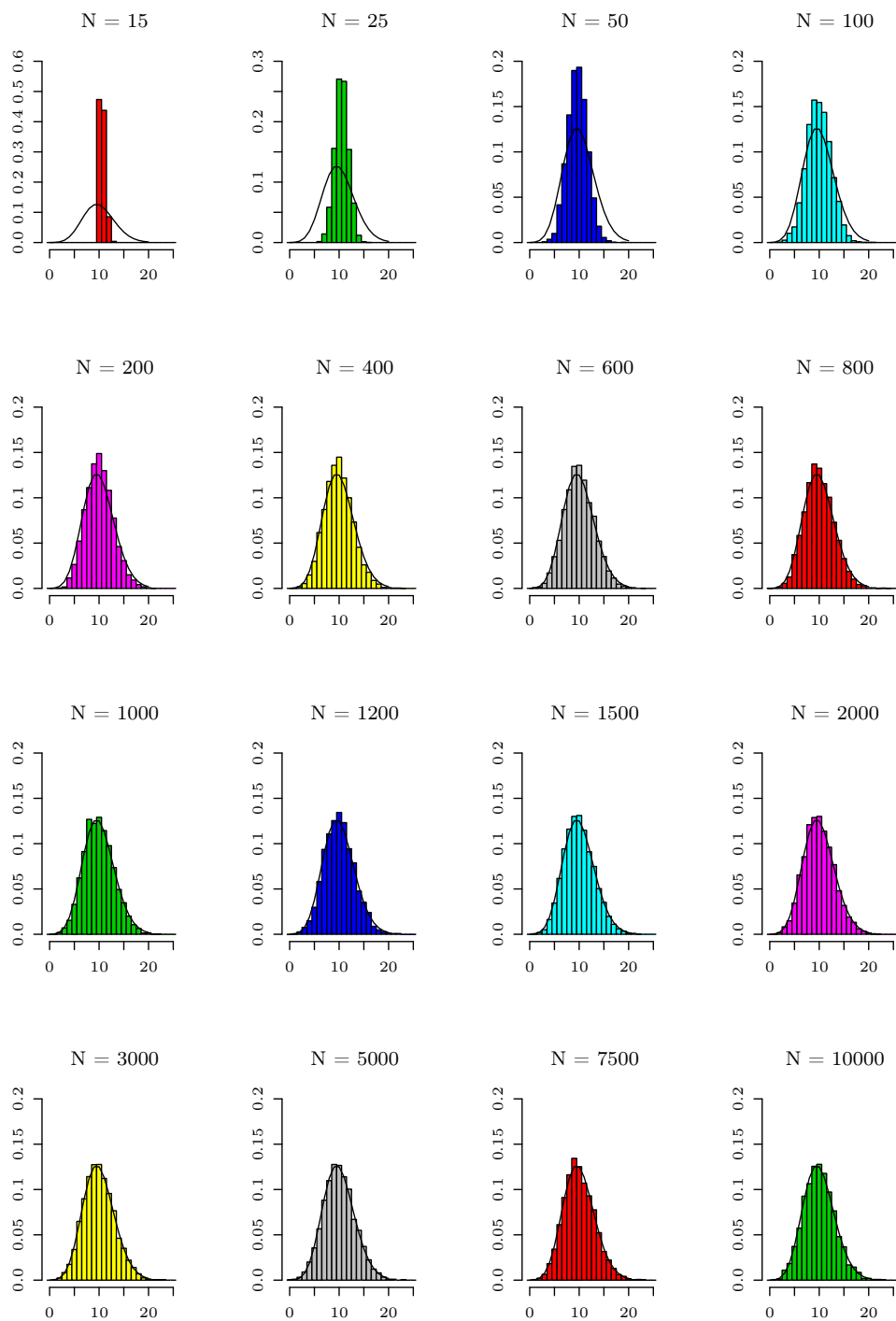
Obrázok 42: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(10)$ pre pevné $p = \frac{1}{5}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



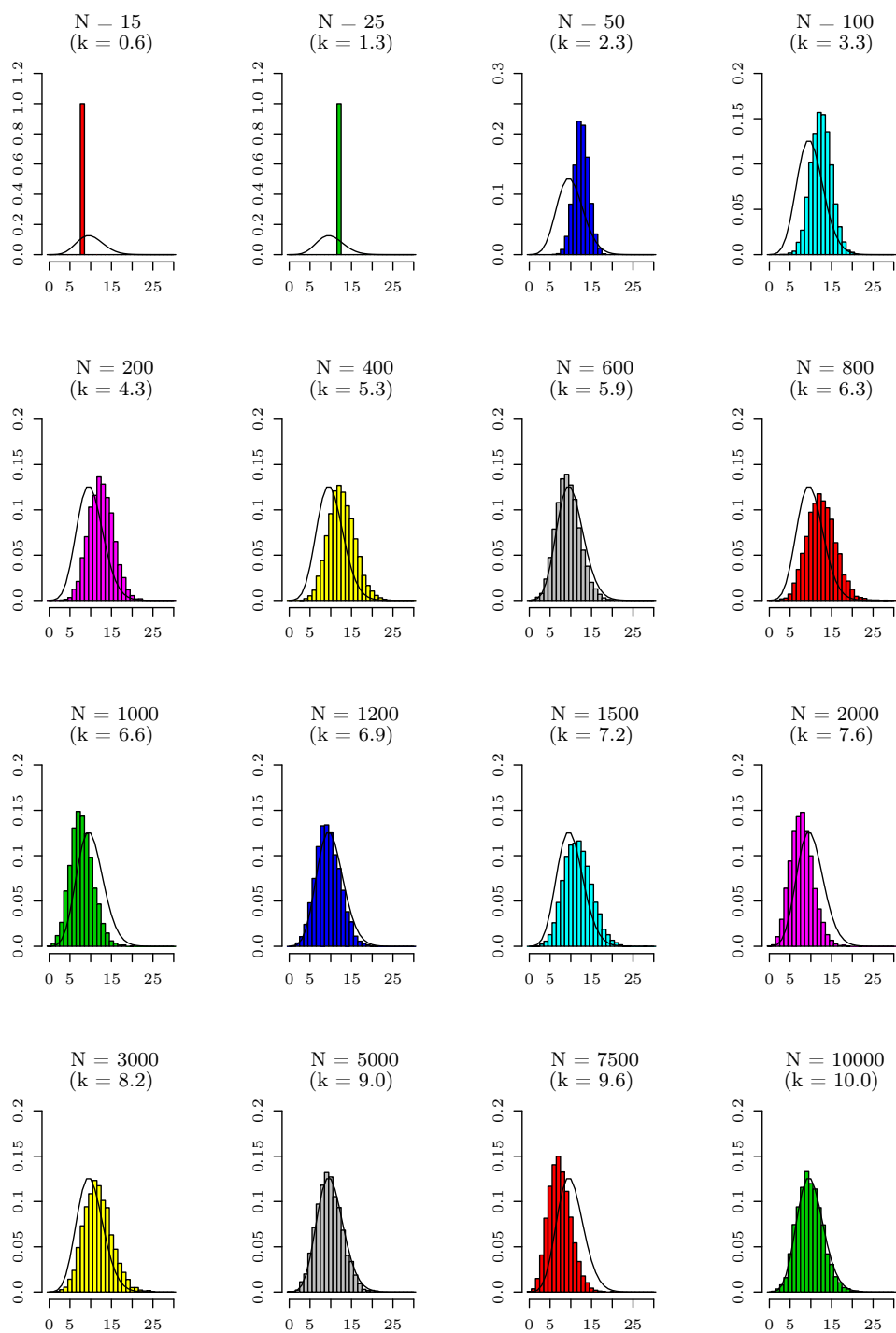
Obrázok 43: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(10)$ pre pevné $k = 4$ a pre rastúce hodnoty N .



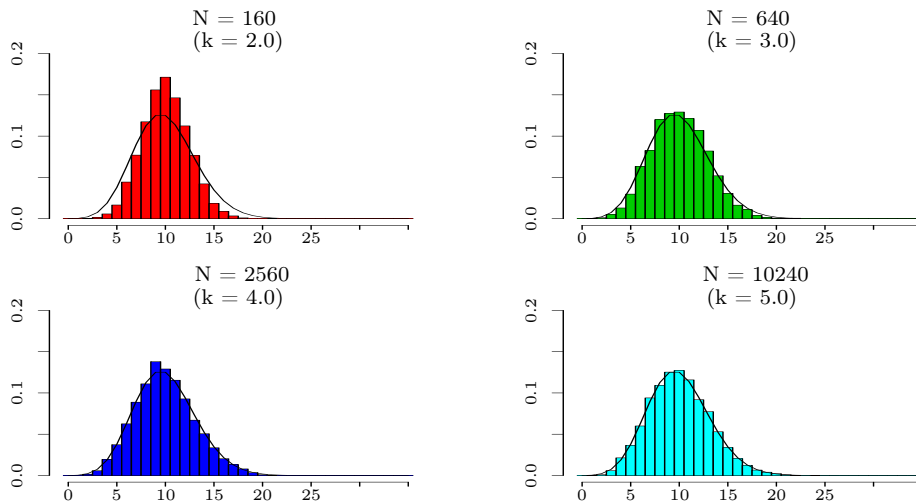
Obrázok 44: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(10)$ pre pevné $p = \frac{1}{4}$ a pre rastúce hodnoty N .



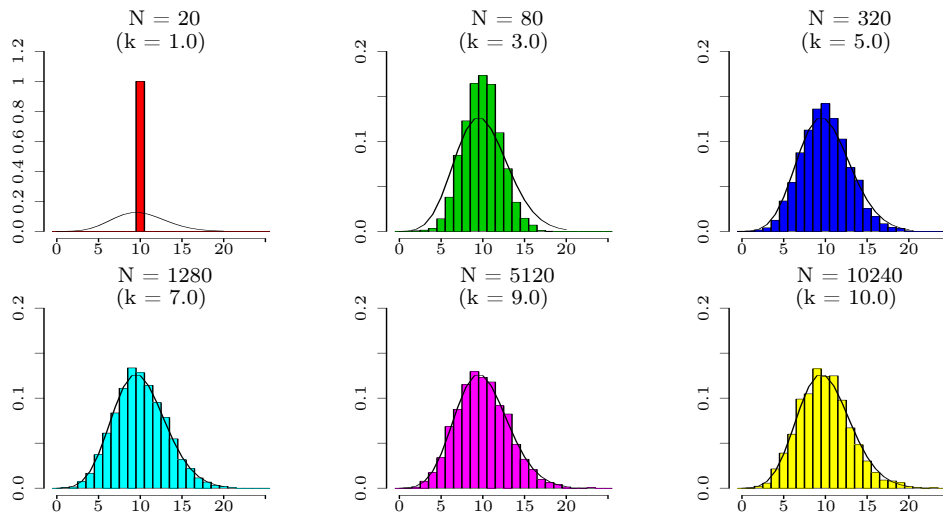
Obrázok 45: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(10)$ pre pevné $k = 5$ a pre rastúce hodnoty N .



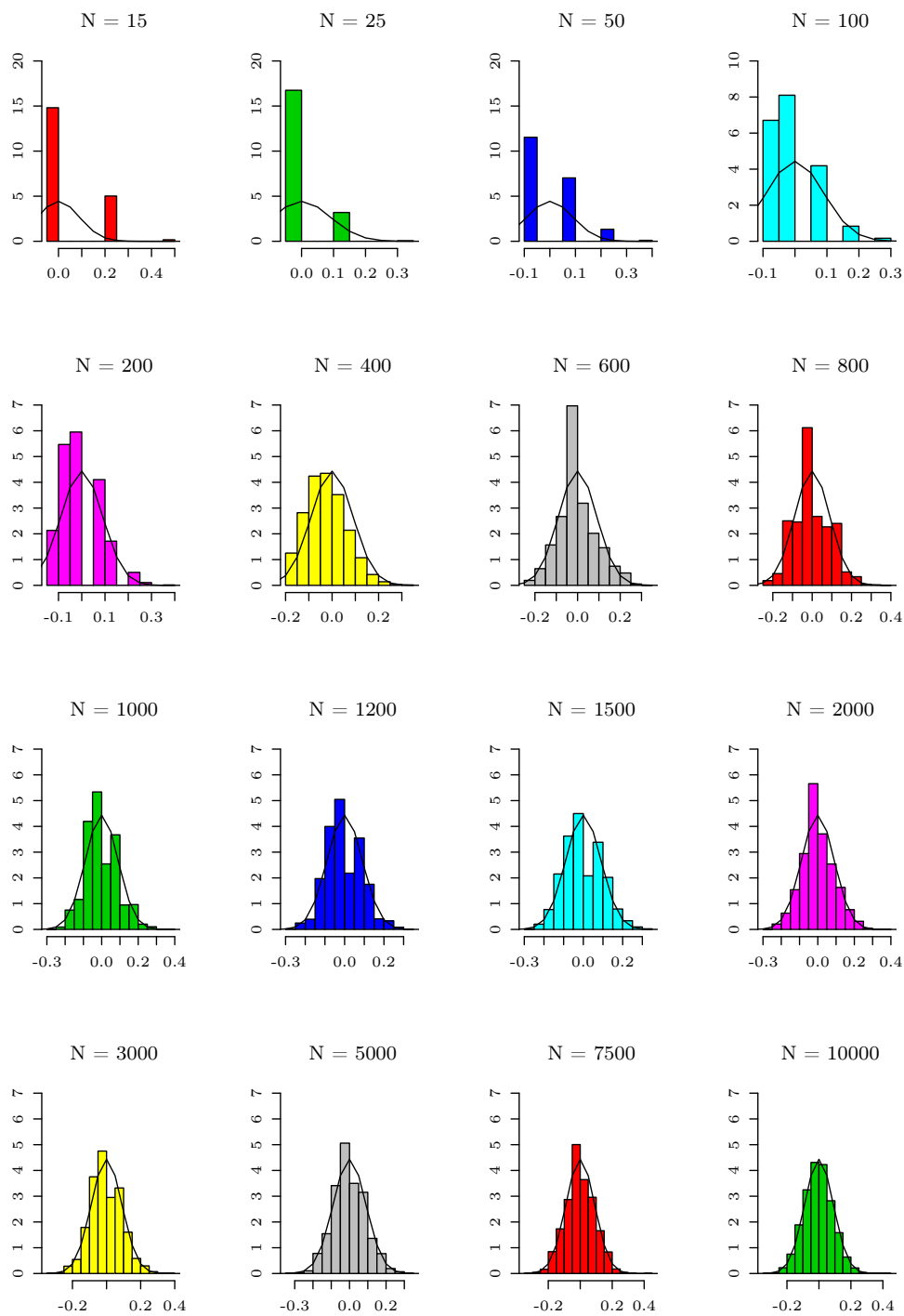
Obrázok 46: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(10)$ pre pevné $p = \frac{1}{2}$ a pre rastúce hodnoty N .



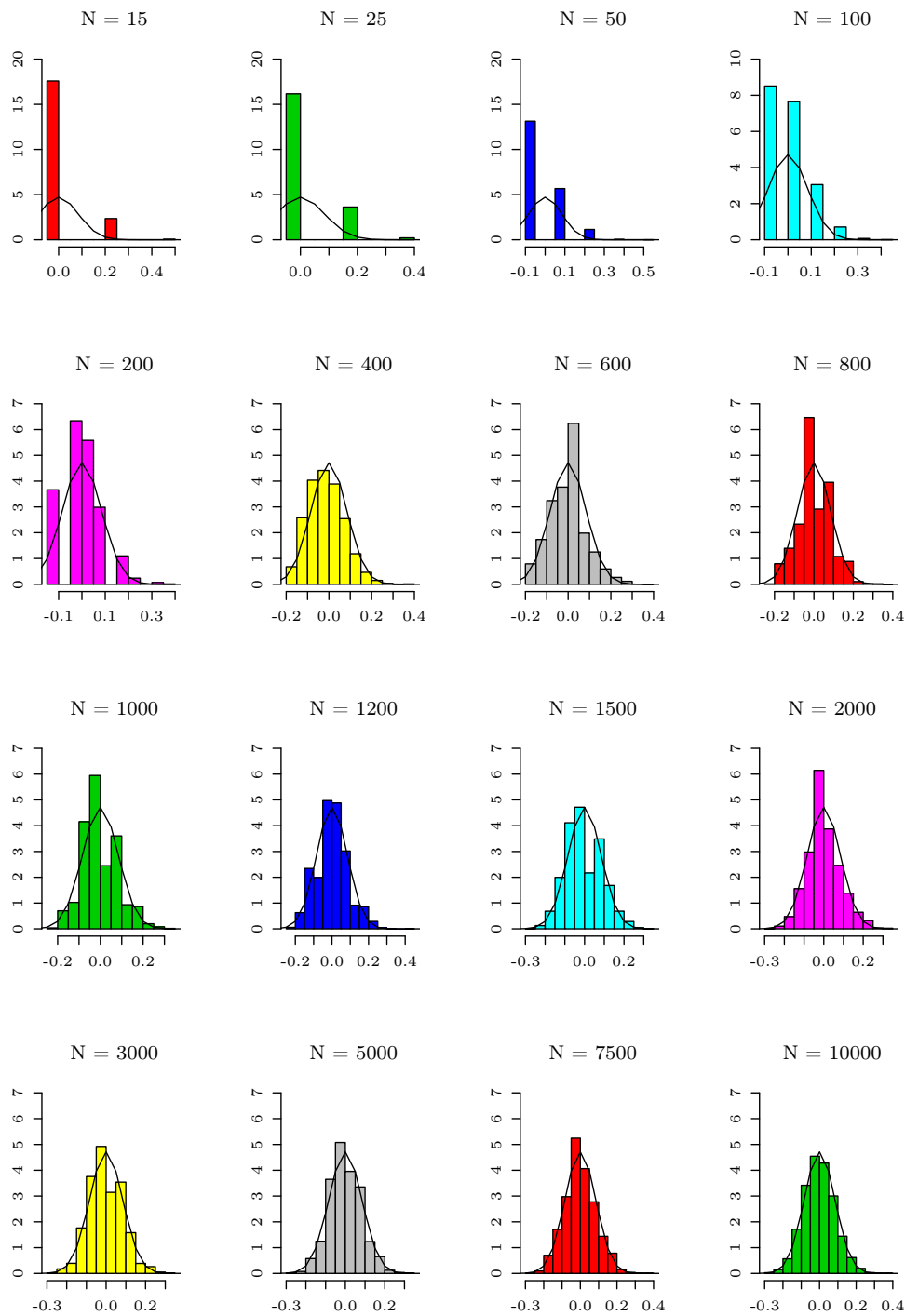
Obrázok 47: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(10)$ pre pevné $p = \frac{1}{4}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



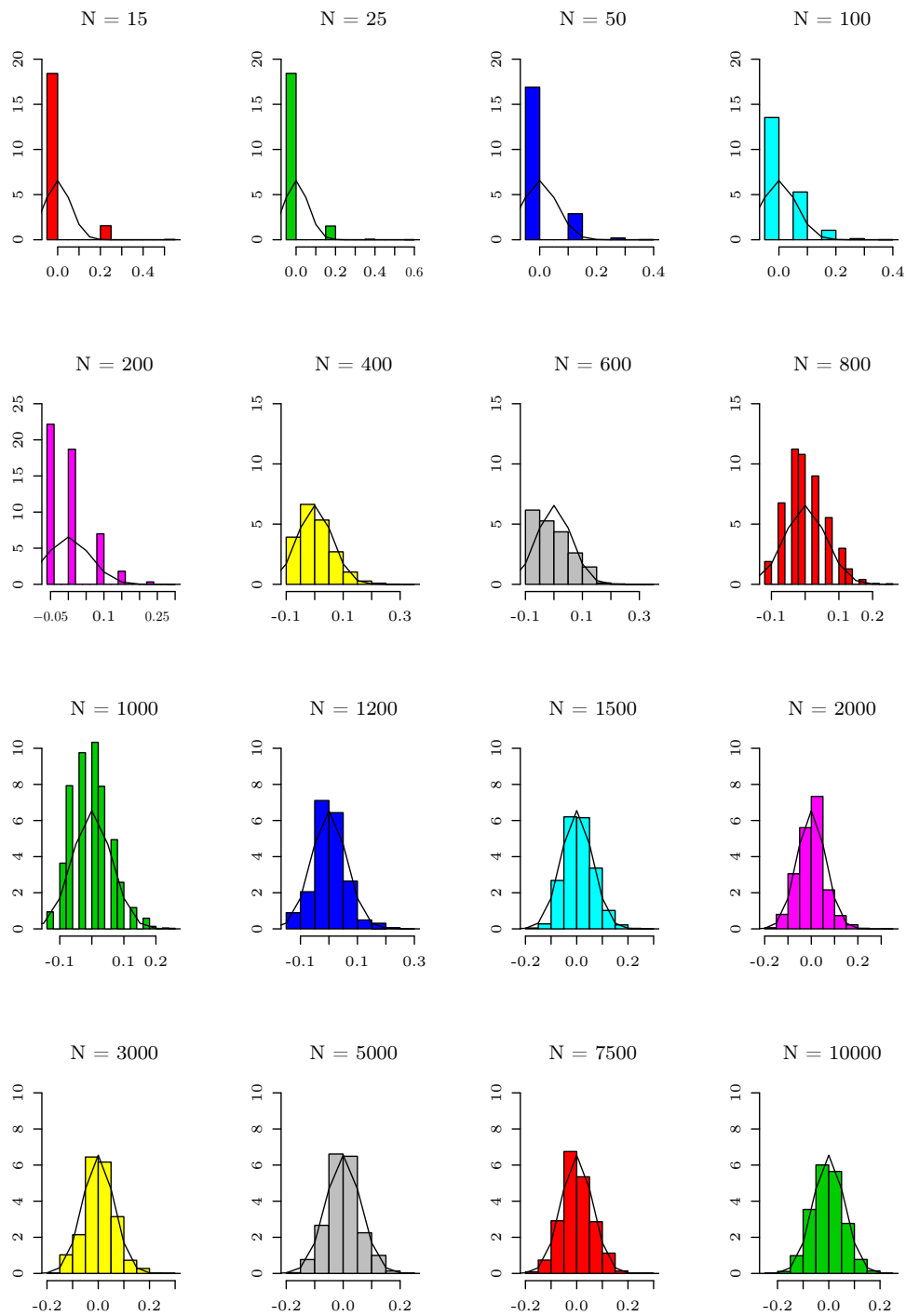
Obrázok 48: Histogramy ilustrujúce rýchlosť konvergencie k Poissonovmu rozdeleniu $Po(10)$ pre pevné $p = \frac{1}{2}$ a pre rastúce hodnoty N volené tak, aby k nadobúdalo hodnoty blízke celočíselným hodnotám.



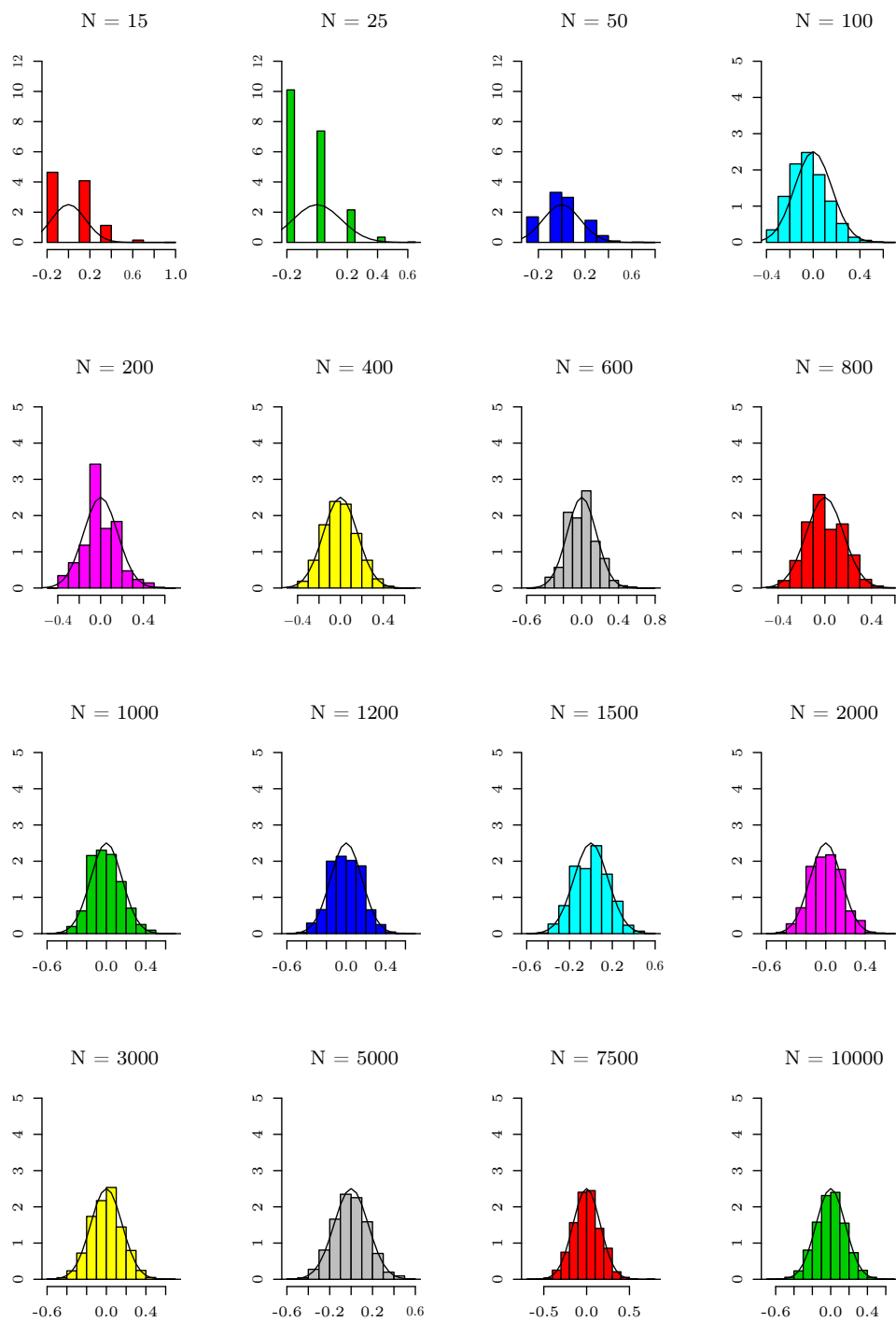
Obrázok 49: Histogramy ilustrujúce rýchlosť konvergencie k normálnemu rozdeleniu $N(0, \rho_2)$ pre pevné $k = 2$, pre pevné $p = \frac{1}{10}$ a pre rastúce hodnoty N .



Obrázok 50: Histogramy ilustrujúce rýchlosť konvergencie k normálnemu rozdeleniu $N(0, \rho_3)$ pre pevné $k = 3$, pre pevné $p = \frac{1}{5}$ a pre rastúce hodnoty N .



Obrázok 51: Histogramy ilustrujúce rýchlosť konvergencie k normálnemu rozdeleniu $N(0, \rho_4)$ pre pevné $k = 4$, pre pevné $p = \frac{1}{4}$ a pre rastúce hodnoty N .



Obrázok 52: Histogramy ilustrujúce rýchlosť konvergencie k normálnemu rozdeleniu $N(0, \rho_5)$ pre pevné $k = 5$, pre pevné $p = \frac{1}{2}$ a pre rastúce hodnoty N .